BIJAN VASIGH
KEN FLEMING
THOMAS TACKER

ASHGATE

Introduction to Air Transport Economics

From Theory to Applications

Introduction to Air Transport Economics

Introduction to Air Transport Economics: From Theory to Applications uniquely merges the institutional and technical aspects of the aviation industry with their theoretical economic underpinnings. In one comprehensive textbook it applies economic theory to all aspects of the aviation industry, bringing together the numerous and informative articles and institutional developments that have characterized the field of airline economics in the last two decades as well as adding a number of areas original to an aviation text. Its integrative approach offers a fresh point of view that will find favor with many students of aviation.

The book offers a self-contained theory and applications-oriented text for any individual intent on entering the aviation industry as a practicing professional in the management area. It will be of greatest relevance to undergraduate and graduate students interested in obtaining a more complete understanding of the economics of the aviation industry. It will also appeal to many professionals who seek an accessible and practical explanation of the underlying economic forces that shape the industry.

'Written in easy-to-read, jargon-free English.... The book will be beneficial to those seeking a deeper understanding of the workings of the major players in the air transport industry, including airlines, airports, and aircraft manufacturers. Readers will learn, for example, how airlines use revenue management to maximize their profits or how airlines and airports forecast passenger traffic.... It can be read as a standalone text by undergraduates and graduate students with limited backgrounds in economics.'

Martin Dresner, Robert H. Smith School of Business, University of Maryland, USA

'This is a rare book that combines both theory and applications in a very meaningful way. As such, I recommend it very strongly to graduate students, policy makers, managers and researchers in aviation.'

Tae H. Oum, Air Transport Research Society (ATRS) and University of British Columbia, Canada

'...provides a look at the industry from the keen perspective of a distinguished economist.... Even in areas such as safety and security, the authors provide clear insight to show how it is economics that drives the successful airlines to adhere to the highest standards of safety and security. I believe that this book will be used by faculties and students worldwide in their effort to understand, and then manage, the activities of the aviation industry.'

Irwin Price, Dubai Aerospace Enterprise University

'This is a landmark textbook in airline economics as, for the first time, disparate information from various economic disciplines is integrated with key institutional factors that affect the industry.... presented in an easy-to-read and understand format with plenty of real-world situations. Much of the material in this book cannot be found in other sources. I highly recommend this book.'

Darryl Jenkins, Ohio State Airline Operations Center, USA

'...offers an overview into the dynamic industry environment while using practical industry applications to explain general economic concepts—a great introduction to the world of aviation and explanation for what keeps airlines in the air.'

Zane Rowe, Senior Vice President Network Strategy, Continental Airlines

'....a detailed introduction to the subject of airline economics, covering both the theory and its application to the modern air transportation industry. Students and those already engaged in the industry will find this an illuminating guide to an ever more complex business.'

Barry Humphreys, Director of External Affairs and Route Development, Virgin Atlantic Airways

ASHGATE

Ashgate Publishing Limited
Gower House, Croft Road
Aldershot, Hampshire
GU11 3HR, England

www.ashgate.com

ISBN 978-0-7546-7081-0

9 780754 670810

Published by
Ashgate Publishing Limited
Gower House
Croft Road
Aldershot
Hampshire GU11 3HR
England

Ashgate Publishing Company
Suite 420
101 Cherry Street
Burlington, VT 05401-4405
USA

Ashgate website: http://www.ashgate.com

British Library Cataloguing in Publication Data
Vasigh, Bijan
Introduction to air transport economics : from theory to applications
1. Aeronautics, Commercial – United States – Finance
2. Airlines – Deregulation – United States 3. Airlines – United States – Cost of operation 4. Aeronautics and state – United States
I. Title II. Tacker, Tom III. Fleming, Ken
387.7'1

ISBN: 9780754670797

Library of Congress Cataloging-in-Publication
Vasigh, Bijan.
Introduction to air transport economics : from theory to applications by Bijan Vasigh, Tom Tacker, and Ken Fleming.
p. cm.
ISBN 978-0-7546-7079-7 -- ISBN 978-0-7546-7081-0 (pbk)
1. Aeronautics, Commercial--United States--Finance. 2. Airlines--Deregulation--United States.
3. Airlines--United States--Cost of operation. 4. Aeronautics and state--United States. I. Tacker, Tom. II. Fleming, Ken. III. Title.

HE9803.A5V37 2008
387.7'10973--dc22

2008003784

Printed and bound in Great Britain by TJ International Ltd, Padstow, Cornwall.

Contents

List of Figures

List of Tables

Acknowledgements

The development and writing of any text book is a long and trying experience—as all aspiring authors know! Our effort has been no exception. And, as with any other text, we have received a great deal of assistance and encouragement from a number of individuals whose efforts we would like to acknowledge.

Many ERAU graduate students helped in the preparation of this manuscript. We owe a special debt to the graduate assistants: Liam Mackay, Christian Vogel, and Duane Miner who participated in data collection, editing graphs, tables and figures. This necessary preparation was often difficult and time-consuming work. Other student assistants who were very helpful in preparing this text were: Megan Praschak, Sean Ross, Charles Nipper, Brian Merzlak, and Andrew Taylor. We would also like to thank the following members of Thomas Tacker's class, who were the first to use the text and made many helpful suggestions: Lavnish Budhiraja, Chih-Han Chang, Constantin Corneanu, Tom Derosiers, Thomas Hilgers, Kevin Hyatt, Virginie Janneteau, Ryan Johnson, Mahmoud Khatib, Niles Klungboonkrong, Brian Lech,Vaida Maleskis, Mary Mankbadi, Nathan Monroe, Jae-Woo Park, Rushi Patel, Naef Saab, Worachat Sattayalekha, Paul Smith, Heli Trivedi, and Odartey Williams.

Ashgate Publishing has done an admirable job in being patient with the text and turning the raw manuscript into an actual textbook.

Thanks also go to many other people who have helped with the book, especially Hossein Tavana and Daniel Petree. Many individuals have read all or part of the manuscript and made helpful suggestions. Among these are Tae Oum, Irwin Price, Martin Dresner, Zane Rowe, Barry Humphreys, Darryl Jenkins, and Kevin Cox.

List of Abbreviations

ACI	Airports Council International
ADIZ	Air Defense Identification Zone
ADS-B	Automatic Dependent Surveillance Broadcast
AFC	average fixed cost
AFTA	ASEAN Free Trade Area
ANOVA	analysis of variance
ANSP	air navigation service provider
AOPA	Aircraft Owners and Pilots Association
APEC	Asia-Pacific Economic Cooperation
ASEAN	Association of Southeast Asian Nations
ASM	available seat mile
ATA	Airport Transport Association
ATC	air traffic control/average total cost
ATI	Air Transport Intelligence
ATM	available ton mile
AU	African Union
AVC	average variable cost
AVOD	audio-visual on-demand
BAA	British Airports Authority
BEA	Bureau of Economic Analysis
BLS	Bureau of Labor Statistics
BTS	Bureau of Transportation Statistics
CAA	Civil Aeronautics Administration/UK Civil Aviation Authority
CAB	Civil Aeronautics Board
CACM	Central American Common Market
CANSO	Civil Air Navigation Services Organization
CARICOCM	Caribbean Community
CASM	cost per available seat mile
CEMAC	Economic and Monetary Community of Central Africa
CFIT	controlled flight into terrain
CIS	Commonwealth of Independent States
COMESA	Common Market for Eastern and Southern Africa
CSME	CARICO Single Market and Economy
DOC	direct operating costs
DOT	Department of Transportation
DR-CAFTA	Dominican Republic-Central America Free Trade Agreement
DRVSM	Domestic Reduced Vertical Separation Minimum

EAC	East African Community
ECJ	European Court of Justice
EBIT	earnings before interest and taxes
EBITDA	earnings before interest, tax, depreciation, amortization
EEC	European Economic Community
EMSR	expected marginal seat revenue
ETOPS	extended-range twin-engine operations
EU	European Union
FAA	Federal Aviation Administration
FC	fixed cost
FTA	US–Australia Free Trade Agreement
FTAA	Free Trade Area of the Americas
GAFTA	Greater Arab Free Trade Area
GAO	Government Accountability Office
GDP	gross domestic product
GDS	global distribution systems
GECAS	GE Commercial Aircraft Services
GNP	gross national product
HHI	Herfindahl Hirschman Index
IAE	International Aero Engines
IATA	International Air Transport Association
ICAO	International Civil Aviation Organization
FR	instrument flight rules
IMF	International Monetary Fund
JAA	Joint Aviation Authority
JFTA	US–Jordan Free Trade Agreement
LCC	low-cost carrier
LHR	London Heathrow
LOC	loss of control
MAD	mean absolute deviation
MC	marginal cost
MLB	Major Baseball League
MSE	mean squared error
NAFTA	North American Free Trade Agreement
NBA	National Basketball Association
NBFU	National Board of Fire Underwriters
NCAA	National Collegiate Athletic Association
NFL	National Football League
NGO	non-governmental organization
NMAC	near mid-air collision
NORAD	North American Aerospace Defense Command
NTSB	National Transportation Safety Board
OAG	Official Airline Guide
O&D	Origin and Destination
OECD	Organization for Economic Cooperation and Development
OPEC	Organization of Petroleum Exporting Countries
PARTA	Pacific Regional Trade Agreement
PPC	production possibility curve

RASM	revenue air seat mile
RPK	revenue passenger kilometre
RPM	revenue passenger mile
RRPM	revenue per revenue passenger mile
RTA	regional trade agreement
RTM	revenue ton mile
SAARC	South Asian Association for Regional Cooperation
SACU	South African Customs Union
SAFTA	South Asian Free Trade Area
SEC	Securities Exchange Commission
TC	total cost/Transport Canada
TCAS	Traffic Collision Avoidance System
TP SEP	Trans-Pacific Strategic Economic Partnership
UEMO	West African and Monetary Union
UNASUL	Union of South American Nations
VC	variable cost
VFR	visual flight rules
VIF	variance inflation factors
WMA	weighted moving average
WTO	World Trade Organization

Preface

While it is undoubtedly true that there are many books that cover different aspects of various economic problems, it is also true that there comes a time when there is a requirement for a textbook that brings together the disparate elements of analysis that are covered in other separate areas. So for example, there are textbooks on labor economics, monetary economics, international economics, comparative economic systems, industrial organization, and numerous other specialized fields. It is our conviction that the time has come to bring together the numerous and informative articles and institutional developments that have characterized the field of airline economics in the previous two decades. While some might argue that this is too specialized an area in economics, we would contend that the unique nature of the economics of this industry make it particularly appropriate for a separate text. We would suggest that these unique features are: the perishable nature of the product and the consequent elasticity of demand and pricing complications; the control of the method of delivering the service by a disinterested third party (namely, air traffic control); the presence of only two major suppliers of the means of providing the service; the unique dominance of this form of transportation for long-haul passenger traffic; the interesting and complicated financial arrangements that are used to provide the service; the existence of quasi-monopolistic entities to jointly deliver the service (airports); and last, but by no means least, the international legal aspects of the industry. All these areas are covered in one place or another, but there is as yet no single text or article that brings them together in such a way that the critically important underlying economics of the industry is made clear to the interested reader.

The underlying foundation of this book is the idea that the reader should be introduced to the economic way of thinking and approaching problems in aviation rather than the more traditional institutional and governmental regulatory approach. In the early chapters the reader is introduced to the elementary ideas of demand and supply and market equilibrium. This is followed by an in-depth presentation of costs and their key applicability to managerial decision-making. The basic economic principles are then applied to a unique analysis of the effect of air traffic control and the governmental ownership of airports on the industry. Following this, there is a thorough discussion of market structures and how they affect the industry. In particular, this section introduces the idea of contestability theory which appears to be particularly applicable to this industry. The international aspects of the industry and global alliances are then discussed in detail.

The final chapters are devoted to what might be called applications of the earlier theoretical chapters. There is an elementary overview chapter on the various types of forecasting that are prevalent in the industry. The next chapter ties together the basic principles of demand (that were covered earlier) in a somewhat more sophisticated presentation of the critically important topic of revenue management. Clear numerical examples are presented, tying this mainstay of the industry to the theoretical idea of elasticity of demand. This discussion is followed by another unique chapter that is entirely devoted to the phenomenon of so-called low-cost carriers. Finally, the book presents a decidedly non-conventional approach to the controversial topic of safety within the industry. That is, rather than the conventional safety at any cost approach (which is in reality not followed anyway), the text adopts a more balanced cost–benefit approach to this important topic.

As economists in a university that specializes in the aviation industry, our preferred approach is to apply economic principles to the industry, and this is the area where we see the unique need. Therefore, we feel that this book will be the first to bring all of these areas together under one cover. In summary, and as discussed above, our approach follows a more or less standardized format. That is, we first present the necessary economic principles that will be used to analyze the industry. We follow this with a discussion of institutional arrangements, particularly in the international area, that make the aviation industry a truly global enterprise. Finally, the closing chapters are devoted to practical applications and comparisons within the industry. It is our hope that the text will appeal to interested readers within the industry, as well as students who intend to enter the industry.

1

The Evolving Air Transport Industry

> If you want to be a millionaire, start with a billion dollars and open an airline. Soon enough you will be a millionaire.
>
> Sir Richard Branson, Founder, Virgin Atlantic Airways

As the comment above implies, in the last 30 years the airline industry's earnings have fluctuated wildly (mostly downward). New carriers such as JetBlue and AirTran in the US, easyJet and Ryanair in Europe, Gol and Volaris in Latin America and a few others have entered the industry, but many others such as Eastern, Pan Am, and Midway have declared bankruptcy and ceased operation. The purpose of this chapter is to describe the evolution of air transport industry, including airlines and airports.[1] The topics include the following:

- The airline industry
- The financial condition of the airline industry
- Airline industry consolidation
- Factors affecting world air traffic growth
- The economic impact of the air transport industry
- The outlook for the air transport industry.

THE AIRLINE INDUSTRY

> People Express is clearly the archetypical deregulation success story and the most spectacular of my babies. It is the case that makes me the proudest.
>
> Alfred Kahn, Professor of Political Economy, Cornell University, 1986

Since the US Airline Deregulation Act of 1978, the US airline industry (and, to a certain extent, the global airline industry) has been characterized by volatility. Periods of high revenues are followed by periods of economic drought. The most recent economic

1 In the United States, more than 164 airlines filed for bankruptcy since deregulation in 1978 and many of these did not survive the bankruptcy proceedings.

"trough" followed the 11 September terrorist attacks in 2001. This volatility produces airline bankruptcies, extensive layoffs or employee pay cuts, loss of shareholder wealth, and great uncertainty in the market. Prior to deregulation, the airline industry was relatively stable with minimal losses and healthy profits; however, it was also clear that this state of affairs was due mainly to government regulation that virtually eliminated any meaningful competition between airlines and certainly prevented new competitors from entering the market. The biggest loser in all of this was of course, the passenger, who had to pay ticket prices that were set to cover average airline costs with no competitive discounts permitted. Therefore, while deregulation may seemingly have caused huge financial losses, it also reflected the fact that airlines were faced with their first bout of meaningful competition and some did not measure up. On the other hand, deregulation also opened up the opportunity for some airlines, such as Southwest Airlines and Ryanair to post some of the greatest profits in the history of the industry. Figure 1.1 graphically displays this trend for the global airline industry overall.

The major reason why the deregulation of the US airline industry had such a large impact on the global airline industry is that the North American airline industry has been the most dominant aviation industry in the world. As Table 1.1 shows, the North American market currently holds the distinction of being the largest market in terms of both total passengers and aircraft movements with percentages of 36.1 per cent and 49.6 per cent, respectively. Europe, with a 30.9 per cent share of total passengers, has narrowed the gap so that the US proportion of world passengers is now much smaller than it was in the 1980s. Other aviation markets, in addition to Europe, have also emerged—in particular, the airline industries of Asia-Pacific and the Middle East.[2] This growth has been a direct result of the economic expansion in those regions, particularly China and India. Furthermore,

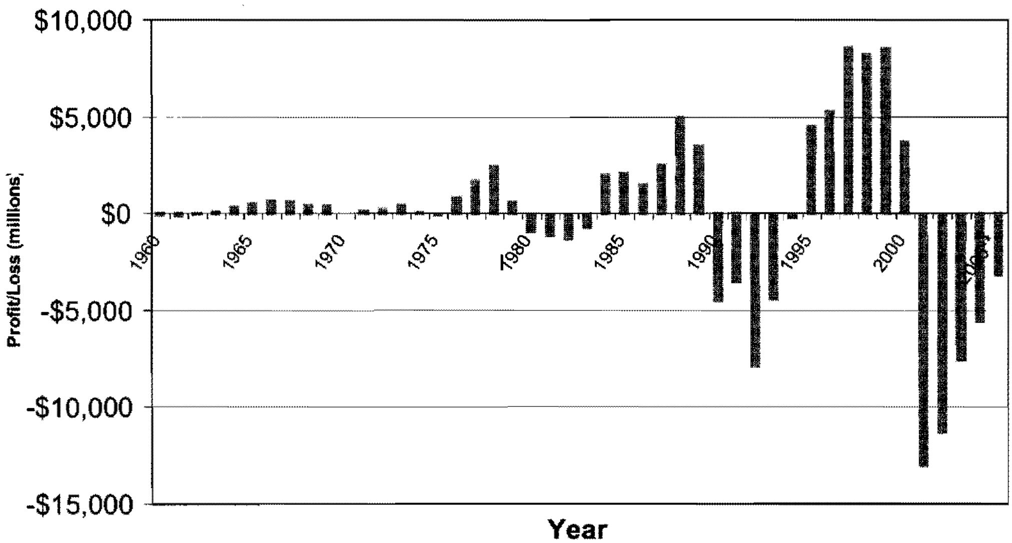

Figure 1.1 World airline operating profits

Source: Compiled by the authors.

2 In 2004 Boeing projected that by 2023 China will require nearly 2,300 new airplanes, making China the largest commercial aviation market outside the United States.

Table 1.1 World airlines passenger and cargo traffic

Region	Passengers	Passengers (%)	YTD Growth
Africa	109,355,929	2.7	6.6
Asia-Pacific	890,903,969	21.8	9.1
Europe	1,261,931,534	30.9	7.2
Latin America	248,393,399	6.1	8.8
Middle East	97,593,687	2.4	10.7
North America	1,476,311,133	36.1	1.8
Total	4,084,489,651	100.00%	5.7%
Region	**Cargo**	**Cargo (%)**	**YTD Growth**
Africa	1,437,612	1.9	10.3
Asia-Pacific	25,868,817	33.6	6.9
Europe	15,795,039	20.5	2.6
Latin America	3,564,103	4.6	0.4
Middle East	3,455,683	4.5	8.7
North America	26,857,094	34.9	-0.3
Total	76,978,348	100.0	3.2
Region	**Movements**	**Movements (%)**	**YTD Growth**
Africa	1,913,945	2.9	1.7
Asia-Pacific	8,290,312	12.5	6.9
Europe	17,787,273	26.8	3.3
Latin America	4,454,798	6.7	2.9
Middle East	973,146	1.5	7.6
North America	32,837,615	49.6	-1.5
Total	66,257,089	100.0	1.3

Source: Airports Council International.

Note: Data is based on the 12 months preceding and including May 2006.

the gap will continue to narrow on the basis of the impressive growth rates of 10.7 per cent and 9.1 per cent in the Middle East and Asia-Pacific regions (experienced from May 2006). Airbus now forecasts that both Asia-Pacific and Europe will surpass North America in terms of seats being flown (Airbus, 2006).

It is interesting to note that in terms of international passengers, both the Asia-Pacific and Europe markets already surpass the United States (Table 1.2). Europe's dominance in international air transportation is mainly a result of its historical ties to former colonial countries and its relatively small geographical area. Because of this small area and significant government support for other surface transportation (mainly railroads), most

Table 1.2 International passenger traffic and traffic growth

	Passengers	Passengers (%)	YTD Growth
Africa	65,901,046	4.0	6.1
Asia-Pacific	337,953,348	20.7	9.3
Europe	889,013,832	54.3	7.3
Latin America	85,672,788	5.2	5.2
Middle East	77,156,487	4.7	12.3
North America	180,340,555	11.0	3.9
Total	1,636,038,056	100.0	7.4
	Cargo	**Cargo (%)**	**YTD Growth**
Africa	1,042,120	2.4	13.1
Asia-Pacific	19,129,620	43.5	7.2
Europe	11,685,033	26.5	3.0
Latin America	2,065,905	4.7	-0.9
Middle East	3,255,882	7.4	9.0
North America	6,843,559	15.5	0.4
Total	44,022,119	100.0	4.8

Source: Airports Council International.

Note: Data is based on the 12 months preceding and including May 2006.

of the domestic aviation industries in Europe are relatively small; therefore, European airlines survive on international travel. Economic growth in Eastern Europe will also undoubtedly increase that region's share of international travel. Thus, the growth in international traffic is certainly a dominant trend in the air transportation industry.

Another trend in the air transportation industry is the growth and expansion of the cargo industry. Tables 1.1 and 1.2 quote cargo (in metric tons) for each region and for international cargo. The Asia-Pacific region is the dominant region for cargo (especially international cargo) and its year-by-year growth rates are quite high. Much of this growth is a result of China's burgeoning economy and the large and growing amount of exports that come from the region. North America's share of total cargo was still the highest as of May 2006, but the negative growth rate indicates that the industry has probably reached a plateau. According to IATA predictions, intra-Asia freight will be 8.3 million tons by 2010; 26 per cent of this total will be international freight and 30 per cent of the growth in freight traffic will come from China.[3]

Another way of looking at the global distribution of air transportation is the number of in-service aircraft. Table 1.3 divides aircraft into three categories based on the aircraft's age: new, middle, and old. While the North American market contains the highest number of commercial aircraft, it also contains a relatively high percentage of older aircraft (24 per cent). For aircraft manufacturers, this is attractive as many of these older

3 Cargo Symposium 2007, Mexico: Remarks by Giovanni Bisignani.

Table 1.3 Aircraft in service by region (2004)

	New	Mid	Old	Total
North America	1,654	2,581	1,301	5,536
Europe	1,768	1,363	237	3,368
Asia	1,154	969	295	2,418
Latin America and Caribbean	238	259	393	890
Africa	162	111	316	589
Middle East	240	144	155	539
Pacific	155	102	15	272
World	**5,371**	**5,529**	**2,712**	**13,612**

Source: Compiled by the authors using Airbus *Global Market Forecast*.
Includes: Western-built passenger jets (greater than 100 seats) and freighters.

aircraft will have to be replaced in the coming years. Europe, on the other hand, does not have very many older aircraft (7 per cent), probably as a result of the stringent noise regulations implemented by the European Union that effectively banned many older-generation aircraft. Finally, some correlation can be made between the number of new aircraft and a region's economic growth. For example, both Africa and Latin America have high percentages of older aircraft (54 per cent and 44 per cent respectively), while more prosperous regions such as Europe and Asia have low ratios (7 per cent and 12 per cent respectively). This means that North America has now become an attractive market to aircraft manufacturers (because of replacement requirements), whereas regions such as Latin America and Africa are less as attractive since airlines in those regions continue to renew their fleets with secondhand aircraft.

A final way of analyzing the composition of the air transport industry is to analyze traffic data on an airport-by-airport basis. Tables 1.4 and 1.5 provide a list of the top 15 airports in terms of total passengers, total international passengers, and total cargo volume. The rankings of the top airports mirror the distribution of passengers by regions as evidenced by the two largest airports in terms of passengers. Since North America is the largest market in terms of passengers, it is not surprising that Atlanta and Chicago are the top two airports. Conversely, since Europe is the top region for international passengers, it should come as no surprise that the top four airports, in terms of international passengers, are in Europe. Also, many of the airports on the international passenger traffic list are airports located in countries that have small or non-existent domestic air travel markets. The impact of the growth of China's economy is clearly shown through the 18.5 per cent growth in total passengers at Beijing. Moreover, the development of Dubai and Emirates Airlines as an international hub is reflected in Dubai's 17.4 per cent increase in international passenger traffic.

The division of airports by region in terms of cargo volume is not so clear-cut, but both North America and Asia-Pacific airports are well represented in the top 15. Memphis is the busiest cargo airport in the world, and this is a direct reflection of the fact that FedEx has its main hub there. A similar situation occurs for Louisville with UPS. In the Asia-Pacific region, the large volume of export trade has spurred cargo growth, especially in Hong Kong and Shanghai.

Table 1.4 Total passenger traffic

Total Passenger Traffic				Total International Passenger Traffic			
Rank	Airport	Passengers	% Change	Rank	Airport	Passengers	% Change
1	Atlanta (ATL)	84,166,330	(1.5)	1	London (LHR)	61,273,527	0.8
2	Chicago (ORD)	76,934,280	0.7	2	Paris (CDG)	50,068,012	6.2
3	London (LHR)	67,955,821	0.0	3	Frankfurt (FRA)	45,045,265	2.6
4	Tokyo (HND)	63,917,082	2.2	4	Amsterdam (AMS)	44,611,014	4.1
5	Los Angeles (LAX)	61,491,371	0.3	5	Hong Kong (HKG)	41,323,000	8.3
6	Dallas/Ft. Worth (DFW)	60,127,713	1.7	6	Singapore (SIN)	31,937,617	8.5
7	Paris (CDG)	55,008,132	5.6	7	London (LGW)	29,272,024	3.7
8	Frankfurt (FRA)	52,346,231	1.9	8	Bangkok (BKK)	28,691,901	10
9	Denver (DEN)	45,397,350	6.3	9	Tokyo (NRT)	27,217,860	0.1
10	Las Vegas (LAS)	44,960,060	5.6	10	Seoul (ICN)	26,223,167	5.7
11	Amsterdam (AMS)	44,773,661	4.0	11	Dubai (DXB)	25,490,762	17.4
12	Beijing (PEK)	44,221,007	18.5	12	Madrid (MAD)	23,189,627	10.3
13	Madrid (MAD)	43,318,353	8.5	13	Munich (MUC)	20,150,718	8.4
14	Hong Kong (HKG)	41,826,000	8.4	14	London (STN)	20,004,197	6.9
15	Phoenix (PHX)	41,635,847	3.1	15	Taipei (TPE)	19,740,019	6.7

Source: Compiled by the authors using Airports Council International data.
Note: Data is based on the 12 months preceding and including May 2006.

Table 1.5 Cargo traffic at major international airports

Rank	Airport	Code	Cargo Tonnes	% Change
1	Memphis	MEM	3,626,425	2.6
2	Hong Kong	HKG	3,523,823	8.8
3	Anchorage	ANC	2,606,108	3.9
4	Tokyo	NRT	2,295,848	(1.8)
5	Seoul	INC	2,205,618	2.6
6	Frankfurt	FRA	2,042,157	8.3
7	Shanghai	PVG	1,968,888	14.0
8	Los Angeles	LAX	1,920,621	0.1
9	Singapore	SIN	1,910,182	5.6
10	Louisville	SDF	1,883,812	8.8
11	Paris	CDG	1,804,300	7.4
12	Miami	MIA	1,766,705	(2.3)
13	Taipei	TPE	1,711,935	0.1
14	New York	JFK	1,641,465	(3.5)
15	Chicago	ORD	1,598,426	2.7

Source: Compiled by the authors using Airports Council International data.
Note: Data is based on the 12 months preceding and including May 2006.

THE FINANCIAL CONDITION OF THE AIRLINE INDUSTRY

> I don't think JetBlue has a better chance of being profitable than 100 other predecessors with new airplanes, new employees, low fares, all touchy-feely ... all of them are losers. Most of these guys are smoking ragweed.
>
> Gordon Bethune, CEO, Continental Airlines, June 2002

The US Airline Deregulation Act of 1978 dramatically changed the global financial condition of the airline industry as other countries began to follow suit and deregulate their own industries. Also, as mentioned earlier, prior to 1978, the industry was relatively stable, chiefly as a result of the government's enforcement of non-competitive regulation and pricing. During the post-deregulation era the industry took on the more cyclical nature of a competitive industry, in which periods of robust financial profitability could be followed by periods of severe economic distress. As in other competitive industries, the financial condition of the airline industry is highly related to economic growth, so it is not surprising that it suffered when the economy stalled.

In the early 1980s, shortly after US deregulation, the airline industry suffered a minor crisis as the economy slowed and competition soared. More specifically, the US domestic industry experienced overcapacity as the many new airlines that were formed as a result of deregulation either went bankrupt or merged with other carriers. The result was four years of global net losses for the industry, largely based on the situation in the United States. A similar situation occurred in the early 1990s as the economy once again experienced a downturn, but this downturn was aggravated by political uncertainty from the first Gulf War and increased fuel costs.

While the early parts of each decade following deregulation have proved to be troublesome for the airline industry, the industry recovered sufficiently to post record short-run profits in the late 1980s and again in the late 1990s. This was partly due to the overall improvement in the global economy, but financial distress and competition also caused airlines to be more innovative and conscious of controlling costs. Tools such as revenue management and frequent-flyer programs were created and developed during these periods. In addition, technological innovations allowed the airlines to improve their profit margins. For example, simpler cockpit design has enabled the reduction in the number of flight-crew members, better engine design has reduced the number of engines required to fly long distances, and fuel costs have been reduced as a result of more fuel-efficient engines. All these technologies have enabled airlines to reduce their costs and/or increase revenue. A more recent technological innovation has been e-ticketing, which allows airlines to reduce their ticket distribution costs.

The post-deregulation airline profitability cycle continued into the twenty-first century, with the global industry experiencing its worst downturn in the history of commercial aviation. Although the 9/11 terrorist attacks were the proximate cause of the global airline industry's financial problems, the root cause was a slowing economy that reduced passenger yields. Added to this were rising jet fuel costs, increased airline operating costs stemming from overcapacity in domestic markets, and increased security costs at commercial airports. In fact, the airline industry was in trouble before the 9/11 disaster, with many airlines losing money and with no significant initiatives to reduce costs and increase productivity. The outcome was record net losses of $13 billion and $11 billion for the airline industry in 2001 and 2002 respectively. With the resurgence of economic growth, the global airline industry returned to profitability in 2004. Table 1.6 graphically portrays these losses for the years in question.

However, political instability in various parts of the world, rising fuel prices, and persistent competition between network and low-cost carriers has meant that the road to recovery has been slow for the airline industry. This situation has been most evident in the North American market where the high-profile bankruptcies of US Airways, United, Delta, and Northwest highlighted the increasing effects of fierce competition from lower-cost airlines and the bloated cost structures of the more traditional airlines. Globally, we see some different patterns. Table 1.7 presents the international air transport market's net profits based on the different geographical regions. The Asian-Pacific and Europe enjoy the highest level of profitability followed by Latin America. In 2006, the US airlines returned to profitability after many years of losses.

There have been some bright spots in the industry as bankruptcy protection has (in some cases) enabled carriers to restructure their costs and receive wage concessions from labor groups. Moreover, the overcapacity issue has been addressed, with carriers not only reducing capacity as a whole, but also shifting capacity to international markets that are less competitive. Only recently, in 2006, has capacity reached pre-2001 levels. Several low-cost carriers have also remained successful and profitable by continuing to expand while

Table 1.6 Schedule airlines' financial performance

(ICAO Contracting States)								
	1996	**1998**	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**
Revenues, $ billion								
Scheduled services	247	258	285	268	266	279	328	360
Non-scheduled flights	12	10	12	10	10	10	12	10
Incidental revenues	24	28	32	29	30	33	39	43
Total Revenues	283	296	329	308	306	322	379	413
Expenses, $ billion								
Flight operations	75	75	99	97	96	103	126	154
Maintenance and overhaul	29	31	34	36	35	35	39	42
Depreciation and amortization	19	18	21	23	22	21	24	25
User charges and station expenses	48	50	55	54	53	55	62	66
Passenger services	29	30	32	33	32	32	36	38
Ticketing, sales and promotion	41	40	40	36	33	32	36	37
General and administrative	17	18	20	23	22	23	27	25
Other operating expenses	13	17	18	18	18	22	24	22
Total Operating Expense	270	280	318	319	311	323	376	409
Net Operating Result	12.3	15.9	10.7	(11.8)	(4.9)	(1.5)	3.3	4.3
% margin	4.4%	5.4%	3.3%	-3.8%	-1.6%	-0.5%	0.9%	1.0%
Non-Operating Items	(4.5)	(2.9)	(4.3)	(4.8)	(8.7)	(4.6)	(6.4)	(5.6)
Profit/Loss before income taxes	7.9	13.1	6.5	(16.6)	(13.6)	(6.1)	(3.1)	(1.3)
Profit/Loss after income taxes	5.3	8.2	3.7	(13.0)	(11.3)	(7.6)	(5.7)	(4.1)
% margin	1.9%	2.8%	1.1%	-4.2%	-3.7%	-2.3%	-1.5%	-1.0%

Source: Compiled by the authors.

keeping costs relatively constant. Innovations such as e-ticketing and fleet rationalization have been instrumental in helping airlines achieve cost reductions, and these reductions have narrowed the cost gap between network airlines and low-cost carriers (LCCs). Figure 1.2 provides a comparison between the network carriers and low-cost carriers in terms of CASM (cost per available seat mile).

Table 1.7 The financial performance of the airline industry

Industry Net Operating Profit: US$ Billion	2004	2005	2006	2007f
Global*	-5.6	-2.1	8.5	8.8
North America*	-10.0	-5.7	4.4	-0.6
Europe	1.1	1.6	2.6	2.4
Asia-Pacific	3.4	2.1	1.9	1.7
Latin America	0.1	-0.1	0.5	0.5
Africa	-0.3	-0.3	-0.3	-0.03

Source: Compiled by the authors using IATA data, 2007.
* Excluding US restructuring costs.

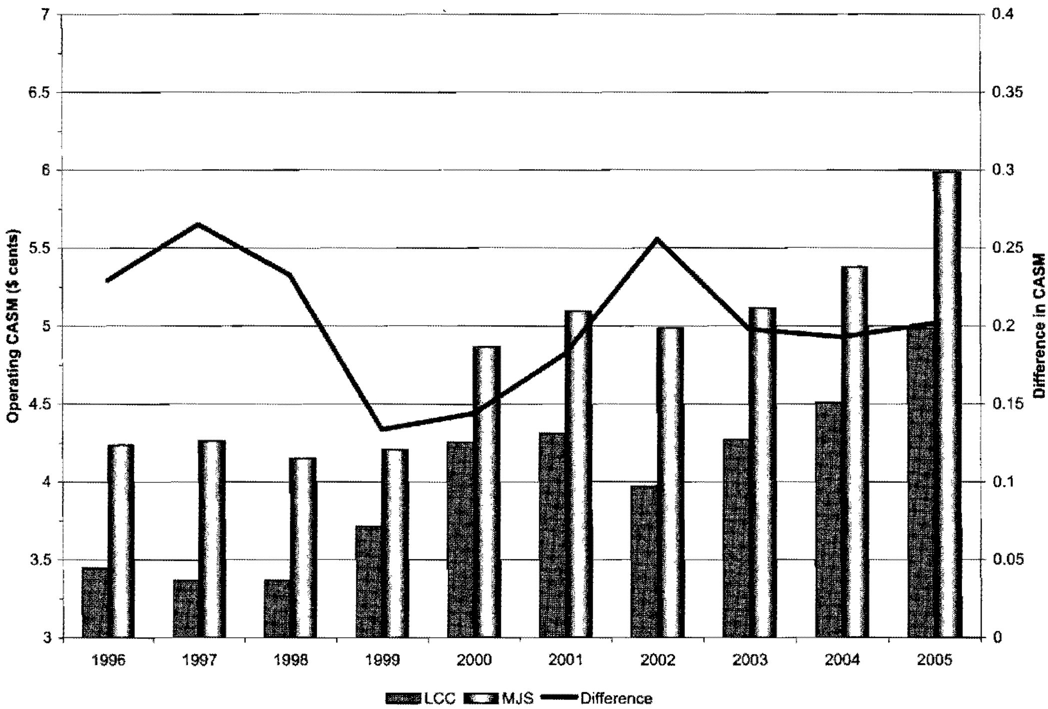

Figure 1.2 Cost comparisons between LCC and legacy airlines

Source: Compiled by the authors based on data taken from ATW World Airline Reports.

As shown by Tables 1.8–1.10, among network carriers, the most profitable airline was cargo giant FedEx. This highlights a trend in the global airline industry where the cargo industry is thriving because of the increased globalization of the business marketplace. Network carriers did well, especially in Europe and Asia-Pacific. Interestingly, Air Canada was the only major North American airline to make the list, while regional

Table 1.8 The top 25 airlines by operating profitability (2004 and 2005)

2004			2005		
Rank	Airline	USD million	Rank	Airline	USD million
1	FedEx Express	1,348	1	FedEx Express	1,596
2	British Airways	967	2	British Airways	1,227
3	Quantas Group	837	3	Air France-KLM	1,130
4	SIA Group	813	4	Quantas Group	855
5	Emirates Group	778	5	Southwest	820
6	Cathay Pacific	673	6	Emirates Group	800
7	Air France-KLM	664	7	ANA	755
8	ANA	662	8	SIA Group	749
9	Air China	556	9	Lufthansa Group	683
10	Southwest	554	10	Cathay Pacific	531
11	Thai	500	11	Air China	455
12	JAL Group	477	12	Ryanair	446
13	Lufthansa Group	454	13	Korean	427
14	Korean	433	14	Air Canada (ACE)	388
15	Ryanair	397	15	UPS Air	294
16	Gol	248	16	Gol	265
17	UPS Air	224	17	Thai	264
18	American Eagle	219	18	TAM	261
19	Iberia	215	19	American Eagle	225
20	TAM	212	20	Skywest	220
21	ExpressJet Holdings	205	21	Atlas Air Worldwide Holdings	193
22	LAN	172	22	SAS Group	173
23	Virgin Blue	169	23	Republic Airways Holding	159
24	China Airlines	168	24	ExpressJet Holdings	157
25	Air New Zealand	165	25	Air New Zealand	148

Source: ATW World Airline Report.

airlines like American Eagle and Skywest were profitable through their codeshare agreements with major network carriers. Low-cost carriers, such as Southwest, Gol and Ryanair, had a very successful 2005, with three low-cost carriers, all from different continents, ranked as the highest commercial airlines, based on operating profit margin.

Table 1.9 The top 25 airlines by operating profit margin (2004 and 2005)

2004			2005		
Rank	**Airline**	**%**	**Rank**	**Airline**	**%**
1	Gol	34.0	1	Gol	23.3
2	AirAsia	30.2	2	Custom Air	22.5
3	Copa Holdings	28.3	3	Ryanair	21.8
4	Ryanair	27.1	4	AirAsia	20.4
5	Custom Air	26.8	5	Lynden	18.0
6	Skywest	24.6	6	Republic Airways Holding	17.5
7	Lynden	22.4	7	Southern	17.4
8	AirBalric	22.2	8	Copa Holdings	17.3
9	Omni Air	22.1	9	Omni Air	16.5
10	Kalitta	20.8	10	Tradewinds	16.1
11	Tradewinds	20.3	11	Sierra Pacific	15.3
12	Ameristar	18.0	12	Kenya	14.6
13	Republic Airways Holding	16.8	13	Jazz Air	12.7
14	Kenya	16.6	14	American Eagle	12.5
15	Emirates Group	16.2	15	Astar	12.5
16	American Eagle	16.1	16	Kalitta	12.4
17	Jet	15.8	17	Air Mauritius	12.2
18	Astar	15.1	18	Emirates Group	12.1
19	Virgin Blue	14.8	19	Atlas Air Worldwide Holdings	11.9
20	Air Mauritius	14.0	20	Wideroe	11.8
21	Air China	13.7	21	AirBalric	11.5
22	ExpressJet Holdings	13.6	22	ATI	11.4
23	Thai	13.5	23	Mesa Air Group	11.4
24	Air Greenland	13.3	24	Skywest	11.2
25	Cathay Pacific	12.7	25	Evergreen	10.9

Source: Compiled by the authors from the ATW World Airline Report, 2005.

The global airline industry is well on the road to recovery, with IATA forecasting a global net profit in 2007—the first since 2000. This global profit is largely spurred by a forecast that the North American market will break even. This is critical to global profitability because this market represents roughly 40 per cent of the world's aviation.

Table 1.10 The top 25 airlines by net profit margin (2004 and 2005)

2004			2005		
Rank	**Airline**	**%**	**Rank**	**Airline**	**%**
1	Custom Air	25.0	1	FL Group	36.2
2	AirAsia	24.5	2	Kingfisher	27.1
3	Copa Holdings	23.6	3	Custom Air	21.6
4	Ryanair	23.1	4	Gol	19.2
5	Gol	22.7	5	Ryanair	18.1
6	Omni Air	22.5	6	Omni Air	18.1
7	Lynden	21.7	7	AirAsia	17.8
8	Kalitta	20.8	8	Lynden	17.5
9	Tradewinds	20.1	9	Southern	16.0
10	Ameristar	18.0	10	Copa Holdings	13.6
11	Emirates Group	15.1	11	Kalitta	12.6
12	Skywest	13.9	12	ATI	11.9
13	Express Net	11.7	13	Emirates Group	11.5
14	ATI	11.7	14	Jazz Air	11.5
15	Aeroflot	11.4	15	Tradewinds	9.7
16	SIA Group	11.4	16	Aeroflot	9.6
17	Cathay Pacific	10.7	17	SIA Group	9.3
18	Virgin Blue	10.5	18	Kenya	9.1
19	Binter Canarias	10.4	19	Sierra Pacific	9.1
20	Jet	10.4	20	Miami Air	9.0
21	TAM	10.2	21	Ethiopian	8.8
22	Kenya	9.8	22	Binter Canarias	8.5
23	Trans States	9.2	23	Air Lingus	8.2
24	Air France-KLM	9.1	24	Iberia	8.0
25	Asiana	8.9	25	Astar	7.6

Source: Compiled by the authors from the ATW World Airline Report, 2005

However, threats to profitability are ever-present, with fuel costs constituting the largest threat. In fact, for many airlines, fuel costs are now higher than labor costs, and airlines have few options in dealing with them. For example, the 15.2 per cent increase in aviation jet fuel from 2005 to 2006 represents roughly a $25 billion increase in the airlines' fuel costs globally.[4] Moreover, since higher fuel costs affect all airlines in a somewhat similar manner, there is little or no competitive advantage to be gained in this area. Therefore, as this extra cost is passed on to the consumer, the relative price structure should remain proportionately the same. In this case, the ultimate question of profitability depends to a large extent on the elasticity of demand for the product and the cost containment ability of the airline's management.

AIRLINE INDUSTRY CONSOLIDATION

The airline industry has been affected by economic recession, rising fuel costs, political uncertainty, and stiff competition. These factors have caused some major carriers, such as Eastern Airlines, Pan Am, and Piedmont into liquidation, and US Airways, United Airlines, Delta Air Lines, and Northwest Airlines into bankruptcy protection.

Table 1.11 displays the market share for various US carriers for the domestic market. In 2000, Southwest Airlines overtook Delta Air Lines to become the largest domestic carrier (in terms of passengers flown) in the United States. This highlights the fact that low-cost carriers are capturing more of the domestic market share while legacy network carriers are losing theirs. This is mainly due to the low-cost carriers' continual expansion and the advantages they possess because of their lower cost structure. Another major trend (in terms of market share) is the emergence of regional carriers. In 1998 ExpressJet, American Eagle, and Skywest had less than 1 per cent combined market share, yet in early 2006 they had acquired 7.5 per cent of the total US domestic market. In this case, the reason was the increased use of regional jets by legacy carriers to open up new markets and combat low-cost carriers.

The industry Herfindahl-Hirschman index (HHI) is a measure of how consolidated the domestic US market is. As Table 1.11 shows, since 1998 the industry has become less consolidated.[5] This spreading out of competition usually equates to lower fares and increased service. In fact, an American Express travel survey has shown that average US domestic airfares have steadily declined since 2000 (Amex, 2006).

Historically, mergers rapidly increased following deregulation in 1978. For the ten years following deregulation there were 51 airline mergers and acquisitions (Dempsey, 1990). The result of these mergers and acquisitions was the creation of six legacy carriers from the 15 independent carriers that had close to 80 per cent US market share in 1987 (Dempsey, 1990).

Although the number of mergers reduced during the 1990s, critics argue that most mergers were still part of well-planned strategies to reduce competition in various markets. Consequently, starting in 1985, the US Department of Transportation (DOT) assumed approval authority for all airline mergers. To approve the merger, the DOT must now balance the consumer benefits resulting from mergers against the possibly negative effects of increasing concentration (Dempsey, 1990). On the other hand, the extraordinary financial problems of legacy carriers suggest that reductions in capacity, whether through mergers or alliances, may be inevitable. Some economists argue that less intense competition, through consolidation and coordination, can actually benefit

4 Information obtained from IATA's Jet Fuel Price Monitor; see http://www.iata.org/whatwedo/economics/fuel_monitor/index.htm.

5 For more information on the HHI see Chapter 9.

Table 1.11 US airline industry domestic market share

Airlines	1999	2000	2001	2002	2003	2004	2005	2006*
Southwest Airlines	14.32%	15.22%	16.56%	16.76%	17.05%	16.73%	17.26%	18.32%
American Airlines	9.87%	10.06%	9.78%	12.75%	11.60%	11.25%	11.35%	11.36%
Delta Air Lines	15.54%	14.68%	13.89%	13.83%	13.04%	12.43%	11.31%	0.00%
United Airlines	12.91%	11.86%	11.18%	10.32%	9.93%	9.66%	8.61%	8.14%
Northwest Airlines	7.11%	7.11%	7.06%	7.12%	7.21%	7.27%	7.14%	6.71%
US Airways	10.76%	10.74%	10.02%	8.60%	7.82%	6.85%	6.41%	16.40%
Continental Airlines	6.21%	5.85%	5.87%	5.51%	5.35%	5.01%	5.02%	5.44%
JetBlue	0.10%	0.23%	0.71%	1.38%	2.10%	2.56%	3.00%	3.61%
America West Airlines	3.25%	3.29%	3.59%	3.87%	3.83%	3.41%	3.31%	0.00%
American Eagle	1.41%	1.48%	1.58%	1.69%	1.86%	2.22%	2.49%	2.66%
Skywest Airlines	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.59%	2.55%
Alaska Airlines	2.28%	2.22%	2.48%	2.54%	2.66%	2.57%	2.56%	2.36%
ExpressJet Airlines	1.05%	1.16%	1.30%	1.50%	1.84%	1.99%	2.16%	2.29%
AirTran Airways	0.43%	0.56%	0.61%	0.77%	1.05%	1.17%	1.22%	1.37%
Industry Herfindahl Index	957.02	932.28	918.13	940.59	890.24	838.02	808.58	804.03

Source: Compiled by the authors using Back Aviation Solutions O&D data.
* 2006 figures represent only Q1 data.

consumers by allowing airlines to build more efficient networks with greater economies of scale, scope, and density.

Figure 1.3 provides a framework of the major airline mergers that have occurred in the United States since deregulation.

While mergers and acquisitions have slowed in the United States, the relatively low HHI and the poor financial condition of the legacy carriers indicates that there is still the potential for additional mergers and acquisitions within the industry. Moreover, airline mergers are not limited to the United States; they have played a large part in international aviation. A few recent international mergers include:

- Air France and KLM
- Lufthansa and Swiss Air
- Air Canada's acquisition of Canadian Airlines
- Japan Airlines' purchase of Japan Air System.

Historically, mergers have not been very successful in the aviation industry. Many mergers do not realize the benefits envisioned when planned, and one-off merger costs, such as aircraft painting and IT harmonization, end up being far more costly than planned. Airline mergers also bring difficulties in dealing with labor groups, especially with regard to such issues as merging seniority lists. Corporate culture can also be a

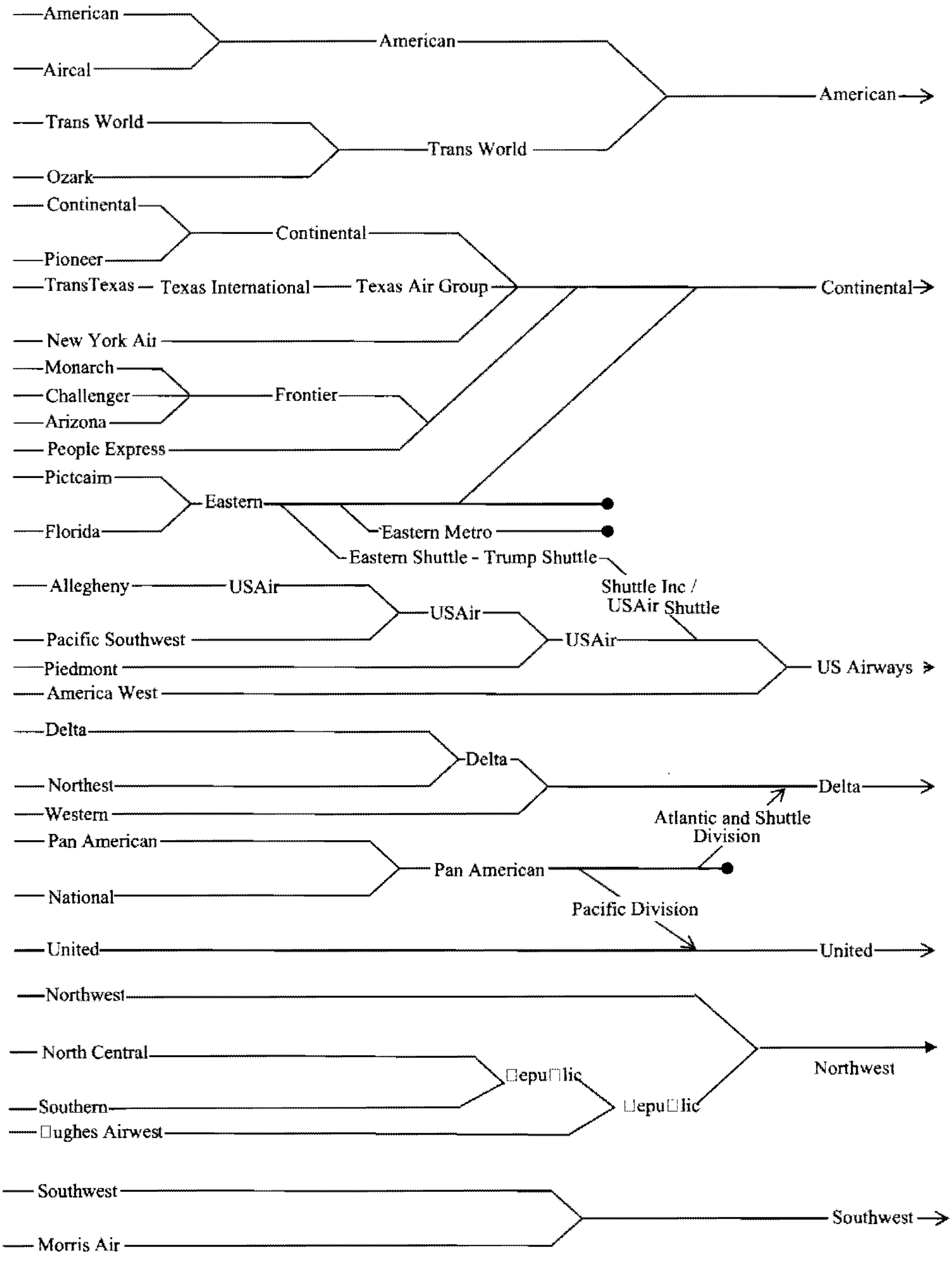

Figure 1.3 The evolution of the US airline industry

much underestimated barrier to successful mergers as different companies' cultures may impede merger success. Finally, one of the greatest challenges a merger faces is managing multiple and powerful stakeholders; these can include, but are not limited to, politicians, regulators, labor leaders, and consumers (McKinsey & Company, 2001). Many of these stakeholders are suspicious of the mergers because they fear lessened competition and increased travel prices (McKinsey & Company, 20001). Many potential mergers have been thwarted by regulators, and one of the key measures that regulators use in analyzing potential mergers is the planned mergers' affect on the HHI.

However, there are, as mentioned, potential benefits from airline mergers. McKinsey & Company (2001) estimate that a merger of two mid-sized carriers could unlock synergies in excess of 7 per cent. The principal benefit of mergers is cost rationalization. Since the airline industry exhibits large economies of scale, merged airlines are able to spread their high fixed costs over a greater network. In addition, the new merged carrier can increase its bargaining power with key suppliers and merge such functions as parts inventories, back office functions, and sales forces (McKinsey & Company, 2001).

Another major benefit of mergers is network harmonization. This can include a variety of things, but in the final analysis the merged airline's route network is greater than that of the individual airlines. A good example of this is the America West–US Airways merger. America West was predominantly a west-coast airline while US Airways was primarily an east-coast airline, but the merged airline had a strong route network on both coasts. Without a merger either carrier would have found it difficult to in increase its presence on the opposite coast. The economies of scope that resulted from the merger allowed US Airways to widen its customer base and strengthen its market power throughout the United States.

Another way to look at consolidation in the domestic US industry is to look at it at the airport level. After deregulation the major carriers adopted a hub-and-spoke system, funneling passengers through a few airports (Dempsey, 1990). This in turn led to some carriers holding dominant positions at certain hub airports throughout the United States. Table 1.12 depicts the consolidation of carriers by enplanements and operating carrier at the ten largest airports in the United States.

The general trend in airport consolidation from 1998 through 2005 is one where the largest carrier has become less dominant. This has occurred for all the airports presented in Table 1.12, except for Dallas. In Dallas, Delta's withdrawal left American Airlines as the only major airline still operating; this also left Dallas as the most consolidated airport of the top 10 domestic US airports in 2005.

The general reduction in consolidation at US airports can largely be attributed to two factors: first, increased competition, particularly from low-cost carriers; and, second, major carriers pushing more flying to regional affiliates. Low-cost carriers such as Frontier and AirTran have situated themselves in the dominant hubs of Denver and Atlanta, and have been successful at taking away market share in those airports. With the emergence of regional jets, major carriers have been pushing capacity towards regional carriers in an effort to reduce costs.

In May 2007 Bombardier Aerospace introduced the next-generation versions of its CRJ700, CRJ900 and CRJ1000 regional jets. These new CRJ NextGen aircraft have featured significant operating cost improvements and the increased use of composite materials.

Since the Form41 data used in Table 1.12 break data down by operating carriers, regional carriers are treated separately. For example, ExpressJet operations at Houston are separate, even though the flights are marketed by Continental. This could potentially distort the level of consolidation at airports with large regional carrier presence.

As shown in Table 1.12, the HHI indicates a decline in consolidation for the US domestic airline industry. This mirrors a similar trend when consolidation is analyzed on an airport basis. However, it is important to remember that the level of consolidation at major US airports is much greater than the level of consolidation of the airline industry. The least

Table 1.12 Major US airports concentration (with enplanements by operating carrier)

		1998	1999	2000	2001	2002	2003	2004	2005
Atlanta (ATL)	Largest Carrier %	66.50%	68.13%	68.47%	63.66%	64.40%	62.90%	66.28%	66.82%
	Largest Carrier	DL	DL	DL	DL	DL	DL	DL	DL
	Herfindahl Index	4470	4732	4807	4189	4307	4152	4611	4794
Chicago (ORD)	Largest Carrier %	48.41%	48.97%	46.55%	46.23%	46.46%	42.91%	40.69%	36.53%
	Largest Carrier	UA	UA	UA	UA	UA	UA	UA	UA
	Herfindahl Index	3456	3469	3356	3256	3303	2848	2716	2421
Dallas (DFW)	Largest Carrier %	64.90%	63.71%	62.81%	62.43%	64.16%	63.85%	65.84%	73.79%
	Largest Carrier	AA	AA	AA	AA	AA	AA	AA	AA
	Herfindahl Index	4585	4433	4306	4226	4397	4286	4521	5607
Denver (DEN)	Largest Carrier %	70.88%	67.29%	62.23%	61.11%	56.80%	52.09%	48.84%	44.46%
	Largest Carrier	UA	UA	UA	UA	UA	UA	UA	UA
	Herfindahl Index	5133	4671	4077	3935	3465	3021	2729	2432
Detroit (DTW)	Largest Carrier %	71.18%	71.29%	70.70%	69.78%	70.05%	67.52%	67.45%	64.63%
	Largest Carrier	NW	NW	NW	NW	NW	NW	NW	NW
	Herfindahl Index	5173	5202	5115	4996	5023	4686	4684	4358
Houston (IAH)	Largest Carrier %	77.82%	76.76%	74.69%	75.07%	74.05%	70.02%	66.73%	64.57%
	Largest Carrier	CO	CO	CO	CO	CO	CO	CO	CO
	Herfindahl Index	6158	6000	5716	5782	5662	5205	4855	4714
Las Vegas (LAS)	Largest Carrier %	30.10%	32.06%	31.48%	34.22%	34.80%	35.15%	33.89%	33.69%
	Largest Carrier	WN	WN	WN	WN	WN	WN	WN	WN
	Herfindahl Index	1649	1678	1600	1745	1781	1802	1705	1701
Los Angeles (LAX)	Largest Carrier %	31.55%	30.59%	30.58%	28.03%	24.17%	22.27%	21.91%	20.47%
	Largest Carrier	UA	UA	UA	UA	UA	UA	UA	UA
	Herfindahl Index	1564	1516	1564	1471	1380	1278	1268	1290
Minneapolis (MSP)	Largest Carrier %	74.36%	74.83%	72.97%	73.33%	74.75%	71.28%	69.43%	67.05%
	Largest Carrier	NW	NW	NW	NW	NW	NW	NW	NW
	Herfindahl Index	5635	5699	5431	5461	5676	5176	4930	4625
Phoenix (PHX)	Largest Carrier %	42.63%	42.09%	42.87%	44.25%	44.88%	40.87%	38.06%	37.55%
	Largest Carrier	HP	HP	HP	HP	HP	HP	HP	HP
	Herfindahl Index	2635	2684	2732	2872	2876	2550	2387	2441

Source: Compiled by the authors using Back Aviation Solutions Form41 data.

consolidated US airport in the list, Los Angeles, still has an HHI score well above the industry's level of consolidation. Therefore, in at least a few markets, potential airline mergers could have a much greater and controversial effect on the level of consolidation at some major airports.

FACTORS AFFECTING WORLD AIR TRAFFIC GROWTH

The factors that affect world air traffic growth are numerous and complex; they occur on global, national, regional, and civic levels. This complexity helps explain why air travel can be growing significantly in one country or city and why it is stagnant or floundering in another—chief among these factors is the level of prosperity in the region. This amount of economic prosperity is measured by such indicators as gross domestic product (GDP) or gross national product (GNP). GDP is the total market value of all final goods and services produced in a country in a given year. Increased prosperity derives increased demand for air travel in two separate, but concurrent, ways. First, increased economic activity helps generate employment, which ultimately causes an increase in business travel, the most important segment of travelers for airlines. Business travel is the primary reason why world financial centers such as London and New York have experienced strong air traffic growth. In addition, increased economic activity also spurs air cargo growth.

The second result of economic prosperity is a decrease in unemployment and a concurrent increase in household income. People have more discretionary income and are able to afford more leisure travel trips. A good example of this has been China where a growing middle class has fueled a large expansion of air travel within the country.

A decrease in the real cost of air travel will also create air traffic growth. This was first experienced in the 1970s when deregulation resulted in a dramatic decline in the cost of air travel. A greater number of people could now afford air travel, and they took advantage of the opportunities. Low-cost carriers generated increased air travel with low fares, and airports experienced tremendous growth in their passenger statistics as soon as a low-cost airline initiated service. This phenomenon has been coined the "Southwest Effect". Ryanair is accomplishing similar feats in Europe where almost everyone can now afford weekend getaways.

Another factor influencing world air traffic growth is population growth rates. Strong population growth rates in developing countries such as India and China have helped spur air travel growth. However, population growth must generally be accompanied by income growth for this factor to significantly affect air travel.

Economic liberalization is another major factor impacting on air transportation. Government restrictions on an economy, such as wage and price controls or excessive regulation, ultimately constrain demand. When such artificial barriers are lifted, the marketplace dictates demand for goods and services, and increased air travel is almost always the result. The reason for this is the fact that government regulation in the aviation industry usually involves ticket prices and market access—that is, favored airlines (usually a national airline) are granted monopoly access with some sort of a fare structure that is structured to cover average costs. This effectively eliminates competition and restricts the growth of air traffic. A good example of economic liberalization is the United States itself. Following deregulation, airfares plummeted and air traffic growth increased significantly. Moreover, the freedom for airlines to fly to whatever destination they wished made flying more convenient for passengers by providing more non-stop flights with greater frequency. Recent air transport liberalization in Europe and India has led to a tremendous growth in air traffic in these countries.

Politics and political stability also play a role in air travel. It is not surprising that countries that choose extremely protectionist and radical policies do not experience great air traffic growth; in these cases, the government restricts air travel as a matter of political policy. Political instability can also greatly influence air travel, since people do not want to travel to regions where they feel unsafe. It is likely that political instability can be blamed for the poor air traffic growth rates in those parts of Africa where governments are in constant turmoil. Finally, political instability reduces and/or restricts business activities within the country.

Terrorist attacks can also affect air travel. After the tragic events of 11 September 2001, air travel dropped off drastically as passengers no longer felt safe traveling. In addition, many felt that it was also unsafe to travel to other international destinations in case some other terrorist attack should occur.

Finally, the amount of leisure time people have can affect the demand for leisure flights. Typically, individuals who possess greater discretionary free time have a greater demand for leisure and/or vacation flights. Tourism promotion can also help spur an increased demand for air travel to a particular destination. For example, Walt Disney World has turned Orlando into the number one destination airport in the United States.

THE ECONOMIC IMPACT OF THE AIR TRANSPORT INDUSTRY

Commercial aviation comprises two primary segments: large commercial air carriers and regional/commuter air carriers. Since the deregulation of the commercial airline industry in 1978, both the large commercial and regional/commuter air carriers have enjoyed more robust growth than the domestic economy. Generally, the commercial airline industry has closely followed the movement of the domestic economy. Since deregulation, the large US commercial air carriers have averaged an annual growth of 4.8 per cent in terms of revenues, compared to a 2.6 per cent average growth rate of the US GDP. During the same period, US regional/commuter air carriers grew at an annual growth rate of 14.3 per cent.

Despite the negative impact of the events of 9/11 and the concurrent recession, the US Federal Aviation Authority (FAA) forecasts long-term growth in enplanements for the large US air carriers to average 3.3 per cent through 2013, while long-term growth will average 5.5 per cent for the US regional/commuter airlines. And, as has been mentioned, international aviation continues to grow with IATA forecasting higher growth rates than the US industry (Pearce, 2006). This growth has been largely spurred by the soaring economies in the Asia-Pacific region. So, what economic impact is the growth of air transportation likely to have on the economy?

Economic impact can be divided into three categories: direct, indirect, and induced. Direct impact represents economic activities that would not have occurred in the absence of air transportation. In the air transportation industry, both airlines and airports provide the economy and local communities with a direct economic impact. Examples of direct economic impacts include the salaries of airline personnel, fuel purchased, landing fees, salaries of airport personnel, and other similar purchases and expenditures. Indirect economic benefits include the financial benefits that are attributed to airport/airline activities. Examples of indirect economic impacts for air transportation include hotels, restaurants, and other retail activities. There is usually a causal relationship between the industry and indirect impacts. For example, if a community experienced a reduction in air travel, the hotel industry in that community would most likely suffer a fall in room occupancy rates as well. Finally, induced economic impacts are the multiplier effects of the direct and indirect impacts. Induced impacts account for the increased employment and salaries that come from secondary spending that results from the direct and indirect economic impacts. The total of these economic impacts measure the importance of an industry in terms of the employment it provides and the goods and services it consumes. The following sections explore the effect of air transportation on each of these economic impacts.

Direct Impact

Direct economic impacts are the consequences of what might be termed first-tier economic activities carried out by an industry in the local area. In the air transportation industry, airports provide the greatest direct impact to local economies. The reason for this is the more or less obvious fact that the economic activities that take place at the airport directly involve the local economy. Most direct impacts, like airport employment and fixed-based operations, occur at the airport; others, like the local production of goods and services for use at the airport, may occur off-site.

Expenditures by airlines, fixed-based operators and tenants also generate direct impacts, but only those expenditures that lead to local business activity are relevant for a regional economic assessment. For this reason, it is important to distinguish between the local value-added component of expenditures and the regional import component. Thus, airline expenditures on fuel generate local fuel storage with distribution systems and also contribute to the importation of fuel into the region. In most parts of the country, only the former component is relevant for any local economic impact analysis. Therefore, the direct economic impacts of air transportation for a community are usually measured on the basis of the airport's immediate economic activity. In addition, large aircraft manufacturers can generate a huge direct economic benefit by locating their production facilities in a given community or state. For example, the direct economic impact of the Boeing 787 Dreamliner project on Washington State in 2006 has been estimated at approximately 11,470 jobs, with an economic output of $2.268 billion (Deloitte, 2004). There are, of course, numerous other examples of large direct economic impacts provided by the air transportation industry.

Indirect Impact

Indirect impacts derive from off-site economic activities that are attributable to air transportation activities. For example, indirect economic impacts include services provided by travel agencies, hotels, rental car companies, restaurants, and retail establishments. These enterprises have a strong relationship to the air transportation industry and, like airport businesses, employ labor, purchase locally produced goods and services, and invest in capital projects. Indirect impacts differ from direct impacts because they originate entirely off-site. Typically, indirect economic impacts are generated by visitors to the area, who are traveling by air. A good example of an industry that has a strong indirect economic impact relationship with air transportation is the hotel industry. Airlines provide economic benefits to the hotel industry by requiring hotel rooms for passengers who have business, or are vacationing, in a city. This increased demand for hotel accommodation in the city creates employment and may require the construction of more hotels, thereby creating more economic impact. The large demand for hotel accommodation caused by air transportation is one of the main reasons why areas around major airports almost always contain numerous hotels.

Induced Impact

As mentioned earlier, induced economic impacts are the multiplier effects that are caused by the increases in employment and income generated from the direct and indirect economic impacts of air transportation. A simple example will help make this concept clear. Imagine a new airline employee who purchases a house in the local community. The builder of the house then uses this income to purchase other goods and services, and the income to the suppliers of these goods and services is also spent. This framework of expenditures is the basis behind the multiplier effect—that is, one transaction leads to multiple economic transactions.

More economically self-sufficient regions tend to have higher multipliers than do regions that are more dependent on regional imports, since more of the spending and re-spending is done within the region. Therefore, the larger the region under consideration, the higher the multiplier will be.

Total Impact

Total economic impact is defined as the sum of direct, indirect, and induced impacts. It is usually expressed in terms of economic output, earnings, or employment (sometimes full-time equivalents). The basic formula for total economic impact is:

$$\text{total impact} = \text{direct impacts} + \text{indirect impacts} + \text{induced impacts}$$

Table 1.13 provides a comparison of the total economic impact in terms of employment for six airports located in the United States. The report for each airport was done independently and at different times, but the methodology used for each is similar. Although the six airports vary in size, they all provide strong economic impacts for their communities. When normalized in terms of commercial departures, Wichita and Memphis both generate one job for every departure. In Memphis one additional daily flight would generate approximately 365 new jobs for the region. Wichita's extremely high ratio is probably attributable to the large manufacturing and maintenance facilities for Cessna and Bombardier. The presence of FedEx in Memphis explains its high economic impact to departure ratio. And, finally, much of Seattle's total economic impact can be attributed to the simultaneous indirect economic impact of the presence of aircraft manufacturing giant, Boeing, and of tourism.

These disparate examples highlight the diversity (cargo operations, manufacturing, and tourism) and strength of the economic impact of the aviation industry.

THE OUTLOOK FOR THE AIR TRANSPORT INDUSTRY

Since demand for the air transport industry is highly correlated with overall economic growth, it is not surprising that the global outlook for the air transport industry mirrors the global economic outlook. Therefore, the air transport industry is expected to grow significantly in regions where economies are developing, such Asia-Pacific, while other regions' air transport outlook is expected to be steady. GDP and economic growth are strong leading indicators of the air transport industry's growth, so, in the short term, these measures can be used to assess the industry.

Table 1.13 The economic impact of selected airports

Airport	Year of Report	Total Jobs	Total Jobs per Commercial Departure
Cincinnati/Northern Kentucky	2000	78,573	0.6510
Memphis	2004	165,901	1.0096
Greenville-Spartanburg	2003	5,787	0.2461
Minneapolis — St. Paul	2004	153,376	0.6299
Seattle — Tacoma	2003	160,174	0.9639
Wichita	2002	41,634	3.1835

Source: See chapter references for all six source documents.

However, direct correlations between GDP and air transport growth are never exact, due to a variety of issues. For example, structural barriers in the air transport industry can cause drastic differences between economic growth and the growth of the air transport industry. A good example of this was the effect of deregulation in the United States; deregulation was a major structural change that caused a rapid increase in the air transport industry's growth compared to overall economic growth.

Airport capacity and, in the US, antiquated air traffic control are also potential structural barriers. Major international airports in the United States and Europe have severe capacity issues with relation to the number of aircraft that they are capable of handling. As these capacity limits are reached, delays at these airports tend to increase exponentially. These delays, especially if they are on an ongoing basis, discourage demand and constrain growth.[6] Similar capacity issues could plague airports in the Asia-Pacific region, especially Indian, Chinese, and Japanese airports. This capacity barrier to air transport growth is a prime reason why Airbus embarked on the creation of its new super-jumbo A380 aircraft.

The two major sources for the long-term air transport outlook are Boeing and Airbus. Each aerospace giant has published its forecasts for the future of the aviation industry. They have similar growth estimates for world air traffic growth, with Boeing forecasting that world revenue passenger kilometers (RPKs) will grow at 4.9 per cent per annum for the next 20 years and Airbus forecasting a 4.8 per cent per annum growth rate (Boeing, 2006a; Airbus, 2006). Airbus (2006) forecasts a greater annual growth rate (5.3 per cent) for the first half of the period than the second half (4.4 per cent). These global RPK forecasts mimic the historical global growth rate of 4.8 per cent from 1985 to 2005. However, the key to understanding the global air transport industry is on a regional basis (Boeing, 2006a). Table 1.14 summarizes the regional growth rate forecasts for Boeing and Airbus.

Both Airbus and Boeing also forecast worldwide demand for new aircraft for the next 20 years. Not surprisingly, each company's forecasts vary slightly, highlighting each company's strategic plan and product offerings. Boeing (2006a) estimates that there will be a demand for 23,760 aircraft seating over 90 passengers in the next 20 years, while Airbus (2006) forecasts a worldwide demand for 21,860 similar-sized aircraft over the same period. Although both companies agree that roughly 70 per cent of the demand for

6 See Chapter 5 for an analytical discussion of this issue.

Table 1.14 Regional economic growth forecast

Region	2006–2015	2016–2025	20 year growth
North America	4.30%	3.60%	4.00%
Latin America	6.30%	5.20%	5.80%
Europe	5.00%	4.20%	4.60%
CIS	6.10%	5.10%	5.60%
Middle East	8.10%	4.80%	6.40%
Asia	7.40%	5.00%	6.20%
Africa	6.00%	4.60%	5.30%
World	5.30%	4.40%	4.80%

Source: Compiled by the authors using *Airbus Global Market Forecast, 2006–2025*.

new aircraft will be for single-aisle aircraft, Airbus (2006) predicts a greater demand for large wide-body aircraft, while Boeing (2006a) believes the remainder of aircraft demand will be for small and medium wide-body aircraft. Airbus (2006) forecasts demand for 1,263 very large aircraft (747s and A380s), while Boeing (2006a) only forecasts 970 aircraft in this segment. Furthermore, the firms differ on where demand for new aircraft will be. Boeing (2006a) still forecasts that the North American market will be the largest market for new aircraft (mostly narrow-body aircraft), while Airbus (2006) forecasts the Asian-Pacific market will order the most aircraft in the next 20 years. In addition, Airbus (2006) foresees greater low-cost carrier growth in this region to spur narrow-body sales.

One other sector of the air transport industry that should be mentioned is the air cargo market. Both Boeing (2006b) and Airbus (2006) forecast world air cargo to grow by about 6 per cent per year for the next 20 years. This worldwide forecast growth outstrips passenger growth forecasts, and this situation is especially true in international markets where the air cargo industry has not developed to the extent of the passenger industry. As a result, demand for cargo aircraft (new or secondhand) is expected to be strong, especially for wide-body aircraft. China is expected to lead the way in air cargo growth, both domestically and internationally. The US domestic air cargo market appears to be mature, with Airbus (2006) forecasting a modest 3.3 per cent annual air cargo growth and Boeing (2006b) forecasting a 3.8 per cent growth rate.

SUMMARY

As the preceding discussion and statistics amply demonstrate, the air transportation industry is a large and growing segment of the domestic and international economies. As such, it is an important area for economic analysis. Although the industry is similar in some ways to other large industries, it has some peculiar characteristics that can best be understood in the context of standard economic analysis. The following three chapters will introduce basic economic theory, including demand, supply, costs, and production analysis. These concepts will be presented in the context of the aviation industry with applicable examples.

REFERENCES

Airbus (2006). *Global Market Forecast 2006-2025.* Retrieved on 14 January 2007 from: http://www.airbus.com/en/corporate/gmf/index.html.

Amex (2006). *2005 US Domestic Airfares for American Express Business Travel Clients Drop to Six-Year Low.* Retrieved on 31 August 2006 from http://home3.americanexpress.com/corp/pc/2006/4q05_monitor_print.asp.

Boeing (2006a). *2006 Current Market Outlook.* Retrieved on 14 January 2007 from: http://www.boeing.com/commercial/cmo/.

Boeing (2006b). *World Air Cargo Forecast 2006-2007.* Retrieved on 18 January 2007 from: http://www.boeing.com/commercial/cargo/index.html.

Center for Economic Development and Business Research (2003). *Wichita Mid-Continent Airport Economic Impact.* Retrieved on 13 September 2006 from: http://webs.wichita.edu/cedbr/AirportImpact.pdf.

Deloitte (2004). *Employment and Income Analysis of the Boeing 7E7 Project.* Retrieved on 13 September 2006 from: http://www.aia-aerospace.org/stats/resources/Boeing_EmploymenAndIncomeAnalysis.doc.

Dempsey, P. (1990). *Flying Blind: The Failure of Airline Deregulation.* Washington, DC: Economic Policy Institute.

Economics Research Group (1999). *The Cincinnati/Northern Kentucky International Airport Economic Impact Analysis.* Retrieved on 13 September 2006 from: www.cba.uc.edu/econed/1998-2011impact.pdf.

John C. Martin Associates (2005). *The Local and Regional Economic Impacts of the Minneapolis/St. Paul International Airport.* Retrieved on 13 September 2006 from: www.mspairport.org/msp/docs/misc/mspimp04_FINAL.pdf.

Martin Associates (2005). *The 2003 Economic Impacts of the Port of Seattle.* Retrieved on 13 September 2006 from: http://www.portseattle.org/downloads/business/POS2003EIS_Final.pdf.

McKinsey & Company (2001). Making Mergers Work. *Airline Business.* Retrieved on 31 August 2006 from Air Transport Intelligence.

Pearce, B. (2006). New Financial Forecast. *IATA Industry Financial Forecast, September.* Retrieved on 31 August 2006 from: www.iata.org/economics.

Sparks Bureau of Business and Economic Research (2005). *The Economic Impact of Memphis International Airport.* Retrieved on 13 September 2006 from: www.memphisairport.org/EcImpactFinal.pdf.

Wilbur Smith Associates (2003). *The Economic Impact of Greenville-Spartanburg International Airport—Update 2003.* Retrieved on 13 September 2006 from: www.gspairport.com/images/downloads/AirImpact.pdf.

2

Principles of Economics

Economics is haunted by more fallacies than any other study known to man. This is no accident. The inherent difficulties of the subject would be great enough in any case, but they are multiplied a thousand fold by a factor that is insignificant in, say, physics, mathematics, or medicine—the special pleading of selfish interests.

Henry Hazlitt

This chapter introduces the "economic way of thinking," primarily through the study of supply/demand and the logic of prices. The example of price controls illustrates how mistaken casual analysis can be and that any calculation of costs must consider opportunity costs, such as waiting in line when artificially low prices create shortages. The chapter also demonstrates how prices efficiently allocate resources and motivate appropriate behavior in the framework of Adam Smith's "Invisible Hand." A similar analysis shows how low landing fees can actually impose greater costs on airlines as the concept of external costs is incorporated. A short section on public choice explains how government policy can sometimes go awry through the impact of rational political ignorance, undue special interest influence, and bureaucratic inefficiency. In this chapter we discuss the following topics:

- Fundamentals of economics
- The economic way of thinking, including:
  - Demand
  - Supply
  - Equilibrium
  - Changes in equilibrium
  - How equilibrium price maximizes consumer well-being
  - Price controls
  - Airport landing fees and airport congestion
- The economics of government, including:
  - Incentives for a voter to be well-informed
  - Undue special-interest influence
  - Bureaucratic inefficiency
  - Government failure versus market failure
  - Reforming government.

Once ancient astronomers had proven that the sun, rather than the earth, was the center of the solar system, several centuries passed before this truth was commonly accepted by people who weren't experts in astronomy. Economists can relate to this. However, we will soon see that common misperceptions about the economy are often more harmful than the mistaken belief that the sun revolves around the earth. Perhaps common economic myths are more comparable to the old notion that illness was caused by "bad blood" and that slicing open a vein to drain off some blood could increase someone's chances of recovering from pneumonia or other serious disease. In reality, of course, blood-drained patients are all the more likely to die. Errors in economic reasoning often have similar results. Today, fortunes are sometimes destroyed, businesses driven to liquidation, and entire economies plunged into depression because of the common belief in errant economic theories—economic fallacies that have been clearly exposed in textbooks for decades and, in a few cases, for centuries. Our aim is to help you avoid many of these self-inflicted wounds, to point out the error of some common economic myths and to provide a foundation for applying economic reasoning to whatever problems are encountered.

FUNDAMENTALS OF ECONOMICS

> If all the economists were laid end to end, they'd never reach a conclusion.
>
> George Bernard Shaw

When one observes two economists on television, most of the time, it seems, they disagree. One insists, say, that the price of oil, and hence jet fuel, has peaked and will soon decline, while the other seems certain that oil and jet fuel prices are headed much higher. Given such media appearances, many people assume that economics is a wildly controversial, unsettled field where one opinion is as good as another. In fact, as is often the case, the popular media is extremely misleading in this regard (Swartz and Bonello, 2003). Although there are, of course, some continuing controversies, there is also a widely accepted body of economic knowledge. Indeed, most of the economic principles we will discuss and apply in this text have been settled for a century or more.

So, why does economics seem more controversial than it actually is? One key reason is the emphasis on economic forecasting. Predicting the exact future behavior of human beings is extremely difficult and therefore inevitably controversial. Knowledge of economics will help you make better predictions of future jet fuel prices, but an educated guess is still a guess. However, the ability to foresee the future in exact detail is an unfairly high threshold. Doctors cannot always predict what diseases a patient might contract in the future, and different doctors might offer different predictions of how severe the symptoms might eventually be, but that doesn't mean that doctors are powerless to help you if your arm is broken.

Similarly, economists can't reliably tell you what the price of jet fuel will be four years from now, but they can lay out the key factors that will determine that price. Returning to the case where two economists offer two very different forecasts, if you asked the two of them to explain the key factors that determine jet fuel price, you would probably find them in complete agreement. They would certainly agree, for instance, that economic growth tends to raise the price of all oil products because a rising standard of living increases demand for oil. They would also agree that some easing of environmental restrictions on oil drilling will tend to reduce the price since this will increase supply. Of course, both economic growth and environmental regulations are affected by government policies and hence political elections. So, to be certain of the future price of jet fuel, one would have to know the outcomes of future elections and exactly how politicians will affect environmental regulations and economic growth—an obvious impossibility.

However, there are problems that go beyond the complexity of forecasting. To put matters very bluntly, economists are not always completely straightforward, particularly in the context of public policy issues. The same may be said, of course, for any other profession. When, for instance, a lawyer proclaims the innocence of his client we all recognize that the lawyer is paid to make that claim and may not believe it at all. When a scientist knows that taking a certain position on global warming will garner an outpouring of favorable media attention and increase the likelihood of obtaining lucrative government grants, we understand that the allure of fame and fortune may trump integrity and scientific objectivity (Agin, 2006). Alas, some economists, often very prominent in the profession, succumb to the same sort of temptation. It is not difficult to find economists employed by politicians sometimes saying things that they know to be untrue. Indeed, economists will sometimes do this even if they are not employed by the politician they defend but in hopes of gaining future employment or just to help the politicians who they believe to be a lesser evil than their opponents (Rubner, 1979). Thus, just as it's possible to find some lawyer to proclaim the innocence of any guilty criminal, it is also possible to find some economist somewhere willing to make false claims against any known economic fact. No wonder most of the public thinks that, no matter how many economists you lay end to end, they can never reach a conclusion!

There is, however, some good news. Economic truths are a good deal easier to understand than quantum physics. It isn't necessary to accept economic principles on the basis of someone's word; with a little work the average person can follow the trail of logic and reach the same conclusion that objective economists reach—to a large extent, you can learn to be your own economist.

THE ECONOMIC WAY OF THINKING

> The only way that has ever been discovered to have a lot of people cooperate together voluntarily is through the free market. And that's why it's so essential to preserving individual freedom.
>
> Milton Friedman

Economics takes as given that people respond to incentives in a generally predictable manner. Though this is sometimes referred to as the fundamental assumption of economics, economists believe that it is not an assumption at all but a simple fact confirmed by common empirical reality. Students spend more time studying material that is guaranteed to be in the test than material that is unlikely be in the test. They are more likely to do an extra credit project if the project is weighed more heavily in determining their grade. If, other things being equal, the price of air travel increases, then people will fly less. People engage in any given activity more if the cost of that activity falls or if its benefits rise.

The economic way of thinking is simply to take this idea of predictable response to incentives and relentlessly follow it to its logical conclusions. We will see that many surprising insights follow from this basic idea.

Demand

Consider the demand for air travel in, say, the continental United States in one week. Figure 2.1 shows the basic shape of a demand curve and illustrates the *law of demand,* which states that price and quantity demanded are inversely related. That is, all else being equal, people will demand more air travel at Price P_2 then at the higher Price P_1. Thus, if the average fare is initially P_1 and is then cut to P_2, the *quantity demanded* increases from $Q^D{}_1$ to $Q^D{}_2$—an example of people responding to incentives in a generally predictable manner. The seller of any product can reliably cut price to increase sales—we'll consider exactly how much sales might increase for a given price cut in later chapters.

You might question how a slight change in price would change quantity demanded. To understand this it's useful to think of a decision-maker who is *on the margin*—that is, one who is almost indifferent as to whether they take a given flight or not. Consider someone, for example, who is planning a trip and choosing between flying or making a four-hour drive by car. Suppose this individual decides to drive but nearly decided to fly. News that the price of air travel has fallen, even by a slight amount, could tip this person into flying instead. Given large numbers of diverse people, it's a virtual certainty that some decision-makers will be in this sort of position and will therefore respond to very slight changes in incentives.

Figure 2.2 depicts a different event, more desirable from the airline's point of view—an increase in *demand.* Note that the entire curve shifts to the right so that, for any given

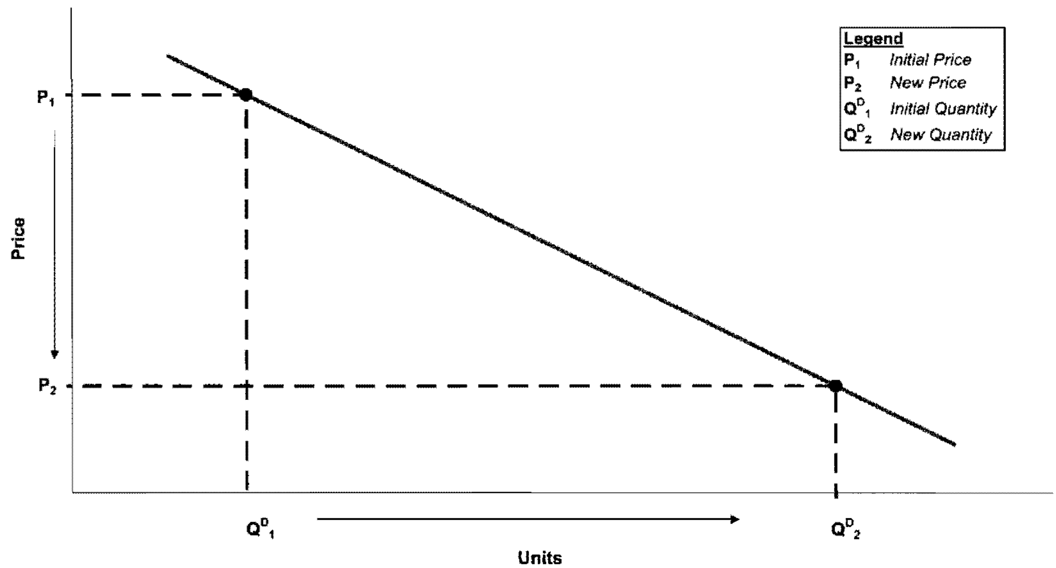

Figure 2.1 Movement along a given air travel market demand curve: a price decrease causes an increase in quantity demanded

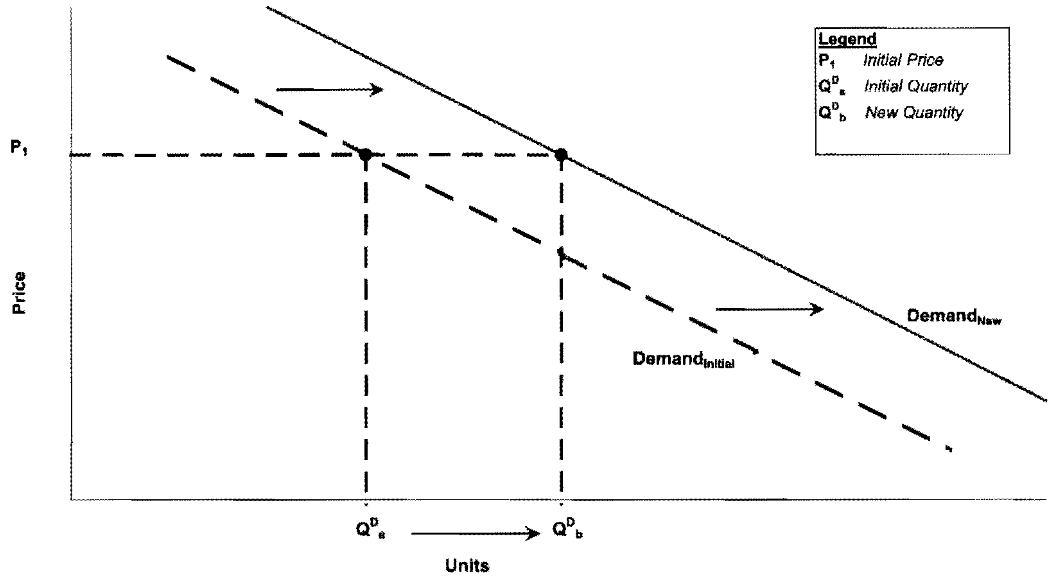

Figure 2.2 Shift of an air travel market demand curve: demand increases, shifts rightward from some change other than the price of air travel

price, passengers wish to buy more air travel than before. If price remains constant at P_1 the amount of air travel bought increases from $Q^D{}_a$ to $Q^D{}_b$. Demand has increased because of some factor other than price; an increase in disposable consumer income, for instance, would trigger such an increase in demand (Colander, 2004).

Factors, such as consumer income, that cause the demand curve to shift are referred to as *determinants of demand*. These can be various, but the most common ones include:

- prices of substitutes (such as travel by rail or car)
- prices of compliments (hotels or rental cars)
- seasonal factors (for example, demand for air travel in surges during holidays)
- general preferences (if, for instance people develop more friendships with others who live far away)
- product quality (safety being especially key in air travel demand)
- random factors (terrorism or natural disasters).

Any change in these determinants that make consumers willing to fly more for a given price will trigger such a demand increase. Likewise, an adverse change—such as rail travel becoming cheaper or hotels becoming more expensive will cause a decrease in demand—a leftward shift of the demand curve (McGuigan, Moyer, and Harris, 2008).

Supply

The supply of air travel, or any other good, is based on production costs. Long-run supply curves can vary tremendously, but short-run supply curves are reliably upward-sloping as in Figure 2.3. We have the same sort of distinctions between movements along the supply curve versus shifts in supply as we had with changes in quantity demanded versus changes in demand. In Figure 2.3, to induce airlines to increase *quantity supplied* from Q^s_1 to Q^s_2, within the space of one week, the average fare must increase from P_1 to P_2. This is reliably the case because many factors of production can't be readily changed in such a short timeframe. For instance, an airline would typically be unable to recruit, hire and train a large number of employees that fast. Thus, to increase, say, available seat miles substantially, the airline would have to pay existing workers overtime; the consequent higher labor costs per available seat mile would mean that the airline would indeed require a higher average fare to be able to cover the higher costs.

An *increase in supply* occurs when the entire supply curve shifts to the right, as in Figure 2.4. That is, it is evident from the figure that airlines are willing to supply more air travel for the same price. In the figure this is shown as an increase in quantity supplied for the same price. This could happen if airline production costs were to fall. Suppose, for instance, that fuel prices decline and airlines are able to lock in these low prices for some time. With this reduction in airline costs markets that weren't quite viable before are now worthwhile, new routes and perhaps additional flights on established routes are added even though the airline can charge no higher fares than before. Although supply often seems less intuitive than demand, supply determinants are in fact straightforward. Supply increases only if production costs fall, or if a new firm enters a given market. Lower production costs might stem from a number of factors: lower taxes on air travel, cheaper aircraft, lower labor costs, and so on.

Similarly, a *decrease in supply*, a leftward shift of the curve, occurs if production costs increase or if a firm exits a given market (Hirschey, 2006).

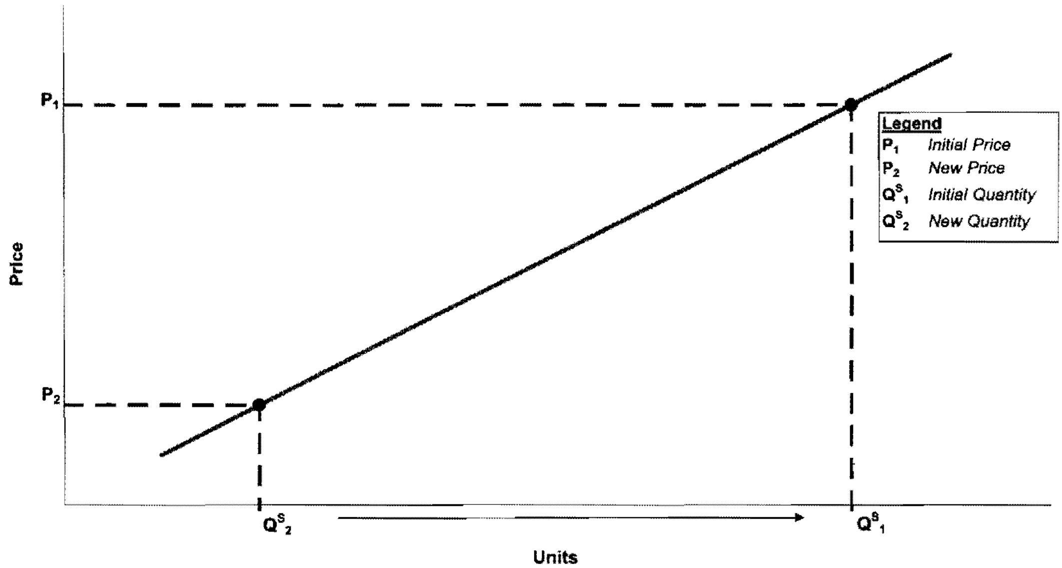

Figure 2.3 Movement along a given air travel market supply curve: a price increase causes an increase in quantity supplied

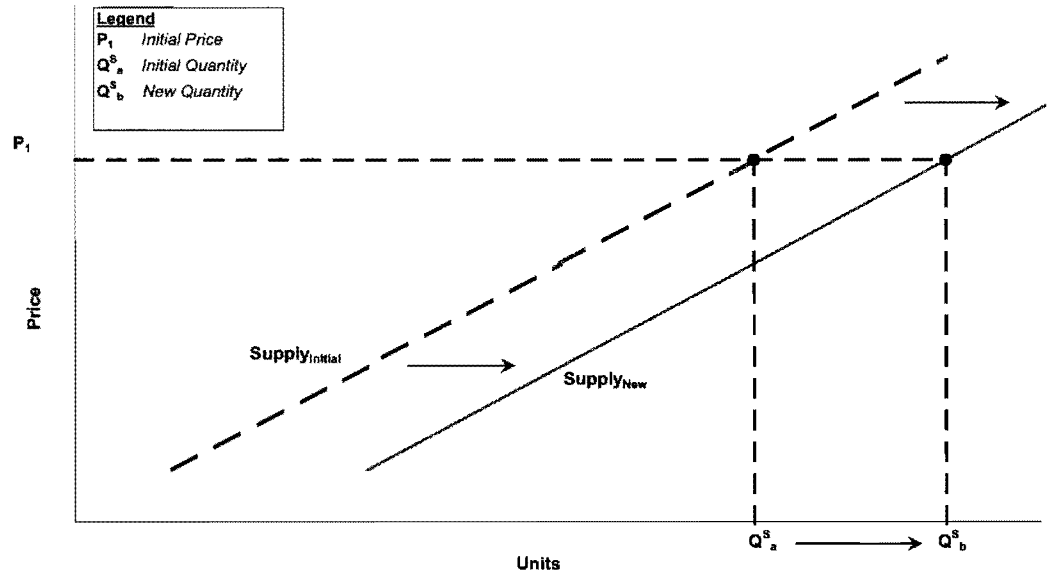

Figure 2.4 **A shift of an air travel market supply curve: supply increases, shifts the curve rightward when production costs decrease or new entry occurs**

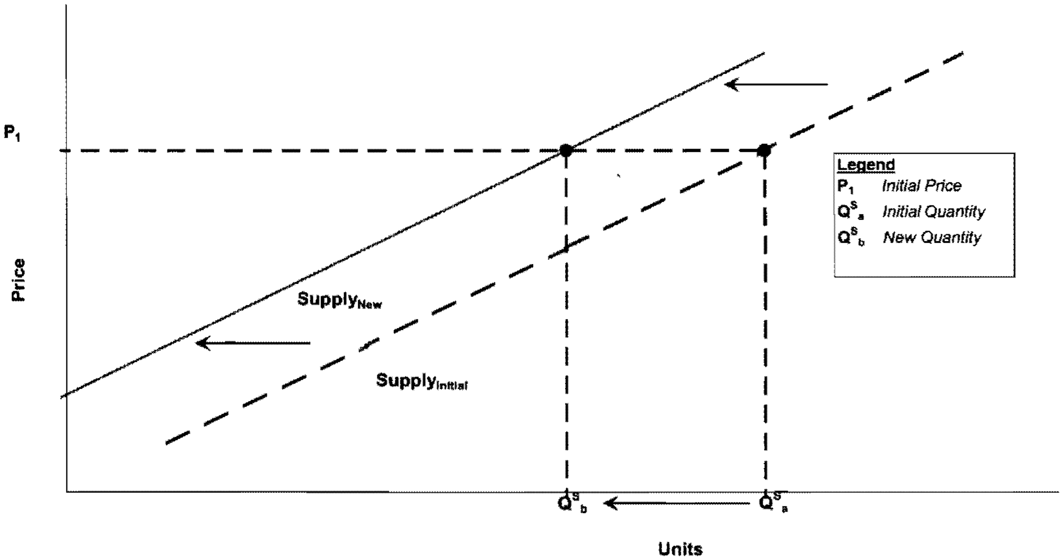

Figure 2.5 **Movement of an air travel market supply curve by itself: supply decreases, shifts the curve leftward when production costs increase or firms exit the market**

Equilibrium

We find *equilibrium price*, as in Figure 2.6, at P^* where supply and demand intersect. At P^* buyers want to buy exactly the same amount that sellers want to sell. Although we tend to think of businesses as being in control of prices it is fair to say that consumers "set" prices just as well as producers. Producers always want high prices whereas consumers always desire low prices, and P^* is a sort of compromise that forces both sides to take into account the other's needs. Producers must receive a high enough price to cover costs and motivate production; consumers must receive a price consistent with their budget.

P^* is achieved through a process of trial and error. The firm estimates demand, plans a level of output and charges a price based on that estimate. Suppose an airline underestimates demand and therefore charges too low a ticket price, P_1, and offers too little output, $Q^S{}_1$, as in Figure 2.6. At this low price the amount demanded by passengers, $D^D{}_2$, is much greater than the output supplied. Airlines, or any other business in this situation, will see tickets being sold at unusually high rates—in other words, the aircraft will begin to fill up its seats much faster than normal. In fact, if airlines don't respond there will be a shortage of seats; they will soon be sold out and the airline will have to turn away numerous customers. However, airlines typically respond fairly quickly. They will raise the price and, because the higher price can cover the higher associated per unit costs, they will increase output. Suppose airlines raise output to $Q^S{}_2$ and increase the average fare to P_2 so that the quantity demanded falls to $D^D{}_1$. This is a step in the right direction, but there is still excess quantity demanded; the price is still too low and seats continue to fill too rapidly. Consequently, price and output will be raised again and will not settle into equilibrium until P^* is reached, at which point Q^S and D^D are equal, at least approximately, at Q^*. Once the average fare reaches P^*, seat inventories behave normally and quantity supplied is brought in balance with quantity demanded as well as possible.

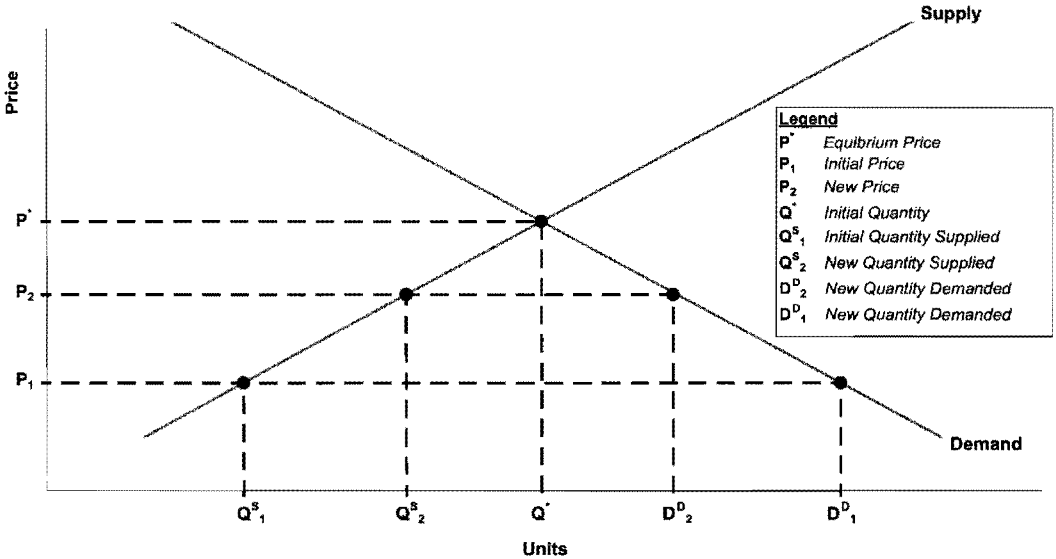

Figure 2.6 **Price movements: average fare is below equilibrium, $D^D > Q^S$, until the price is raised to P^* where $D^D = Q^S$, at least approximately**

Airlines differ from most businesses in that inventories perish—that is, empty seats are worthless once the plane takes off. We will address this complication in some detail in Chapter 11 on revenue management. For now, let us simply acknowledge that selling air travel is more complex than selling, for instance, canned vegetables. The producer of canned vegetables can maintain inventories for some time—whatever is unsold today can be sold in the future. Thus, it is possible to more or less exactly match Q^S to D^D. As airlines' inventories are perishable, it isn't feasible to have 100 per cent load factors—that is, fill every seat on every flight. So, for airlines, Q^s and D^D are only approximately equal.

There is a dynamic process for reaching P^*, equilibrium price, if airlines initially overestimate demand. In that case, the average fare is too high, and very few bookings are received. If this situation persists there will be a surplus of air travel, and aircraft will depart with many more empty seats than normal. Of course, airlines want to avoid this and will therefore bring price down, and reduce capacity until equilibrium is reached at P^*.

Once equilibrium is achieved, price and quantity will remain at P^* and Q^* as long as both supply and demand remain constant. In practice, supply and demand for air travel tend to shift frequently, so we see almost constant changes in price and quantities (Mankiw, 2007).

Changes in Equilibrium

Analyzing changes in equilibrium is straightforward as long as one proceeds by first deducing which curve is shifting in which direction. Suppose, for instance, that the air travel market is in equilibrium initially and then changes when a new event causes passengers to more fully realize how safe commercial flying is compared to other modes of travel. Though the airlines are very pleased with this development, this does not change production costs; therefore we know that the supply curve doesn't move. It stands to reason that a more realistic assessment of airline safety will result in a greater general willingness to fly—for a given price, people will want to fly more than before. This causes an *increase in demand*; the demand curve shifts right, as we see in Figure 2.7. Now, we simply read the graph, based on where the new demand curve, *demand*$_2$, intersects supply to see that both equilibrium price and equilibrium quantity rise. Supply doesn't change because the supply curve is not shifting, but *quantity supplied increases* as we move along the existing supply curve.

Next, suppose that we are initially in equilibrium once again when wages for airline employees increase. This will not shift demand because consumers do not *directly* care much about the details of airline employee compensation. Since higher wage costs do increase production costs, we will shift the supply curve to the left, as presented in Figure 2.8. We see from the intersection of demand and the new supply that equilibrium price rises from P^*_1 to P^*_2. Since consumers do, of course, care about the price of air travel (thus they *indirectly* care about how expensive pilots are, which is embodied in the price of air travel) *quantity demanded* falls, with equilibrium quantity decreasing from Q^*_1 to Q^*_2.

The basic effects of any given shift in supply or demand can be deduced by following these same procedures. As another example, consider the market for oil. The initial supply and demand curves would be at position $\text{Supply}_{\text{Before}}$. When the suppliers decide to collaborate and supply less oil for every price, this causes a backwards shift in the supply curve, to $\text{Supply}_{\text{After}}$.

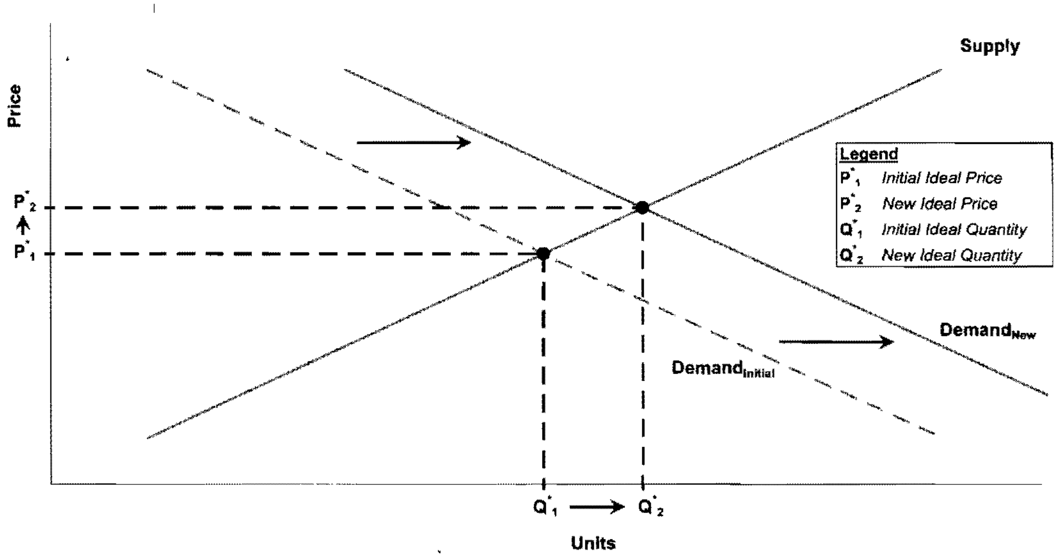

Figure 2.7 **Shift of demand curve: demand increases which causes both equilibrium price and quantity to increase**

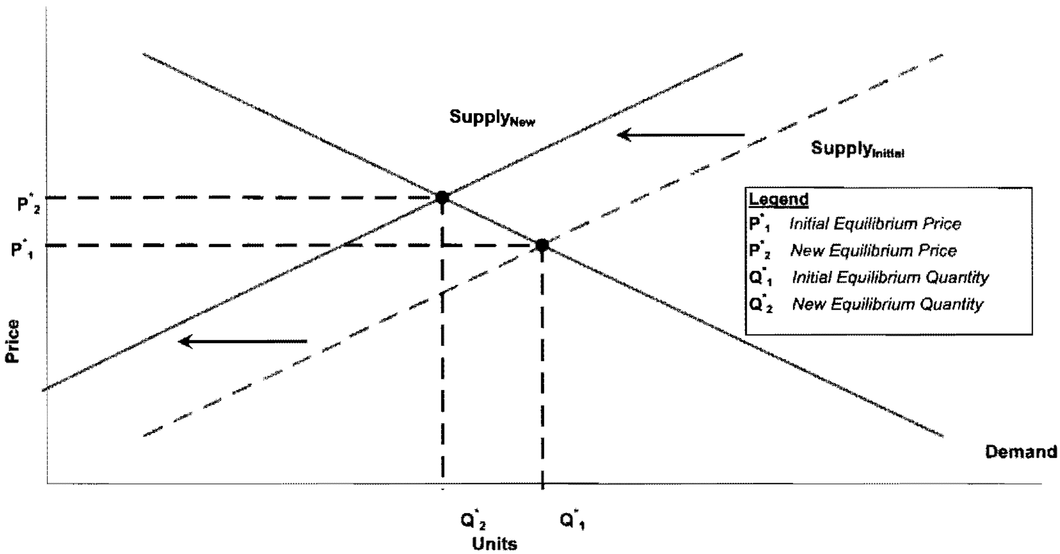

Figure 2.8 **Shift of supply curve: as supply decreases, equilibrium quantity decreases and equilibrium price increases**

Equilibrium Price Maximizes Consumer Well-being

The equilibrium price is generally the best possible price for consumers, given the reality of producer costs. Consumers would love air travel to be provided free of charge, but most have enough sense to realize that a law requiring airlines to give their product away would simply result in airlines shutting down. Price must be high enough to motivate sellers to provide their product at appropriate quality. Sometimes the price may seem higher than necessary to motivate needed production—but things are not always as they seem. It is generally optimal for consumers to let the price move to wherever supply and demand may send it.

It is useful to initially illustrate this point outside the air travel market and then apply a similar logic to a more complex aviation example. Let us first analyze the widely misunderstood case of pricing crucial consumer goods in the aftermath of a natural disaster. Suppose a severe hurricane knocks out electrical power in an area with drastic consequences for, say, the ice market. Most of the ice in the area will melt, while, simultaneously, the demand for ice will surge way above its normal level. The drastic reduction in supply, combined with huge demand, will raise equilibrium price far above the norm, as depicted in Figure 2.9 where we see equilibrium price at $20 for a bag of ice that would normally sell for less than $2.

The natural, emotional reaction to this is to feel that sellers are engaged in "outrageous price gouging," but economic logic leads to a very different conclusion. In this case, the high price actually helps consumers to better deal with the emergency: the problem here is the hurricane; the high price helps people cope far better than a "normal, fair price" would.

The harsh reality is that the hurricane leaves the city with an ocean of demand for ice, but only a few drops of ice available. It is impossible to get the product to all who want it, so it is important to guide this crucial resource to those who have the most urgent needs. Everyone is thirsting for something cold to drink, but a few people have life-saving medicines that will spoil unless they are preserved with ice. The $20 price will convince most of those who are merely thirsty to leave the ice alone, while those who face a literal life-or-death need for ice will not hesitate to pay the exorbitant price. Thus, the high price

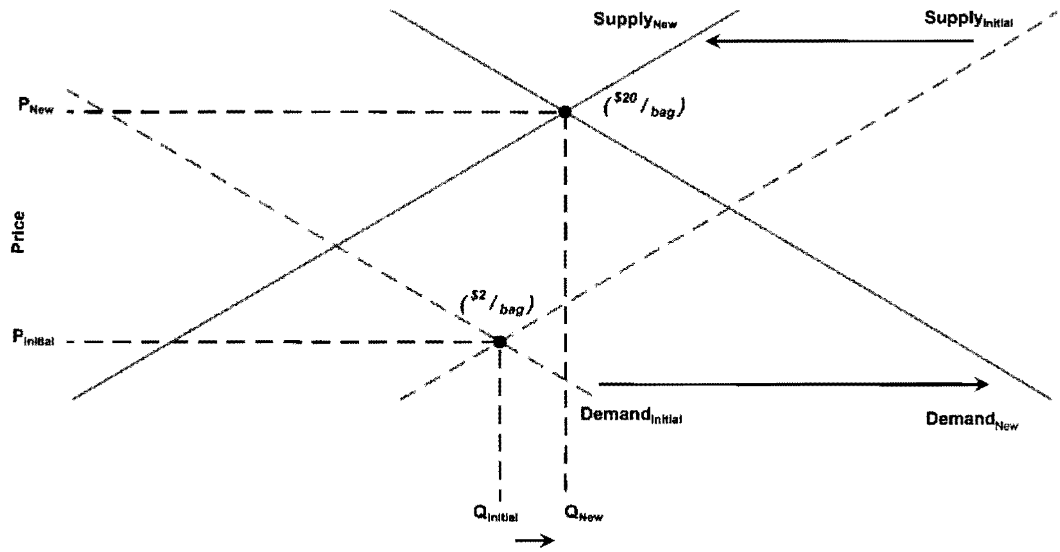

Figure 2.9 Movement of both supply and demand on the supply demand curve: as supply decreases and demand increases, equilibrium can only be reached at a much higher price

rations the good to those who need it most urgently. Of course, to truly minimize human suffering we need more than just a high price; we need also some charity to buy the ice for those who have urgent health needs but are too poor to afford to buy the ice themselves. The high price maximizes their chance of finding and purchasing ice in time to save lives. In fact, any philanthropist rushing in with ice would do more good by selling it for $20 than they would by giving it away randomly—ice selling at $20 is mainly going to those with urgent needs, whereas ice randomly given away is mainly going to thirsty people. Naturally, it would be ideal to give the ice only to those with urgent needs, but it is difficult to quickly identify those with greatest need under such chaotic conditions.

This seemingly unfair, outrageous price also motivates extreme measures to bring in more supplies. In such situations, young pilots have rented aircraft and flown in ice by helicopter and seaplane, a mode of transport normally unaffordable for most young pilots. But with a bag of ice going for $20, pilots of modest means could afford to rent the aircraft and bring in the life-saving supplies.

It would be nice if private charities or, perhaps, government officials were able to miraculously bring in enough ice to solve all problems. But this is inherently not the case—no consumer would pay $20 for something they could readily receive as a free handout from the Salvation Army or some government relief agency. The high price is conclusive evidence that charities and governments are overwhelmed, and an urgent response is needed from anyone capable of bringing ice in quickly.

It may be that some people bringing in the ice do so out of purely selfish motivation; unconcerned about saving lives and minimizing human misery, they are simply rushing in to "make a fast buck." We might fret about the soul of such a person, but if the ice they bring saves the life of a sick child by preserving her antibiotics, then that child is no less alive because she was saved by a selfish money-grubber rather than a generous philanthropist. When generous people can't do enough it's nice to have a high price to get everyone else motivated to help as well.

This whole scenario is an example of what the founding father of economics, Adam Smith, termed the "invisible hand." Voluntary trades in free enterprise often motivate behavior that helps society, even though individuals are mainly trying to help themselves and their own families. A pilot with some spare time looking only to enrich himself is guided, in Smith's phrasing, "as if by an invisible hand" to fly in ice that saves lives. Similarly, even a greedy consumer giving no thought to the crucial health needs of others will tend to leave the ice for more needy neighbors, simply because he refuses to pay such a high price. Working through the price system, the invisible hand defeats the greed of the consumer and redirects the greedy impulse of the pilot into a highly productive service to others. This invisible hand principle is the underlying foundation that makes individual freedom feasible and a free enterprise system so productive.

Invisible hand solutions also tend to be directly proportional to the problems posed. In the immediate aftermath of the hurricane, price is at its highest as the need for careful rationing and efforts at bringing in new supply is most crucial. As power begins to be restored demand will decrease, while the supply of ice increases. Both effects will reduce price so that people with lower priority needs will begin to buy ice. In a few weeks, if power is completely restored, price falls to a normal level, perishable medicines are back in refrigerators, and the typical use for ice is once again to chill drinks at parties. The high price is with us only as long as we need it.

Price Controls

It has been argued that the primary objective of price control is to prevent extreme, runaway inflation and all the evils that go with it, but some other negative consequences may happen as well. For example, on 15 August 1971, the US president, Richard Nixon, imposed price controls to contain inflation. A controlled price will allocate resources, but not in accordance with supply and demand. Suppose government decrees that the price of ice must be reduced from $20 to $2. The siren call of low prices is appealing to consumers anxious to get a bargain, but the result is tragic. As Figure 2.10 illustrates, the low price will drastically reduce quantity supplied, from Q^* to Q^s_g. Renting helicopters or driving refrigerated trucks from far away is no longer so affordable or appealing, with the price at an artificially low $2. As a result, the flow of ice slows to a trickle, making suffering and even death more likely because there is so little ice. At the same time, the quantity demanded now surges Q^d_g because price is depressed—almost anyone is thirsty enough to pay $2 to ice down some drinks. So what little ice is available is now mostly snapped up for casual use.

Perhaps the greatest irony is that the government price control results in consumers typically "paying" more for ice then the market rate of $20. This follows from the fact that there is more to life than cash; time, in a manner of speaking, is money, too. Figure 2.10 shows that the amount of ice available, Q^d_g, could be sold for P^*, which is obviously well above $20. If P^* is, say, $32, then we know consumers would pay that price to buy up all the ice. Normally, consumers compete for scarce products through price, but in this case errant government regulation precludes that. So, consumers compete instead by getting to limited supplies ahead of the crowd—that is, they arrive early and wait in line. Since

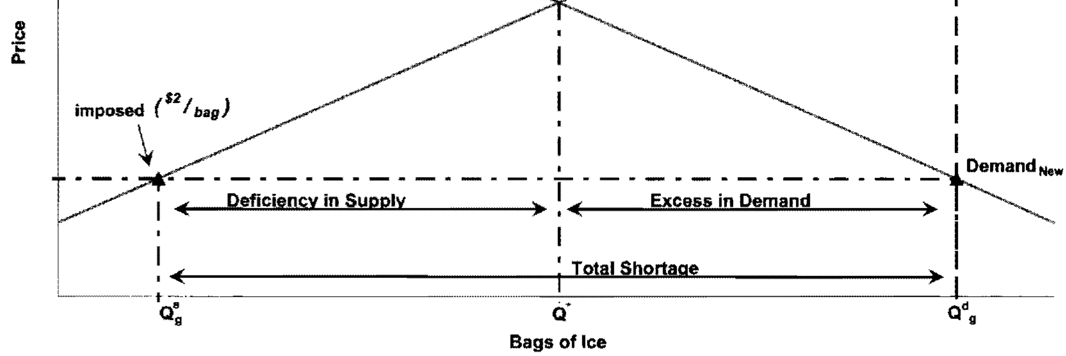

Figure 2.10 Supply shortage: when a price point is imposed on a price below equilibrium on the new supply curve, a supply shortage will occur

the cash price is artificially limited to $2, consumers are willing to use up an additional $30 worth in time. If the average ice consumer values their time at $10 per hour, then the average wait in line will be three hours. Of course, the chaotic uncertainty of the situation will result in some people waiting much longer or giving up altogether; others may luckily stumble into an unexpected delivery and face a much shorter wait. But the average price paid will be $30 worth in time plus $2 cash. One way or another, massive demand in the face of miniscule supplies will bid up price. In the absence of regulation, people pay $20 for ice; with the price control there is much less ice, and people pay a higher "price" for it—only now the greatest cost is time with a much smaller cash cost. But it's that total cost that counts and renders the price control a sort of "fool's gold."

Airport Landing Fees and Airport Congestion

Excessive and persistent airport congestion may be a function of price, the landing fee, if government keeps it too low. (Most of the world outside the United States is moving toward privatizing airports, but major private airports, such as Heathrow, typically have landing fees mandated by government.) Other factors, of course, may contribute to airport congestion—stringent environmental regulation that prevents airport expansion or archaic technology that forces aircraft to maintain wider separation, and so on. But regardless of other factors, the correct price can eliminate excessive congestion. The demand for airport use is, like the demand for anything else, a downward sloping curve; if we increase the landing fee, then fewer aircraft will use the airport and naturally congestion may be alleviated (see Figure 2.11).

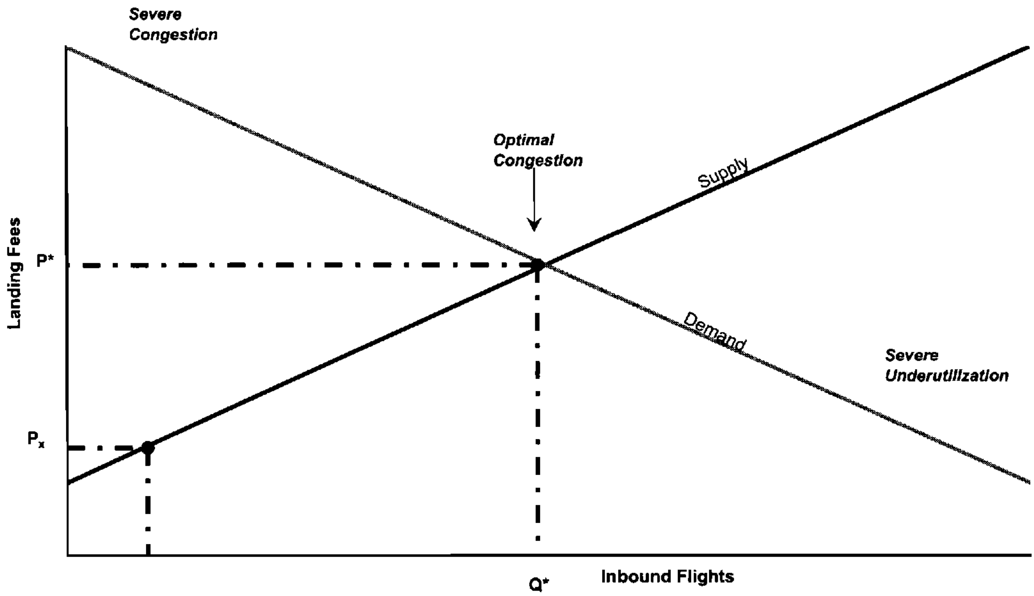

Figure 2.11 The impact of landing fees on airport congestion: increasing landing fees can be a solution to congestion, but it could also result in a deserted airport

Airports are different in nature: for example, demand for London Heathrow (LHR) may be less sensitive to price, more *inelastic,* than other airports. But, at a certain point, demand becomes more elastic and reduces the quantity demanded. If the price is too low, like P_x, then we get an excessive level of airport operations and therefore excessive congestion. To get to optimal congestion we need simply to raise the price to P^* so that we get optimal level of operations, Q^*. Note that optimal congestion does not mean that we achieve some sort of problem-free nirvana; it means simply that we maximize net benefits generated by airport operations, given the facilities we have. We may still wish for more reasonable environmental regulation that will allow airport expansion; we may still spend some time waiting to take off. Problems such as severe weather can still trigger excessive delays, but correct pricing will eliminate routine excess congestion.

This may become clearer if we carefully define optimal congestion. Every time an aircraft takes off or lands it generates both costs and benefits. As long as a given airport operation creates more benefits than costs than that operation should occur. We can define *marginal benefit* in this case as the benefit generated by the latest operation (landing or takeoff). Likewise *marginal cost* is the cost generated by that latest operation. Individuals inherently will engage in an activity if it generates marginal benefit to them greater than marginal cost, and avoid activity that generates marginal benefit to them less than marginal cost. But problems can occur if an individual's activity generates an *external cost*—that is, a cost that falls on someone else. We tend to discount, or even ignore, costs that we impose on others. Thus, external costs often result in excessive activity that generates more cost than benefit.

Consider an airline that is deciding whether or not to schedule a landing at LHR. The airline will consider only the costs that fall on them—the landing fee, extra use of fuel as the aircraft circles while waiting its turn to land, lower productivity of the aircraft and crew because it takes so long to get in and out of the airport, and so on. The airline will ignore the fact that its use of the airport will force other aircraft to wait longer, to burn more fuel, suffer their own productivity losses, and so on. This is how we end up with excess congestion if government keeps the landing fee too low.

In Figure 2.12 we see two different marginal cost curves. The marginal cost system-wide curve embodies all costs, including what an individual airline would view as external costs. As the number of landings/takeoffs increases, congestion problems worsen, so, at some point, this system-wide marginal cost increases; thus, we see the curve slope upward. The marginal benefit curve is essentially a demand curve. An aircraft with many high-paying customers whose final destination is London would enjoy a high benefit from landing at Heathrow and would be at the highest point on the marginal benefit curve. An aircraft with a few holidaymakers on their way to some other destination would generate a small marginal benefit from landing at Heathrow and would be at the low point of the marginal benefit curve. Many aircraft between these extremes would generate marginal benefits and would lie between them on the marginal benefit curve.

As more aircraft use the airport, the marginal cost system-wide and marginal benefit curves move nearer to each other until they equalize at optimal output level Q^*. Again, operating at Q^* does not mean that there are no congestion problems; it merely means that the problems are worth enduring, given the benefits. (This is barely true for the very last aircraft added to the mix right before Q^* is reached.) However, if another landing is added after we reach Q^*, that flight will generate more cost than benefit. Anything beyond Q^* results in excess congestion.

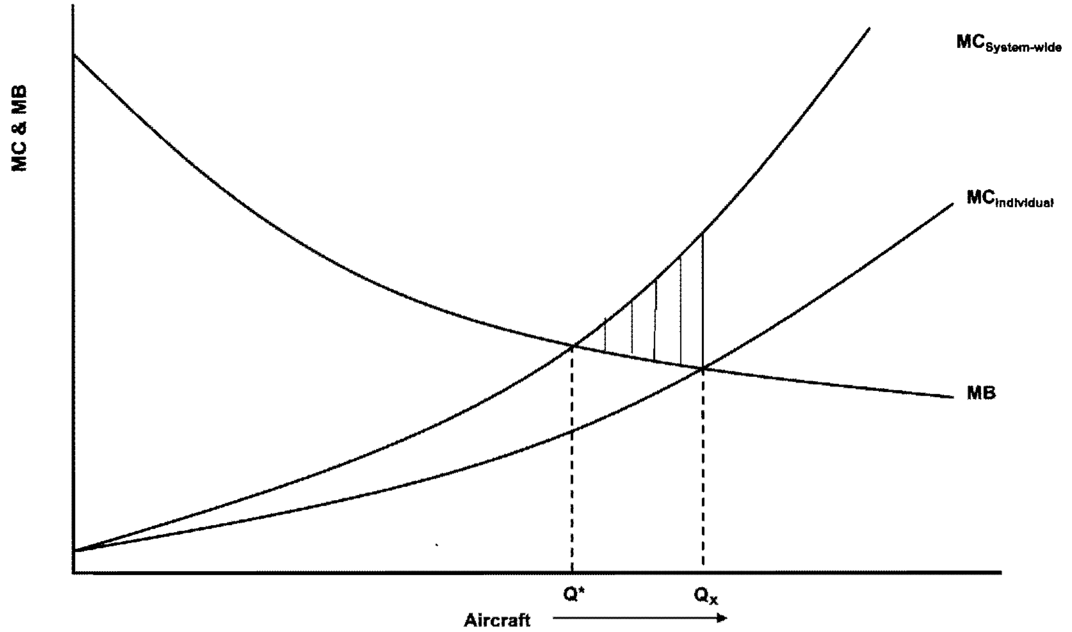

Figure 2.12 Marginal cost system-wide curve, marginal cost individual aircraft curve, and marginal benefit curve

Looking at Figures 2.11 and 2.12 together, if government allows the landing fee to be at P^*, then operations settle at Q^*; optimality is achieved. However, governments seldom, if ever, get this right. A typical landing fee is likely to be equal to P_x.

The marginal cost curve for the individual aircraft does ignore the additional congestion costs that are imposed on others. At a price of P_x the individual marginal cost lies below system-wide marginal cost. Aircraft operating beyond Q^* do more harm than good, but since some of that harm falls on others, the cost is less *for that aircraft* than the benefit. Eventually, the congestion problems get so bad that no one else chooses to use the airport, even though those external costs are ignored; the number of operations settles at Q_X. The shaded, triangular area in Figure 2.12 equals the net loss of wealth caused by the price being too low. Costs here include: wasted fuel, higher aircraft maintenance, lost productivity of aircraft and airline employees from excessive time spent getting into and out of the airport, and the wasted time of air travelers. Note that this last rebounds to airlines in the form of lower revenues because when you waste people's time you can't charge them quite as much. Conversely, if you could get people to their destination quicker, especially busy business travelers, they would be willing to pay higher fares.

If the landing fee is increased, then the individual marginal cost curve shifts leftward since each individual aircraft now experiences higher cost. If we raise the landing fee to P^*, then marginal individual cost and marginal system-wide cost become the same curve, operations settle at Q^* and we're back to optimality. Conceptually, this is an easy thing to do, at least approximately, but the political nature of government decision-making has made proper pricing virtually impossible for many decades. We are stuck with excess congestion costs thus far because governments generally are incapable of pricing correctly (McEachern, 2006).[1]

1 The issue of privatization of airports is discussed in more detail in the second half of Chapter 5.

THE ECONOMICS OF GOVERNMENT

We do not wish to create the impression that all governments are so hapless that government activities always generate more harm than good. Virtually all economists agree that government, though very imperfect, has a crucial role to play in at least establishing and protecting property rights. This entails, at minimum, government authority over the military, police, criminal courts and property right issues such as pollution and patent laws.

However, it is natural to wonder, and worth explaining, why government policy sometimes deviates so radically from economic optimality. Why, for example must wealth be destroyed and people caused to suffer and even die because politicians can't apply basic economic logic to pricing decisions?

Economist James Buchanan won a Nobel Prize for his work addressing this very issue. Buchanan and his frequent collaborator, Gordon Tullock, established a branch of economics, *Public Choice*, which analyzes government decision making (Buchanan and Tullock, 1962). Let us consider a few key insights from this field.

Incentives for a Voter to be Well Informed

It costs a lot to become a well-informed voter, not in terms of money but in terms of time. Most voters are, for instance, fully capable of gleaning all the major policy insights that a study of economics can offer. But in this busy world, the vast majority of voters simple can't find the time to educate themselves in this area.

On the benefit side, it is tempting to assume that a voter who does become well-informed will benefit by "getting better government." But that assumption is unrealistic. Collectively, voters control electoral outcomes. But the chance of a single voter affecting an electoral outcome, especially at the national level, is approximately zero in virtually all cases. This harsh truth is unchanged even if we include the typical individual's actions to influence other people's votes. The unpleasant reality is that an average voter cannot expect to affect important political outcomes; collectively the votes matter, but an individual vote is essentially symbolic. Most people, including the authors, would agree that there is a moral obligation to be a well-informed voter. Most of us enjoy a sense of fulfilling a civic duty when we take a least a little time to study a political issue or candidate before voting. However, this sense of civic duty is, alas, not extremely strong.

Rather than spend a few more hours a week studying political issues in order to cast a better-informed vote, virtually all of us will instead choose to apply those few hours to something that will make a tangible difference in our lives. A few more hours working eases financial pressures, a few more hours spent with our spouse strengthens our marriage, a few more hours spent at leisure refreshes us, or a few more hours studying improves our grades. Time spent in this manner accomplishes something, whereas casting a wiser vote normally has no tangible impact. Economists refer to this as *rational political ignorance* (Downs, 1957). It is rational—meaning logically consistent, given our preferences—to remain ignorant when the cost of acquiring knowledge is greater than the benefit. We are not saying that such ignorance is morally good or in any way desirable, merely that it is a systematic and understandable problem, given the incentives people face.

Public education is sometimes successful in exposing popular myths and erasing ignorance in other areas. For example, the educated world is well aware that, despite all casual appearances, the sun does not revolve around the earth, nor is the earth flat.

Unfortunately, public education has generally failed to inform the citizenry concerning economic policy. Economic regulations widely condemned by economists are often extremely popular. In the world of popular economic beliefs most voters are quite convinced that the world is flat and that their earth is indeed the center of the universe.

So, why do so many governments around the world impose harmful price controls? Because such policies are extremely popular with rationally ignorant voters who simply want lower prices and don't study, or even carefully think through, the consequences of translating impulsive wishes into uncompromising law. Voters, and hence the politicians who must be elected by them, often tend to be driven by the passions of the moment. Good politics is often very bad economics.

Undue Special-interest Influence

Government may also go astray because politicians are sometimes motivated to help certain groups of people, at the expense of the overall economy, in order to, in essence, buy their votes. We will see later, for example, that economic analysis indicates that most, if not all, regulations inhibiting international trade, including those relating to aviation, do far more harm than good. However, a politician may be able to buy the votes of employees working for domestic companies if the politician supports regulation that prevents efficient foreign-based companies from entering the domestic market. Consumers will typically lose far more from the resulting higher prices (this might more fairly be termed "price gouging"!) than the protected employees and major stockholders will gain; the nation is made poorer, but the politicians have their votes. Of course, undue special-interest influence would occur less frequently if voters recognized such corruption for what it is. In other words, rational political ignorance frequently lies at the heart of this problem as well.

Bureaucratic Inefficiency

Airlines owned and operated by government are, on average, substantially less efficient than private airlines. In fact, economists would generally agree that the facts support the conclusion that almost any private organization tends to be more efficient than any comparable government counterpart. One key reason for this is the huge difference in incentives faced by public versus private employees. Employees in a private airline are painfully aware, for instance, that they must satisfy paying customers or else they will soon be out of a job. Employees in government generally count on a lifetime of employment, knowing that taxpayers will likely be forced to subsidize them regardless of how satisfied anyone is.

Government employees are also much more heavily regulated than private workers, often lacking the authority to take actions necessary to solve problems. Evidence suggests that government workers are often overpaid, but there are certain higher-level jobs where the pay seems too low. Supply and demand logic applies in labor markets as well—low pay creates a shortage of workers; managers will not be able to attract and maintain qualified workers. The solution to such a problem is simple and routine in the private sector: increase pay. But government managers are not free to adjust pay to market conditions; pay rates are fixed by strict formulas. If these managers were free to use their own judgment in setting pay for government employees there would be nothing to stop them from paying massive salaries to friends, and to themselves!

On the other hand, private managers face no regulatory barriers to increasing pay and therefore generally avoid serious labor shortages altogether. We know that if, say, a private airline decides to pay employees exorbitant wages and then tries to pass on the inflated labor costs to customers, the customers will refuse to fly with the airline. Where voluntary trade exists, where customers are free to walk away, the invisible hand provides crucial consumer protection. Where taxpayers are compelled to fund bureaucracies there is no such protection; regulation is the only potentially effective way of preventing government workers from greedily enriching themselves at the taxpayers' expense. The same issues present themselves in procuring supplies and all areas of operations—regulation is the only feasible, though still very imperfect, check on government abuse of power.

Pervasive regulation, along with weaker incentives, is a way of life in government bureaucracies. Bureaucratic inefficiency is the natural result.

Government Failure versus Market Failure

Both government and free enterprise are imperfect institutions. The private sector may have less than ideal levels of competition, may fail to provide *public goods* that benefit society broadly but are difficult to charge customers for (such as national defense), may produce less than ideal distribution of income and may fail in various other ways. As mentioned earlier, virtually all economists agree that government should provide for national defense, police, criminal courts, environmental protection, and in general establish and protect property rights. No one expects any government to handle these tasks perfectly, but there is general agreement that the inevitable government failures in these areas will be less severe than the problems that would arise if government did not intervene. Most economists probably favor a good deal of government intervention beyond these basics.

On the other hand, economists also generally agree that the private provision of most goods will be more efficient than government production, that prices should normally be set freely by private companies, and that the best consumer protection is often competition in an open market. Too much government control harms an economy, but a society organized only through voluntary agreements with no government coercion is also impractical. Economists favor a *mixed economy* where most production and resource allocation is directed through the price system of private enterprise, but where government provides some crucial services and is an active regulator in certain areas of the private sector.

Exactly how much free enterprise and government control should be in this mix is a matter of some controversy. However, we do have extensive data that sheds some light on the issue. Countries vary significantly in the degree of free enterprise they allow in their economies, and there are sometimes significant changes in the degree of free enterprise in a given country over time. A number of authors have extensively studied this history and examined how economies perform under various mixes of government control and economic freedom.

Figure 2.13 shows the strong correlation between a lesser degree of government control and higher per person income (higher per capita real GDP).

Higher economic freedom is defined as less government intervention—less government ownership of the economy, lower taxes, lower government spending, less government regulation, less interference in free international trade and less currency inflation. In Figure 2.13 countries are divided into quintile rankings of economic freedom. Countries in the top quintile are in the top 20 per cent for economic freedom, exhibiting less government control

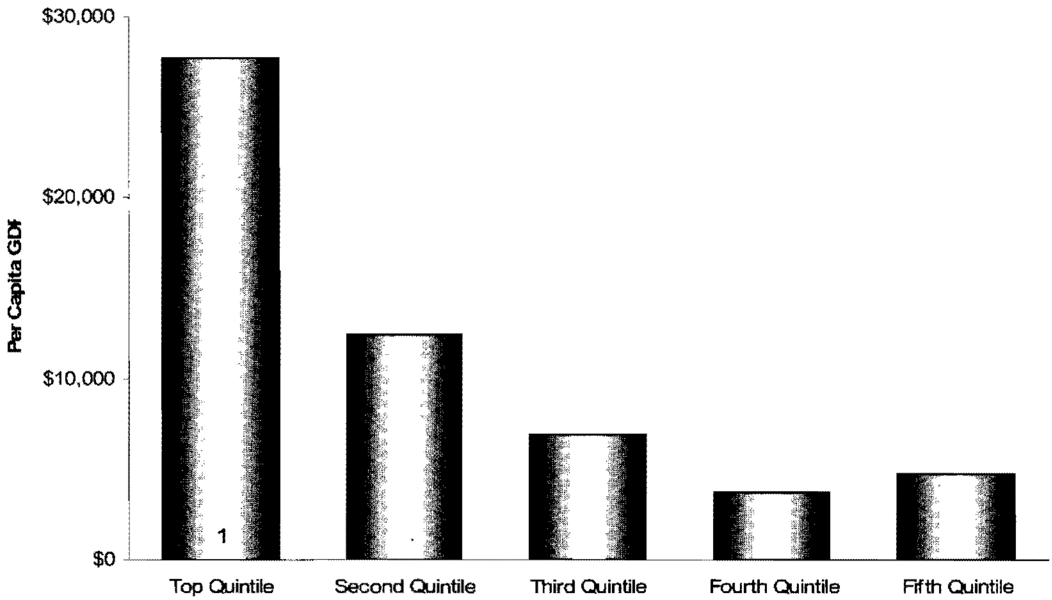

Figure 2.13 Per capita GDP values in countries ranked in descending order of economic freedom. The correlation shows that per capita GDP is directly proportional to economic freedom

Source: Kane *et al.*, *2007 Index of Economic Freedom* (published by the Heritage Foundation).

than the other 80 per cent of countries. The second quintile is composed of countries with substantially more government intervention than the top quintile but less than the remaining 60 per cent of countries, and so on for the other quintiles. As you can see, income falls dramatically as government power over the economy rises. This result holds across many decades and, allowing for adjustments of many other factors, the data clearly support the view that most countries today probably suffer from too much government control of their economies.

Though very robust, such data do not, of course, end the debate. These studies attempt to adjust for all other relevant factors, but this is always difficult to do perfectly. Also, even a country suffering from too little economic freedom generally might be in need of more government intervention in a particular area. Furthermore, this historical record does not have clear implications for countries that already rank comparatively high in economic freedom. For example, only a few countries consistently enjoy more economic freedom than the United States—principally the islands of Hong Kong and Singapore—so we don't have extensive examples to follow of more economically free nations outperforming the US economy.

Nevertheless, it is fair to say that most countries would probably benefit from more economic freedom; there is a fairly general trend in that direction, particularly within aviation.

Reforming Government

Most alleged government "reforms" accomplish little or nothing because they do not fundamentally address rational ignorance, bureaucratic inefficiency, or even special-interest problems. However, there is a set of reforms, especially prevalent in aviation, which does—privatization.

Privatization comes in various forms, but essentially entails delegating some degree of operations to the private sector; government establishes private–public partnerships and transforms itself into more of an overseer and less of an operator. Even the purest form of privatization does not remove all government influence. Airlines, for instance, are increasingly run as private businesses rather than government entities. However, every private airline in the world faces a myriad government regulations relating to safety, personnel, and finances.

Moreover, even if a state-owned airline is sold, it is easy to attach special regulations to the newly privatized firm. It is politically essential, for instance, to require that newly privatized firms not have any layoffs, at least for several years. Otherwise, public employee unions tend to successfully sabotage any movement away from government ownership. Many airports have been privatized, but, generally, special government regulations, for better or worse, continue to control landing fees. The broader point, of course, is that it is not necessary for politicians to surrender all control to the private sector in order for us to enjoy at least some of the increased efficiency that free enterprise can bring.

It is also not necessary to eliminate government funding to obtain gains from privatization. The mildest form of privatization is to retain government ownership and funding, but to *subcontract* at least some operations to private firms. Thus, if it is not possible to sell an airport outright, the government may retain ownership of an airport but hire a private firm to run it, as is the case in Burbank and Indianapolis (Vasigh and Haririan, 2003). This can virtually eliminate bureaucratic inefficiency since personnel, procurement, and other decisions are now made by a business rather than by politicians. Also, if an airport is not well run, the firm can be fired and a new one hired—unlike the situation in government where a poorly performing bureaucracy is simply given more money to "fix problems." If complete subcontracting is not politically feasible, then at least some tasks—janitorial services, maintenance, or whatever—can be contracted to private firms.

Of course, privatization, like all human endeavors, is also imperfect. Subcontracting, in particular, has substantial potential for government corruption; since politicians ultimately choose which company gets the contract, they may select less efficient providers to win political favor or in exchange for bribes. In the United States, defense subcontracting always seems to be embroiled in some alleged scandal. But the only alternative to hiring private firms to build defense systems is to hire a full-time government bureaucracy to do it instead. This alternative, apparently, is unanimously rejected. It is understood that establishing a permanent government monopoly would be worse than contracting work to competing private firms, whatever problems there may be in subcontracting.

Government has a role to play in enforcing certain regulations, in subsidizing certain activities. But, given the problems stemming from rational political ignorance and bureaucratic inefficiency, we can see why government is usually not an efficient producer of anything. It is important for the general economy, and aviation in particular, to limit government influence and to allow the private sector to handle the operations for which it is so well suited.

SUMMARY

This chapter introduced the economic way of thinking through the principles of supply and demand. These principles were then used to describe the process by which market equilibrium is achieved. The market equilibrium price was then shown to be the maximum utility position from an economic point of view. Examples from the aviation industry were then presented to illustrate the concept of equilibrium. Finally, the economics of government control and regulation were discussed and analyzed within the context of the incentives that exist in a market system versus a government setting.

REFERENCES

Agin, D. (2006). *Junk Science: How Politicians, Corporations, and Other Hucksters Betray Us.* New York: Thomas Dunne Books.

Buchanan, J.M., and Tullock, G. (1962). *The Calculus of Consent: Logical Foundations of Constitutional Democracy.* Ann Arbor, MI: University of Michigan Press.

Colander, D.(2004). *Microeconomics* (5th edn). New York: McGraw-Hill.

Downs, A. (1957). An Economic Theory of Democracy. New York: Harper and Row.

Hirschey, M. (2006). *Managerial Economics* (11th edn). Mason, OH: South-Western.

Kane, T., Holmes, K., O'Grady, M.A. (2007). *2007 Index of Economic Freedom.* Washington, DC: The Heritage Foundation.

Mankiw, N.G. (2007). *Principles of Microeconomics* (4th edn). Mason, OH: South-Western.

McEachern, W. (2006). *Economics: A Contemporary Introduction* (7th edn). Mason, OH: South-Western.

McGuigan, J., Moyer, R., and Harris, F. (2008). *Managerial Economics: Applications, Strategies, and Tactics* (11th edn). Mason, OH: South-Western.

Rubner, A. (1979). *The Price of a Free Lunch: The Perverse Relationship between Economists and Politicians.* London: Wildwood House.

Swartz, T., and Bonello, F. (2003). *Taking Sides: Clashing Views on Controversial Economic Issues.* New York: McGraw-Hill.

Vasigh, B., and Haririan, M. (2003). An Empirical Investigation of Financial and Operational Efficiency of Private versus Public Airports, 8(1), *Journal of Air Transportation* , pp. 225–36.

3

Market Demand Analysis and Demand and Supply for Airline Services

Economics has many substantive areas of knowledge where there is agreement but also contains areas of controversy. That's inescapable.

Ben Bernanke

This chapter builds on the demand and supply concepts introduced in the previous chapter by examining airline demand in much more detail. Whereas Chapter 2 gave a general picture of the laws of demand and supply, this chapter will cover the more technical and quantitative aspects of airline demand and supply. More specifically, airline demand analysis is concerned with understanding passenger behavior, measuring and characterizing the airline response to a change in ticket prices or incomes, and deriving the demand-side information necessary to make sound business decisions. On the other hand, airline supply refers to airlines' ability and willingness to provide a specific number of seats at alternative prices in a given time period in a given market.

The chapter *combines theory, applications, and exercises to illuminate the importance of the theory of demand and supply for airline companies. Other topics include: elasticity of demand and its managerial application;* different types of elasticity related to passenger demand, including how they react to changes in economic environment; and how to calculate elasticity and the various determinants of price, income, and cross-price elasticities of demand. A full list of topics covered is as follows:

- Basis for demand
- Demand schedule
- Demand curve
- Demand function
- Determinants of demand for air transportation
- Characteristics of demand for air transportation
- Source of demand
- Elasticity of demand, including:
    - Price elasticity

  - Cross-price elasticity
  - Income elasticity
- Supply of airline services
- Factors affecting supply of airline services
- Characteristics of supply for airline services
- Airline supply and demand equilibrium.

BASIS FOR DEMAND

As a more formal definition of demand, we can say that demand is the ability and willingness to buy specific quantities of a good or a service at alternative prices in a given time period under *ceteris paribus* conditions.[1] Understanding demand theory and the demand function is one of the more important aspects for any business, since sales are the reason for the company's existence. In the airline industry, demand is usually assessed in terms of the number of passengers, revenue passenger miles (RPMs), or revenue ton miles (RTMs).[2] For aircraft manufacturers, demand would be represented as the number of aircraft sold. Although demand varies from industry to industry, its characteristics remain similar, and its importance to business is always high. Therefore, it is critical to fully understand the nature of demand.

DEMAND SCHEDULE

As stated earlier, the law of demand states that, *ceteris paribus*, as price increases, the quantity demanded decreases. In order to understand the law of demand's practicality, consider a transcontinental flight from New York to Los Angeles. For this round-trip flight, what would be the maximum a passenger would be willing to pay? $500? $1,000? $5,000? The fact is that, at some price, the passenger would consider it too expensive to fly and not take the trip. This decision to not fly is the law of demand in practice, that is, at some price the quantity demanded will decrease.

When this decision is presented to all possible travelers there will be different responses because people have different purposes for travel and varied incomes. These different responses help create a demand schedule, which is simply a table showing the quantities of a good that customers are *willing and able* to buy at alternative prices in a given time period, *ceteris paribus*. Such a table outlines the number of customers who would purchase a product or service at the given price. It is important to remember that demand is cumulative—that is, a consumer who is willing to pay $1,000 for the flight would certainly also be willing to pay $500 for the same flight. Table 3.1 provides a hypothetical demand schedule for the New York to Los Angeles flight.

The demand schedule contained in Table 3.1 highlights the law of demand, since the quantity demanded (number of passengers) for the $200 airfare is significantly more than the quantity demanded for the expensive $5,000 airfare. At an airfare of $5,000 very few people are willing and able to pay for the ticket—perhaps only a few extremely wealthy and/or time-sensitive passengers.

---

1 *Ceteris paribus* is a Greek term for all else being equal or for everything being held constant.

2 A revenue passenger mile (RPM) represents one seat occupied by a revenue generating passenger who is carried one mile. A revenue ton mile (RTM) represents one ton of revenue cargo carried for one mile.

Table 3.1 Hypothetical demand schedule for the New York to Los Angeles flight

Ticket Price of New York to Los Angeles Round Trip	Quantity Demanded (No. of passengers)
$200	735
$500	690
$1,000	615
$2,000	465
$3,000	315
$5,000	15

DEMAND CURVE

From the demand schedule above, a demand curve can be constructed. A demand curve graphically describes the demand schedule and the quantities of a good that customers are willing and able to buy at alternative prices in a given time period. The law of demand can then be derived from this schedule, and it states that the demand curve will always be downward-sloping.[3] It should be noted that the demand curve does not portray actual purchases, but only what consumers would be willing and able to purchase. Figure 3.1 provides the demand curve for the New York to Los Angeles trip. In this example, the demand curve is a linear negative sloping line.

In order to highlight the effect of the ticket price on the demand curve, refer to Figure 3.3. A change in the ticket price is always defined as a *ceteris paribus* movement along the demand curve. For example, if the current ticket price for the flight was set at $3,000 and then lowered to $2,000 (all other things remaining constant), the only effect this change would have would be on the quantity demanded. This is reflected in the demand schedule shown in Table 3.2 in column 2 where the quantity demanded moves from 315 to 465. This movement is displayed graphically in Figure 3.3, where the quantity demanded moves from point A to point B along the demand curve.

DEMAND FUNCTION

Using information from the demand schedule and/or the demand curve, a demand function can be constructed. A demand function is simply the functional relationship between the quantity demanded and factors influencing demand. There are numerous factors that influence the demand for air travel, for example, in the New York to Los Angeles market, and therefore price is not the sole determinant of demand. (However,

3 Stated exceptions to the law of demand usually involve confusion between the perceived quality and/or prestige that a high-priced good confers on the purchaser. A simple thought experiment confirms this. Imagine that the exact same quality and/or prestige could be achieved at a lower price. Rational consumers would always select the lower-priced good for the exact same quality and/or prestige. Therefore it is the quality and/or prestige that the good has, and not the high price, that the consumer is responding to.

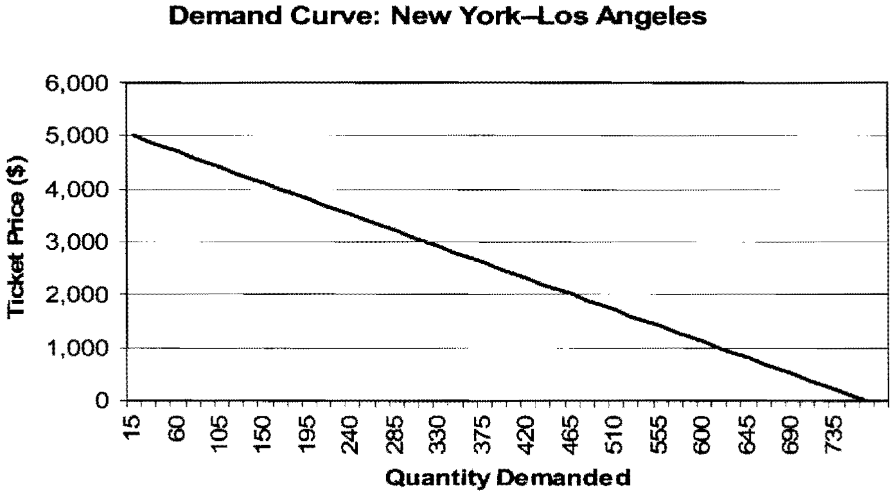

Figure 3.1 Demand curve for the New York to Los Angeles flight

when the other determinants are held constant, the relationship between price and quantity demanded is always negative.) These functions can be classified into two categories based on their composition: *implicit* or *explicit*. Implicit demand functions simply state a general relationship between the quantities demanded and the factors affecting demand. Explicit demand functions are mathematical relationships between the quantity demanded and the various variables impacting on demand. Implicit functions do not have the actual mathematical relationships, but rather a more generalized statement of the factors affecting demand. For example, the implicit demand function for the New York to Los Angeles flight could be:

$$D_{NY\text{-}LA} = f(P_X, P_Z, Y, H)$$

where: P_X is the own ticket price
P_Z is the competition's ticket price
Y is the annual income or state of the economy
H is a composition of other factors, such as service, customer loyalty, and random factors.

The implicit demand function simply states that a relationship exists between the dependent variables and the independent variables, but it does not state the extent to which the variables are related. The numerical relationship that displays the degree of influence that each factor has on the quantity demanded is the explicit demand function. Using the information obtained from the demand schedule and the demand curve, an explicit demand function for the New York to Los Angeles flight could be written as:

$$D_{NY\text{-}LA} = 15000 - 2P$$

Based on this linear demand function, two statements can be made about the nature of demand for a round-trip New York to Los Angeles flight. First, when the price of the ticket is $7,500, then demand for this trip drops to zero. In other words, demand for any product or service is always limited by the extent of the market demand (in this case, any price above $7500).

The other statement concerning the linear demand curve is that the negative price coefficient (or slope) of –2, which means that for every dollar increase in the ticket price, the demand drops by two passengers. This change in demand occurs for all price points, creating a constant negative slope.

While the linear demand function is simplistic and clear to understand, the fact remains that the demand schedule rarely has a perfect linear form. This makes intuitive sense since passenger demand does not drop off evenly with price increases, but instead drops off in steps. Often there is a major inflection point where demand decreases dramatically.

Two other types of demand function are the semi-log function and the log-linear function.[4] The general forms for the two functions are:

Semi-log: $LnQ_D = \beta_o + \beta_1 P$

Log-linear[5]: $LnQ_D = \beta_o + \beta_1 LnP$

where: Q_D is the quantity demanded,
P is the ticket price, and
β_o and β_1 are the coefficients.

For the airline industry, it is usually assumed that the typical demand function takes the log-linear shape. The general shape of a log-linear function is illustrated in Figure 3.2.

Unlike a linear demand function where the slope of the function is constant throughout, the slope of a log-linear demand function changes. In Figure 3.2 the initial slope of the function is fairly steep, indicating that a unit drop in the ticket price does not generate a similar increase in the quantity demanded. This is partly a result of the fact that the ticket price is still considered expensive by the majority of potential customers. Eventually, as

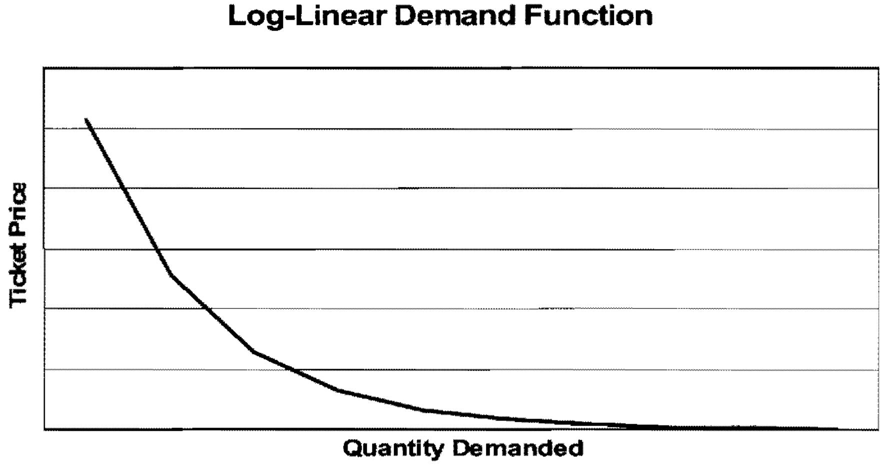

Figure 3.2 Conceptual log-linear demand curve

4 The log-linear function is also sometimes referred to as the log-log function

5 Note that for log-linear demand functions, the elasticities of demand are the coefficient of the variable and are constant. For instance, the price elasticity of demand would be simply β_1.

the price decreases, the quantity demanded for the product or service becomes greater and greater to the point where a small drop in the price generates a large increase in the quantity demanded.

DETERMINANTS OF DEMAND FOR AIR TRANSPORTATION

While price is a major determinant of demand for any product or service, it is not the sole determinant of demand for air transportation. There are many other factors that affect demand for air transportation; however, discussing each factor in detail could consume an entire book. Realistically there are only a few major determinants that affect demand for air transportation, and these will be explored in this section. They include:

- ticket price
- competitor's ticket price
- passenger income
- state of the economy
- availability of other modes of transportation
- customer loyalty
- in-flight amenities
- frequency of service
- safety
- random factors, such as SARS and 9/11 or the threat of terrorism.

To illustrate the impact that these factors might have on the demand curve, consider the previously used demand schedule for the New York to Los Angeles flight. Table 3.2 provides an updated version of that demand schedule.

In order to highlight the effect of the ticket price on the demand curve, refer to Figure 3.3. As we said earlier, a change in the ticket price is always defined as a *ceteris paribus* movement along the demand curve. For example, if the current ticket price for the flight was set at $3,000 and then lowered to $2,000 (all other things remaining constant), the only

Table 3.2 Updated demand schedule for the New York to Los Angeles flight

Ticket Price of New York to Los Angeles Round Trip	Quantity Demanded (No. of passengers)	New Quantity Demanded (No. of passengers)
$200	735	960
$500	690	915
$1,000	615	840
$2,000	465	690
$3,000	315	540
$5,000	15	240

effect this change would have would be on the quantity demanded. This is reflected in the demand schedule shown in Table 3.2 in column 2 where the quantity demanded moves from 315 to 465. This movement is displayed graphically in Figure 3.3, where the quantity demanded moves from point A to point B along the demand curve.

Ticket price is the only determinant of demand that causes a movement along the demand curve. Changes in the other determinants of the demand cause a shift in the entire demand curve. Continuing with the New York to Los Angeles example, suppose that the competition increases its airfares. This will cause demand for this airline's flight to increase as some consumers switch to this airline. This increase in demand is reflected in Table 3.2's third column, where the quantity demanded at each price level increases by 225 passengers from column 2. The increase in the quantity demanded across all price levels creates a rightward shift in the demand curve from D_1 to D_2, as depicted in Figure 3.4. For every price level the quantity demanded has increased by 225. Conversely, a negative impact on demand would create a leftward shift in the demand curve.

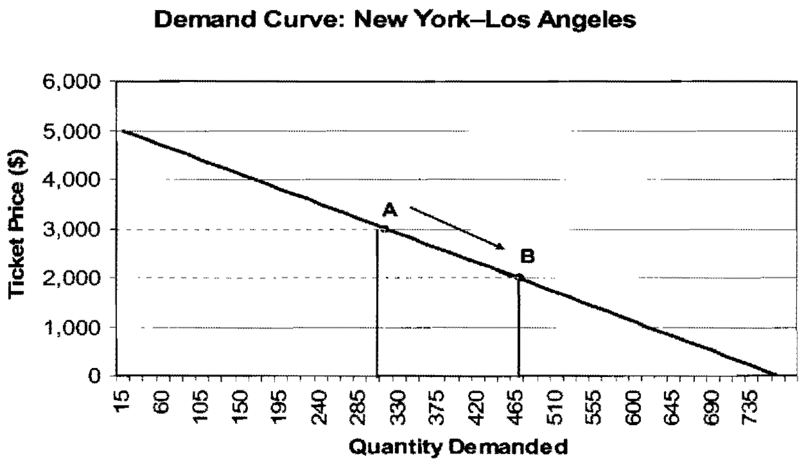

Figure 3.3 Change in quantity demanded

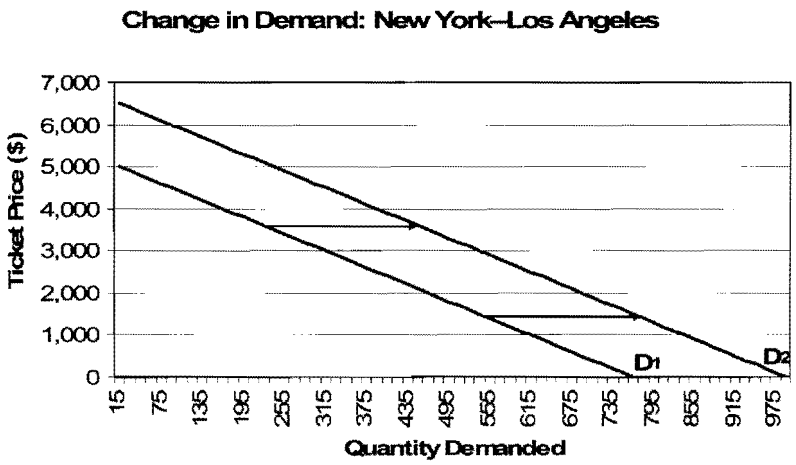

Figure 3.4 Change in demand

The demand for a single flight is affected by multiple factors, but for many, especially price-sensitive leisure travelers, the price of the flight and the price of competing flights are probably the most important variables. With the advent of the Internet, where airline ticket price information is readily available to potential consumers, price and competitors' prices have probably become even more important. However, these price variables affect different segments of the population. For time-sensitive travelers, ticket price versus a competitor's ticket price may not be as important as for price-sensitive travelers.

For the competitor's price variable, the coefficient would be expected to be positive, since an increase in the competitor's price would make the competition's product less competitive, and ultimately increase demand for the company's product. Consider airlines A and B which both have flights on the same route and have introductory fares of $400. If airline B were to increase its airfare to $450, this would make it less competitive and ultimately increase demand for airline A. Conversely, if airline B were to drop its airfare to $350, this would make it more competitive, which would decrease demand for airline A and increase demand for airline B. Regardless of the viewpoint, the price coefficient would be negative and the cross-price coefficient would be positive. The general form of a log-linear demand function can be expanded to include the competition's price:

$$LnQ_D = \beta_o + \beta_1 LnP_X + \beta_2 LnP_Z$$

where: β_o is a constant
P_X is the own price and
P_Z is the competitor's price
β, β_2 and β_3 are the coefficients.

As was discussed in Chapter 2, income is another important determinant of demand. Consumers with higher incomes are able to purchase more goods and services; therefore, an increase in disposable income will provide an increase in demand for air travel. In addition, increased consumer income is usually correlated with increased business activity, indicating a higher demand for business travel. Because of this direct relationship between demand and income, the coefficient for consumer income is positive. The general log-linear demand function for air transportation can now be expanded to include income:

$$LnQ_D = \beta_o + \beta_1 LnP_X + \beta_2 LnP_Z + \beta_3 LnY$$

where: β_1, β_2, and β_3 are the coefficients

Although own price, substitute price, and income are the three major determinants of demand for air transportation, there are other factors that affect demand. An important one of these is the availability of substitutes. For air travel, this includes other modes of transportation. In the United States, driving is a reasonable substitute for many short flights, while in Europe high-speed rail can greatly affect the demand for select air services. For example, the high-speed Brussels to Paris rail line has created a situation where there is little demand for Brussels to Paris air transportation. However, where there is a lack of other modes of transportation—for example, air services to a remote Caribbean island—demand for air travel can be expected to increase. On the basis of this discussion, the expected coefficient for the availability of substitutes could be either positive or negative.

The final pricing variable that impacts on the quantity demanded is the price of a complementary product or service—that is, a product or service that is usually used jointly with the primary good. A good example of a complement for the airline industry is the hotel and rental car industries. Since many leisure and business travelers have to stay in hotels while on their trip, the price of the hotel will affect the demand for air travel. This is a result of demand for air transportation being derived (that is, being generated by) from something other than air travel). For example, if the average price of a night's stay in Cancun were to increase, fewer people would want to take a vacation in Cancun. Therefore, the demand for flights to Cancun will be reduced. On the basis of this typical example, one would expect the coefficient for the price of a complement to be negative.[6]

In a competitive market environment, the frequency of an airline's service between two cities will also affect demand. Flight frequency is especially important for business travelers since they are generally more time-sensitive than leisure travelers. An airline with several flights between two cities has a greater probability of meeting a travelers' schedule than an airline with only a few flights. Moreover, a robust flight schedule provides the traveler with greater flexibility in case of schedule changes. This is a primary reason why regional jets have become more popular—they enable airlines to provide increased flight frequency while holding steady the total number of seats offered in the market. We would therefore expect a positive coefficient between flight frequency and demand.

Whether an airline has a nonstop or connecting service between two cities will also affect demand. The availability of a nonstop flight will generally increase demand because passengers usually prefer nonstop flights over connecting flights. However, this assumption may not apply to all markets, especially ultra-long-haul markets in which passengers may appreciate a stopover. Therefore, we would expect that a nonstop flight variable would generally have a positive relationship with demand (with some possible exceptions on long-haul flights).

Customer loyalty is another key determinant of demand for air transportation. In this regard, one of the more successful marketing tools that the airline industry has implemented has been loyalty or frequent-flyer programs. By offering free flights and perks for loyalty, airlines have been successful in obtaining repeat business, especially among business travelers.

While customer loyalty is important, it is also important to attract new customers. Here, airlines have emphasized service which can incorporate a host of factors, including aircraft seat placement (more room), in-flight entertainment, food and beverages, airport amenities, baggage handling, and, most importantly, friendly customer service. Although service is ultimately an intangible variable, it does impact on demand. Airlines that are perceived to provide a high level of service will have a greater demand for their flights. Airlines such as Virgin Atlantic, Emirates, and Singapore Airlines have been very successful at generating increased demand through a perceived level of strong service.

Finally, there are stochastic and random factors that can materially affect demand. For example, the terrorist attacks of 11 September 2001 crippled demand for air travel for quite some time.

6 While the coefficient for a competitor's price is positive.

CHARACTERISTICS OF DEMAND FOR AIR TRANSPORTATION

The demand for air travel also has many unique characteristics that present problems for the airline industry. While all of the following characteristics shape the demand for air transportation, only the first five will be examined in detail:

- constant fluctuation
- cyclicality
- seasonality and peaking
- directional flow
- perishability
- schedule wait time
- airport access time
- flight time
- hub connection time
- denied boarding time.

The first major characteristic of demand for air transportation is that, unlike the demand for many products, it is constantly fluctuating. Because of the numerous determinants highlighted in the preceding section, demand for individual flights is constantly changing. Moreover, no two routes exhibit the same properties of demand, making every route unique in its demand characteristics.

Cyclicality, a second characteristic of demand for air transport, refers to a long-term trend of peaks and troughs of economic activity. The national economy has long been known to experience such cyclicality. And, since the airline industry is highly correlated with the national and global economy, it is not surprising that it also experiences some cyclicality.

Figure 3.5 depicts the number of passengers carried by US airlines since 1970. While the chart shows overall steady growth, there remain three pronounced drops in US airline industry enplanements: the early 1980s, early 1990s, and post-9/11.

Another major characteristic of air-transport demand is peaking, more commonly called seasonality. Unlike cyclicality, which is a long-term cycle, peaking is more of a short-term event where demand spikes. The most common form of peaking is seasonality where demand increases during the summer months and then declines during the winter months. This trend is particularly true of leisure destinations, where the weather is more favorable and individuals have more time off. For example, Mediterranean resort destinations are in high demand during the summer months, but demand declines during the winter months. This situation also applies to most domestic routes in the United States. As a specific example, consider the Chicago to Seattle route. Figure 3.6 displays a 10 per cent sample of the route's passenger enplanements per quarter since 2002. As Figure 3.6 shows quite clearly, the route experiences tremendous demand during the third quarter of every year, while demand is quite low for the first quarter of every year. This is a fairly regular pattern partly attributed to Seattle's summer cruise ship industry and the Pacific-Northwest's more favorable summer weather. In addition to seasonality, other peaks in demand include Thanksgiving, Christmas holidays, and other statutory holidays.

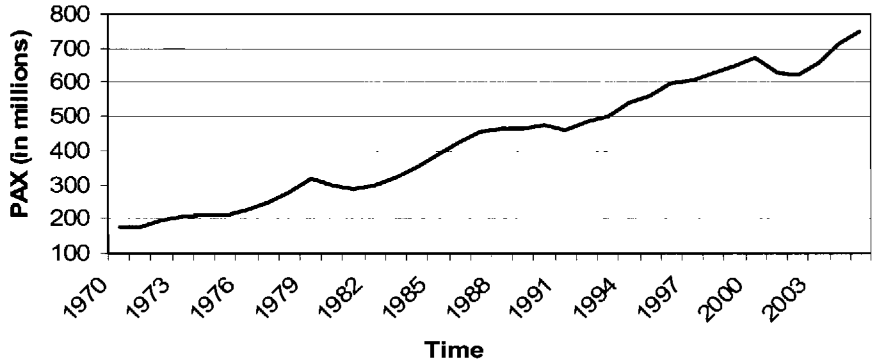

Figure 3.5 US airline industry passenger enplanements since 1970

Source: Compiled by the authors using Back Aviation Form41 data.

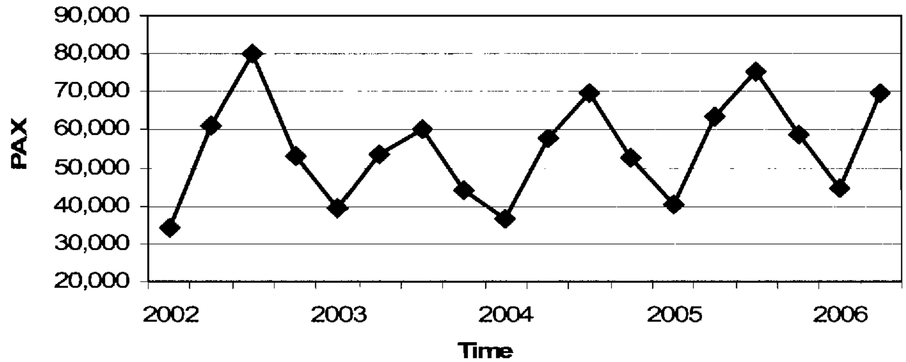

Figure 3.6 Seasonality of Chicago to Seattle enplanements

Source: Compiled by the authors using Back Aviation O&D data.

Since peaking is fairly predictable, airlines can add capacity, if they choose, by either increasing the frequency of the flights on an existing route, or by introducing seasonal-only service. However, in order for airlines to be able to add seasonal capacity, they either have to take the aircraft off other routes or have excess capacity in reserve, which is very expensive. Ideally, airlines would like to add aircraft during the summer months and retire them during the winter; however, short-term leases for aircraft are rare. This means that seasonality can present a sizeable financial, operational, and scheduling burden for airlines, as they want to be able to meet the seasonal demand, but must also bear the assets for the remainder of the year. This is one reason why North American carriers have robust aircraft maintenance schedules during the winter months when their schedule is not as busy. Some carriers, such as American Airlines and Air Canada, have been successful at moving capacity to Central and South America during the northern hemisphere winter season.

Another characteristic of airline demand that is similar to peaking is directional flow which relates to the increased demand of passengers in one direction for a period of time. While cyclicality spans decades and peaking spans years, directional flow is usually assessed on a weekly basis and is fairly short-term. An example of directional flow could include customers flocking to a city a few days before a major sporting event (such as the Super Bowl) and then immediately demanding to leave after the event has finished. The key factor is that directional flow is essentially one-way for a short period of time. This creates another unique scheduling problem for airlines, since, in order to accommodate the directional flow of passengers, some aircraft will be flown in a relatively empty condition in the opposite direction.[7]

The principal problem with cyclicality—peaking—and directional flow is that demand for air transportation is perishable. The moment the plane leaves the gate, any empty seats are lost as revenue-generating products—a situation that does not, of course, pertain to a manufacturing company, which can keep its product in inventory for sale on another day. Consequently, the close matching of demand and supply is essential for success in the aviation industry. Because of situations such as peaking and directional flow, airlines are faced with the prospect that a good portion of their seats will go unsold, simply due to the nature of demand and the structure of their operation. Because of this, pricing is extremely important in order to help offset issues related to the structure of demand. Airline pricing policy and revenue management will be covered in Chapter 11.

SOURCE OF DEMAND

The source of demand is very critical to understanding the nature of demand for air transportation. Depending on the source, demand can be categorized as *direct* or *derived.*

Direct demand refers to demand that directly satisfies a consumer's need. For example, a consumer's need for a dessert can be satisfied by a cheesecake, while another consumer's need for entertainment can be satisfied in a variety of ways, such as by a movie or live sporting event. In both these examples the product or service directly satisfies the need. In the airline industry there is no direct demand since passengers do not purchase tickets just for the sake of flying. Instead, passengers fly in order to arrive at a resort, visit family or conduct business, as well as for a variety of other reasons, and this makes the demand for air transportation a derived demand. Derived demand occurs when the demand for a product or service depends on the demand for other products or services. This means that airlines are generally not able to directly affect the demand for their products. And this further creates a situation in which the airlines largely react to demand. On this basis, airlines must be continually aware of the multitude of factors that can impact on their demand curve. (Strategies for dealing with this problem are covered in more detail in Chapter 11.)

7 While this previous example highlights a one-time directional flow event, rush-hour traffic is an example of a continuous directional flow in demand. In most major cities, the roads are clogged with people attempting to head into the city center in the morning, while the other direction is usually fairly empty. For continuous directional flow in the aviation industry, Las Vegas is probably the best example. On Friday evenings there is demand for air travel to Las Vegas, while on Sunday evenings there is demand for travel out of Las Vegas as people want to spend a weekend in Las Vegas without skipping work. Directional flow of demand presents problems for airlines as they may want to capture the one-way demand, but will have problems filling their aircraft on the return sector.

ELASTICITY OF DEMAND

An important economic principle that can aid airline or airport managers in their economic decision-making process is elasticity. The formal definition of elasticity is the percentage change in the dependent variable (quantity demanded) resulting from a 1 per cent change in an independent variable (factor of demand). Informally, elasticity measures the responsiveness of one variable to changes in another.

The basic formula for elasticity is:

$$Elasticity = \frac{Percentage\Delta Y}{Percentage\Delta X}$$

In measuring elasticity, there are two types of variable: endogenous and exogenous. Endogenous variables are variables that the airline can directly control, whereas exogenous variables are variables that are out of the company's control. In the airline industry both price and service would be endogenous variables, while factors such as consumer income, competitor's price, and complement's price are exogenous variables. It is useful to know the effects of these exogenous variables on demand, since this information allows the airline to manage capacity and demand more efficiently. The three major elasticities that will be explored in greater detail are:

- price elasticity
- cross-price elasticity
- income elasticity.

Price Elasticity

While there are numerous types of elasticity, the price elasticity of demand is probably one of the most useful for airline managers. Using the general definition of elasticity, price elasticity is the percentage change in the quantity demanded resulting from a 1 per cent change in price. Therefore, price elasticity enables managers to perform "what-if" scenarios to see the effects on the quantity demanded that a change in price would have.

Since elasticity is not constant throughout the entire demand curve there are two ways to measure elasticity. *Point* elasticity measures the elasticity of the function at a specific value, while *arc* elasticity measures the elasticity of the function over a range of values. Thus, arc elasticity is an average of the elasticities over a specified range of values, while point elasticity is the exact level of responsiveness at the specific price. The basic formulae for point price elasticity and arc price elasticity are:

Point price elasticity:

$$\varepsilon_p = \frac{Percentage\Delta Q}{Percentage\Delta P}$$

$$\varepsilon_p = \frac{\delta Q}{\delta P} \times \frac{P}{Q}$$

Arc price elasticity:

$$E_p = \frac{\dfrac{\Delta Q}{AverageQ}}{\dfrac{\Delta P}{AverageP}} = \frac{\Delta Q}{\Delta P} \times \frac{P_2 + P_1}{Q_2 + Q_1}$$

In order to calculate price elasticity, the derivative of the demand function needs to be calculated.[8] As a numerical example, consider the following explicit short-run demand function for a flight, where we assume that the competitor's ticket price is $120 and that the average annual income in the market is $40,000.

$$Q_D = 800 - 2P_X + 1.5P_Z + 0.0005Y$$

The first step is to re-compute the demand function based on the assumptions concerning the external market. Substituting the values into the demand function yields:

$$Q_D = 800 - 2P_X + 1.5(120) + 0.0005(40{,}000)$$
$$Q_D = 800 - 2P_X + 180 + 20$$
$$Q_D = 1{,}000 - 2P_X$$

On the basis of this new demand function that is only related to price, price elasticity can now be calculated. The first step is to take the first-order derivative of the demand function. This results in a value of -2 (see footnote 8). The next step is to find the quantity demanded based on a single price for point elasticity, or for multiple prices for arc elasticity. With a ticket price of $100, the quantity demanded would be 800 passengers. Using the point price elasticity formula, the point price elasticity is:

$$\varepsilon_P = \frac{\delta Q}{\delta P} \times \frac{P}{Q} = -2 \times \frac{100}{800} = -0.25$$

A point price elasticity value of −0.25 means that, for every 1 per cent increase in the ticket price, from the $100 level, the quantity demanded would decrease by 0.25 per cent.

The arc price elasticity for ticket prices ranging from $100 to $200 can be calculated as follows:

$$E_p = \frac{\Delta Q}{\Delta P} \times \frac{P_2 + P_1}{Q_2 + Q_1}$$

$$E_p = \frac{-200}{100} \times \frac{200 + 100}{600 + 800}$$

$$E_p = -2 \times \frac{300}{1400}$$

$$E_p = -0.43$$

8 For those readers without calculus, the derivative is simply the change in the dependent variable for a one-unit change in the independent variable. In this case that would be $\Delta Q/\Delta P$ or −2. That is, a one-unit change in P will produce a minus two-unit change in Q.

Price elasticity is usually categorized into one of three groups based on its numerical value and its impact on demand:

$$|E| > 1 \text{ Elastic}$$
$$|E| < 1 \text{ Inelastic}$$
$$|E| = 1 \text{ Unitary Elastic}$$

Price elasticity with an absolute value of less than one is termed inelastic. Inelastic demand occurs when a 1 per cent increase in price results in a less than 1 per cent decrease in demand. (In the above example, since the absolute value of elasticity was calculated as less than one, the price elasticity would be inelastic). In these situations, consumers have a strong desire to purchase the good or service and therefore price is not a central concern. In price-inelastic situations, firms can increase the price to increase total revenue as the price effect dominates the quantity effect. However, since point elasticity is not constant throughout a linear demand curve (in fact, as will be discussed later, every linear demand curve has both an elastic and inelastic region), this practice can only continue up to a certain price at which the demand becomes less inelastic (or more elastic). A good example of a price-inelastic product is demand for air travel in the short run. In the few days leading up to the day of departure, the majority of travelers who need to take the flight are willing to pay for an expensive ticket, since they are likely to have an important reason for traveling, and they rank the convenience of the flight more highly than its price.[9]

At the opposite end of the spectrum is elastic demand. Elastic demand occurs when the coefficient of elasticity has absolute values greater than one. With elastic demand, consumers are more sensitive to changes in price, so that a 1 per cent decrease in price will be offset by a greater than 1 per cent increase in the quantity demanded, and total revenue increases with price cuts. However, as above, the benefits of price cuts will be exhausted at some point as the firm will eventually reach a portion of the demand curve where demand becomes inelastic. Longer-term demand for air travel is more price elastic than short-term demand because many passengers (especially leisure travelers) will choose to take a flight based solely on price.

The final category of elasticity is unitary elastic demand, which has an absolute value equal to one. Under unitary elastic demand, the quantity and price effects are equal, creating a situation where a 1 per cent increase in price is directly offset by a 1 per cent decrease in the quantity demanded.

As the preceding discussion indicates, the point at which the company achieves the optimal price level is the point of unitary elasticity. Therefore, the managerial rules of thumb are simple: if the demand for the product is inelastic, then the company should raise prices; if the demand for the product is elastic, then the company should lower prices; if the demand for the product is unitary, then the company should retain the present price. These pricing decisions are displayed graphically in Figure 3.7. For simplicity, the graph focuses only on revenue maximization; the true goal, profit maximization, will be discussed more fully in subsequent chapters.

9 That is part of reason why last-minute tickets are expensive and the airlines use revenue management techniques to save seats for last-minute time-sensitive passengers.

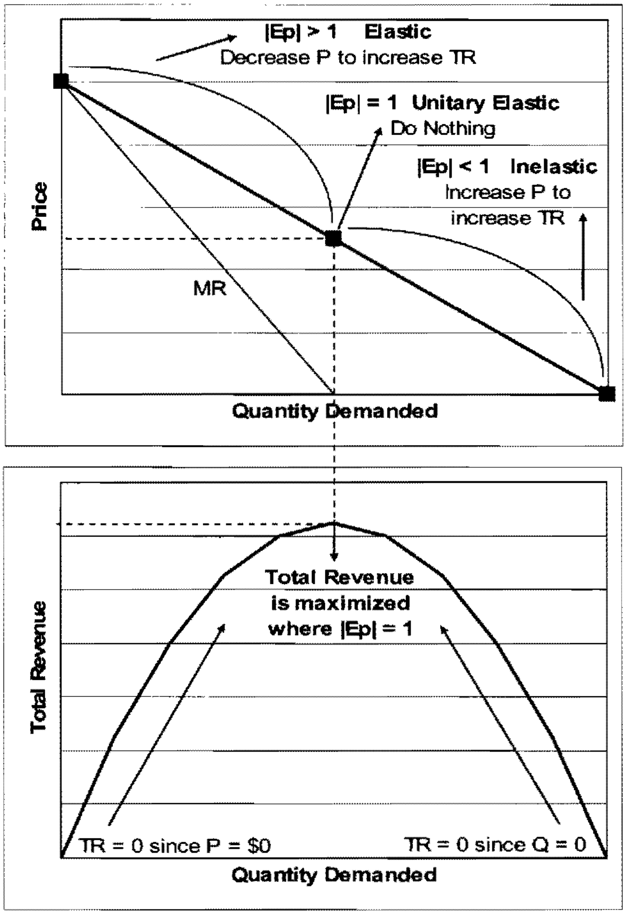

Figure 3.7 Pricing decision based on elasticity

Since a clear relationship exists between price and total revenue for different states of elasticity, Table 3.3 summarizes the impact that a change in price has on total revenue. As mentioned previously, an increase in price will increase total revenue for goods with inelastic demand, while decreasing total revenue for products which are price-elastic. The direct opposite is true for a decrease in price. For unitary elastic products, a change in price will not affect total revenue since total revenue is maximized at the point where unitary elasticity is achieved.

While elasticity varies along a linear demand curve, certain products and services can be categorized on the basis of their normal own-price elasticity. As mentioned earlier, air travel exhibits tendencies of both inelastic and elastic demand, depending on the timeframe. Table 3.4 provides a list of various services and products with their corresponding estimated price elasticity.

Cross-price Elasticity

Another type of elasticity is cross-price elasticity of demand, which measures the impact of a related firm's price on demand. Cross-price elasticity of demand helps determine

Table 3.3 Relation between price changes and total revenue for different states of elasticity

	Elastic	Unitary Elastic	Inelastic
Price Increase	↓ in TR	No Impact	↑ in TR
Price Decrease	↑ in TR	No Impact	↓ in TR

Table 3.4 Estimated price elasticities of demand for various goods and services

Goods	*Elasticity of Demand*
Inelastic	
Salt	0.10
Toothpicks	0.10
Airline travel, short-run	0.10
Gasoline, short-run	0.20
Gasoline, long-run	0.70
Coffee	0.25
Tobacco products, short-run	0.45
Physician services	0.60
Automobiles, long-run	0.20
A pproximately Unitary Elastic	
Movies	0.90
Housing, long-run	1.20
Private eductaion	1.10
Tires, short-run	0.90
Tires, long-run	1.20
Elastic	
Restaurant meals	2.30
Foreign travel, long-run	4.00
Airline travel, long-run	2.40
Automobiles, short-run	1.20–1.50
Fresh tomatoes	4.60

Source: Compiled by the authors from Anderson, McLellan, Overton, and Wolfram (1997).

whether the related firm is either a substitute (competitor) or a complement. The basic formulae for both point and arc cross-price elasticity are as follows.

Point cross-price elasticity:

$$\varepsilon_{xy} = \frac{Percentage\Delta Q}{Percentage\Delta P_y}$$

$$\varepsilon_{xy} = \frac{dQ}{dP_y} \times \frac{P_y}{Q}$$

Arc cross-price elasticity:

$$E_{xy} = \frac{\dfrac{\Delta Q}{AverageQ}}{\dfrac{\Delta P_y}{AverageP_y}}$$

$$E_{xy} = \frac{\Delta Q}{\Delta P_y} \times \frac{P_{y2} + P_{y1}}{Q_2 + Q_1}$$

Using the above formulae and the preceding example, the point cross-price elasticity of demand can be found, assuming that the ticket price of the firm's flight is $100, the competitor's ticket price is $120, and the annual income is $40,000:

$$Q^D = 800 - 2P_X + 1.5P_Y + 0.0005M$$
$$Q^D = 800 - 2(100) + 1.5P_y + 0.0005(40{,}000)$$
$$Q^D = 580 + 1.5P_y$$

$$\varepsilon_{xy} = \frac{\delta Q}{\delta P_y} \times \frac{P_y}{Q}$$

$$\varepsilon_{xy} = 1.5 \times \frac{120}{760}$$

$$\varepsilon_{xy} = 0.24$$

From this example, the cross-price elasticity of 0.24 indicates that the related competitor can be considered a mild substitute to the firm, since the cross-price elasticity of demand is greater than zero. In this situation, a 1 per cent increase in the competitor's ticket price will cause a 0.24 per cent increase in the firm's quantity demanded. Since the two firms are competing against each other, an increase in the rival firm's price would make them less competitive, increasing demand for the company's product. In the airline industry, where price competition is fierce, the cross-price elasticity of demand is undoubtedly highly positive. With information from the Internet, consumers are now able to view almost all the pricing options for their flights, and therefore some passengers can be extremely price-conscious (especially since there is little difference between the product, namely an airline seat).

If the cross-price elasticity of demand is found to be less than zero, then the related firm's product is determined to be a complementary good. A complementary good is one which increases the demand for the firm's good. Examples of complementary goods to the airline industry are both hotels and rental cars, as the price of accommodation and transportation directly relate to the demand for air transportation. As an example, if the cross-price elasticity of demand was found to be -0.55, then a 1 per cent increase in the complementary good's price would create a 0.55 per cent decrease in the quantity demanded. As in the case of own-price elasticity, cross-price elasticity can be categorized into one of three groups based on its numerical value and the impact of a related firm's price on demand:

$$E_{X,Y} > 0 \text{ Substitute}$$

$$E_{X,Y} < 0 \text{ Complement}$$

$$E_{X,Y} = 0 \text{ Independent}$$

Income Elasticity

A third type of elasticity is income elasticity. Income elasticity determines the sensitivity that changes in the annual income of consumers have on the quantity demanded for a product. Disposable income, personal income, and gross domestic product (GDP) are all good measures for this variable. Disposable income is income available to spend on leisure travel, while business income is a part of the equation since increased business activity will probably spur an increased need for business travel. Since the income variable comprises two parts for air transportation, GDP is the best proxy variable for income, since it takes into consideration both household disposable income and business activity. The formulae for income elasticity are similar to both the price and cross-price elasticity formula.

Point income elasticity:

$$\varepsilon_Y = \frac{Percentage\Delta Q}{Percentage\Delta Y}$$

$$\varepsilon_Y = \frac{dQ}{dP_y} \times \frac{Y}{Q}$$

Arc income elasticity:

$$E_Y = \frac{\dfrac{\Delta Q}{AverageQ}}{\dfrac{\Delta Y}{AverageY}}$$

$$E_Y = \frac{\Delta Q}{\Delta Y} \times \frac{Y_2 + Y_1}{Q_2 + Q_1}$$

Using the same example, assuming that the ticket price is \$100 and the competitor's ticket price is \$120, the arc income elasticity between \$40,000 and \$50,000 is:

$$Q^D = 800 - 2P_X + 1.5P_Y + 0.0005Y$$
$$Q^D = 800 - 2(100) + 1.5(120) + 0.0005Y$$
$$Q^D = 780 + 0.0005Y$$
$$Q^D_1 = 780 + 0.0005(40{,}000) = 800$$
$$Q^D_2 = 780 + 0.0005(50{,}000) = 805$$

$$E_Y = \frac{\Delta Q}{\Delta Y} \times \frac{Y_2 + Y_1}{Q_2 + Q_1}$$

$$E_Y = \frac{5}{10{,}000} \times \frac{50{,}000 + 40{,}000}{805 + 800}$$

$$E_Y = 0.0005 \times \frac{90{,}000}{1{,}605}$$

$$E_Y = 0.028$$

As with own-price elasticity and cross-price elasticity, the good or service can also be classified according to its income elasticity:

$$E_Y > 1 \text{ Superior good}$$
$$E_Y > 0 \text{ Normal good}$$
$$E_Y < 0 \text{ Inferior good}$$

If the product's income elasticity is greater than zero, then the good is categorized as a normal good—that is, the quantity demanded of a normal good increases with any increase in the consumer's income. As we might expect, the vast majority of goods and services can be classified as normal goods, as in the above example. A subcategory of normal goods are superior goods, whose income elasticity is greater than one. Superior goods have a proportional increase in the quantity demanded that is greater than the increase in consumer income. Superior goods usually encompass high-end luxury products such as fancy sport cars and business jet travel.

The other goods categorized according to income elasticity are inferior goods. Inferior goods have income elasticity values less than zero, indicating that, for any increase in income, the quantity demanded decreases. This peculiar situation occurs when products have a price advantage over competitors but are generally not perceived as quality goods. Therefore, when consumers' income increases, they are more willing to purchase the perceived better product. Examples of inferior goods might include generic products versus brand names or, in some markets, coach travel versus first-class seats.

The concept of elasticity is critical to understanding pricing policies of any industry, especially the air transportation industry. And, as the earlier discussion has shown, elasticity can be used to determine the optimum price level where total revenue is maximized. Ultimately, revenue management has its foundation in this concept since elasticity can be used to help manage both pricing and capacity.[10]

SUPPLY OF AIRLINE SERVICES

Chapter 2 introduced us to the principles of supply and demand, and we will now apply the supply part of those principles to the air transportation industry. One of the key reasons why the airline industry has faced financial difficulties and has profit margins well below many other industries is that its demand fluctuates constantly but its supply is relatively fixed. Lack of flexibility in the supply function makes it very difficult to manage capacity effectively. Therefore, it is important to understand the factors impacting airline industry supply in some detail. The next section introduces those factors, while Chapter 4 will discuss airline production in greater detail.

10 The topic of yield management will be discussed in full in Chapter 11.The principles of elasticity enable managers to see the impact that changes in competitors' prices, advertising campaigns, and economic booms and recessions will have on the airline's operations. From a foundation based on elasticity, a competitive plan and strategy can be created for the airline. Therefore, the concept of elasticity is invaluable to decision-makers in all industries, especially highly volatile industries such as air transportation.

FACTORS AFFECTING THE SUPPLY OF AIRLINE SERVICES

As mentioned in the introduction to this chapter, airline supply refers to airlines' ability and willingness to provide a specific number of seats at alternative prices in a given time period and in a given market. In the air transportation industry, supply is usually expressed in available seat miles (ASMs) or available ton miles (ATMs). An ASM is simply one seat carried through the air for one mile, regardless of whether it contains a passenger or not. The presence of a revenue passenger in the seat is the key difference between RPMs (demand) and ASMs (supply), as RPMs only measure seats that have a revenue passenger in them. A similar distinction can be made between ATMs (supply) and RTMs (demand).

An implicit supply function for the airline industry can be written as:

$$Q_S = f\{P, P_{RES}, Tech, Comp, Rand, Govt\}$$

where:
- P = ticket price
- P_{RES} = price of resources, including aircraft costs, fuel, labor costs, maintenance costs, and so on
- Tech = technological improvements
- Comp = behavior of the competition
- Rand = random factors
- Govt = government regulation.

The major determinant of supply, just like demand, is the ticket price of the good or service. This relates to the law of supply in that the quantity of a good supplied in a given time period increases as its price increases, assuming all else is held constant. In the airline industry this simply means that airlines are willing to supply more seats as ticket prices increase. Based on the law of supply, the supply curve slopes upward, so that any change in price is simply a movement along the supply curve.

Table 3.5 provides a hypothetical supply schedule for the New York to Los Angeles flight.

The next major determinant of supply is the price of resources. For the air transportation industry, production resources include, but are not limited to, aircraft, fuel, maintenance, labor, and landing fees. All these factors, and many more, impact on supply because they affect the cost of production. If the cost of production increases, the airline's total costs increase, causing a leftward shift in the supply curve. In the airline industry, an increase

Table 3.5 Supply schedule for New York to Los Angeles

Ticket Price of New York to Los Angeles Round Trip	Quantity Supplied (No. of passengers)
200	15
500	315
1000	465
2000	615
3000	690
5000	735

in the costs of production may force the airline to cut some flights that would no longer be profitable. This reduction in flights is a leftward shift of the supply curve—that is, fewer seats are offered at the same ticket price. Conversely, if the price of resources decreases, the supply curve shifts to the right—that is, more seats are offered at the same ticket price.

Applications of the impact that the price of resources have on supply are numerous. For example, during bankruptcy protection, Delta Air Lines was able to significantly reduce its costs by receiving wage concessions and reducing aircraft leasing rates. As a result of this, Delta dropped many domestic destinations and redeployed resources to new international markets. Similarly, with skyrocketing fuel costs, airlines eliminated a number of flights that were previously viable.

The next major factor that determines supply for air transportation is technology. The impact of technology on the supply of air transportation has been vast. Technology has advanced civil aviation to the point where it is one of the safest modes of transportation. The introduction of the Boeing 747 in 1970 created a rightward shift in the supply curve as the jumbo jet was able to carry more passengers on a flight than any other aircraft (Boeing, 2007). The new Airbus 380 will also cause a shift in the supply curve of airlines that receive the aircraft. However, perhaps one of the best examples of technology's impact on air transportation supply is the introduction of ultra-long-range aircraft, such as the Boeing 777-200LR and the Airbus A340-500, which allowed routes such as Singapore to Los Angeles and New York to Dubai to be be operated on a nonstop basis. Finally, the introduction of ETOPS (extended-range twin-engine operations) enabled a greater supply of air transportation over the Atlantic and Pacific oceans through the use of smaller twin-engine aircraft.

Competitive factors are another important determinant of supply for air transportation. Since airlines have historically aggressively competed over market share, competition from other airlines has had a considerable impact on supply. Airlines regularly adjust capacity and supply in markets in response to competition and changing market forces, although there have been cases where they have taken this to an extreme. For example, in 1999 American Airlines faced antitrust lawsuits over its competitive actions against smaller rivals, particularly Vanguard Airlines out of Love Field (American Cleared, 2003). The lawsuit alleged that American Airlines had "dumped" capacity on routes where

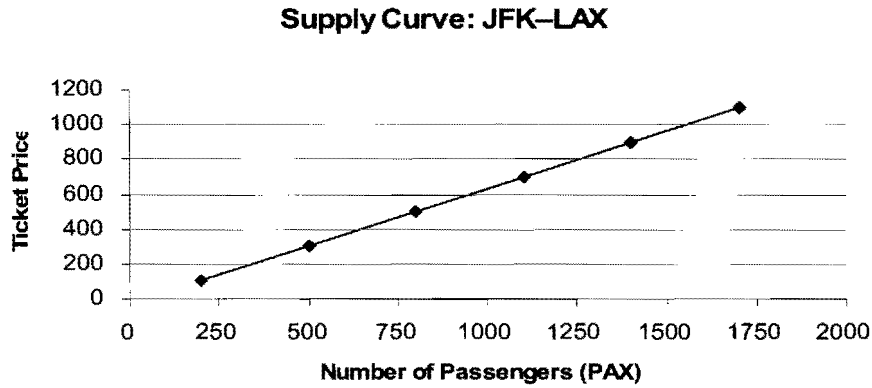

Figure 3.8 Supply curve

Vanguard Airlines was competing with American (American Cleared, 2003). Eventually American Airlines was cleared of the charge. Similarly in 2000, El Salvador-based Grupo TACA accused Continental Airlines of capacity dumping following a liberalization of the US–El Salvador bilateral air agreement (Knibb, 2000). Since the financial downturn in the US aviation industry, airlines have begun to compete less on market share and focus more on profits.

Just as in demand, random factors play a large role in affecting supply. The tragic events of 11 September 2001 not only created a dramatic one-off shift in the demand curve, but also impacted on the supply curve, albeit not to the same extent. Because of the terrorist attacks, all commercial flights were grounded for two days, then, when flights resumed, security procedures were added that added some costs. In addition, the two carriers involved in the accidents—American Airlines and United Airlines—lost aircraft, thereby reducing their fleet size and ability to transport passengers.

The deregulation of air transport has also significantly affected the supply curve. Since regulation generally prohibits market forces from determining supply, an artificial cap is usually set on supply. Therefore, when air transport deregulation occurred, the artificial cap on supply was withdrawn, and supply subsequently increased.

A more current example would be the United States–China air transportation agreement. The current agreement limits the amount of supply between the two countries, thereby placing an artificial cap on supply. However, whenever additional rights are granted, the supply of air transportation between the two countries will shift to the right. Further discussion of international aviation is contained in Chapter 6.

The final factor impacting on air transportation supply is government regulation. Despite the fact that aviation industry has gone through massive privatization and financial deregulation, almost all aspects of safety and operations remain highly regulated. On 20 January 2005 the FAA mandated that the vertical separation between aircraft above the United States at altitudes from 29,000 to 41,000 feet be reduced from 2,000 feet to 1,000.[11] This action has made more routes available to airlines and has had the effect of shifting the supply function to the right (Figure 3.9).

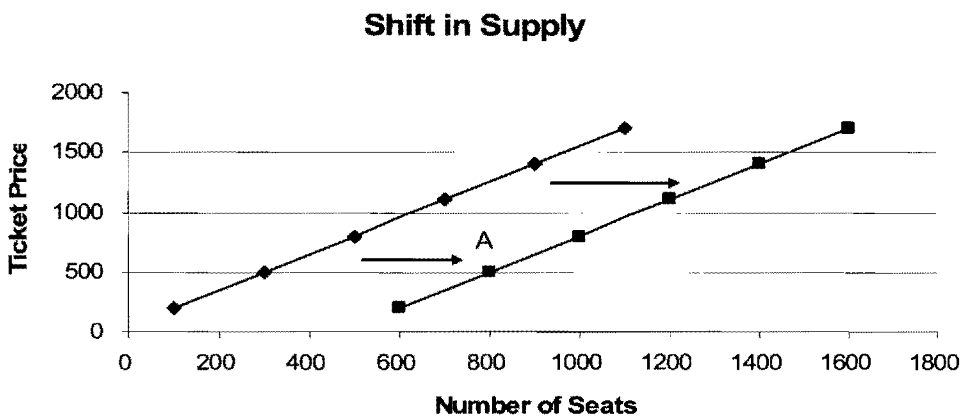

Figure 3.9 Impact of deregulation on supply

11 Domestic Reduced Vertical Separation Minimum (DRVSM); the rule was designed to increase airspace capacity.

CHARACTERISTICS OF SUPPLY FOR AIRLINE SERVICES

Two important characteristics of supply that help shape the air transportation industry are seasonality and rigidity. Both these characteristics make it difficult for airlines to match supply and demand.

One major feature of air transportation demand is its constant fluctuation, and airlines must react to this by adjusting the supply to match the passenger demand. In order to accommodate seasonality, airlines need to either pull capacity off existing routes, or have idle capacity available to accommodate additional flights. Both these options entail embedded costs, and , airlines typically use a mix of both. Although the increased costs of a seasonal schedule can be offset by increased revenues, airlines have greater difficulty adjusting supply on a short-term basis due to the second major characteristic of airline supply—rigidity.

An airline's supply is fairly rigid as it can be difficult for it to reduce and/or increase supply dramatically. Airlines create their schedules at least six months in advance, and accept bookings up to a year in advance, so they must adhere to their schedules or face re-accommodation fees. Fixed costs, such as investment in infrastructure at hub airports, aircraft leases, and labor contracts have to be paid regardless of the schedule, making it impractical for airlines to reduce capacity at short notice. This is a particular problem for those major US carriers that operate in a hub-and-spoke network and is one reason why a non-hub carrier, such as Southwest Airlines, has greater flexibility with supply. Ultimately, this rigidity in supply limits the airlines' ability to match supply and demand effectively

AIRLINE SUPPLY AND DEMAND EQUILIBRIUM

While Chapter 2 introduced the conceptual framework of market equilibrium, here we present a concrete example of both supply and demand functions to determine market equilibrium price and quantity. Suppose that the demand and supply functions for DirectJet's flights from New York to Seattle are the following:

$$Q^D = 500 - 5P_X + 2P_Z + 0.01Y$$
$$Q^S = 800 + 3P_X - 2P_{RES}$$

Assuming that the competitor's ticket price is $300, the annual average income is $50,000 and the cost of resources for the flight is $100, both the demand and supply functions can be rewritten solely in terms of the ticket price:

$$Q^D = 500 - 5P_X + 2(300) + 0.01(50{,}000)$$
$$Q^D = 500 - 5P_X + 600 + 500$$
$$Q^D = 1{,}600 - 5P_X$$

$$Q^S = 800 + 3P_X - 2(100)$$
$$Q^S = 800 + 3P_X - 200$$
$$Q^S = 600 + 3P_X$$

In order to find the market equilibrium price for the flight, the demand and supply functions need to be set equal to each other:

$$Q^D = Q^S$$
$$1{,}600 - 5P_X = 600 + 3P_X$$
$$1{,}000 = 8P_X$$
$$P_X = \$125$$

Based on this calculation, the market equilibrium price for DirectJet's flight between New York and Seattle is $125. At this price point, the quantity demanded equals the quantity supplied. Since the supply and demand curves are equal to each other at the equilibrium price, either the demand or supply function can be used to determine the market equilibrium quantity. Both functions are used below to illustrate this point.

$$Q^D = 1{,}600 - 5P_X$$
$$Q^D = 1{,}600 - 5(125)$$
$$Q^D = 1{,}600 - 625$$
$$Q^D = 975$$

$$Q^S = 600 + 3P_X$$
$$Q^S = 600 + 3(125)$$
$$Q^S = 600 + 375$$
$$Q^S = 975$$

As depicted in Figure 3.10 below, at a uniform price of $125, 975 consumers will demand DirectJet's flight between New York and Seattle,[12]

The appendix to this chapter contains a mathematical derivation of an advanced pricing and elasticity application.

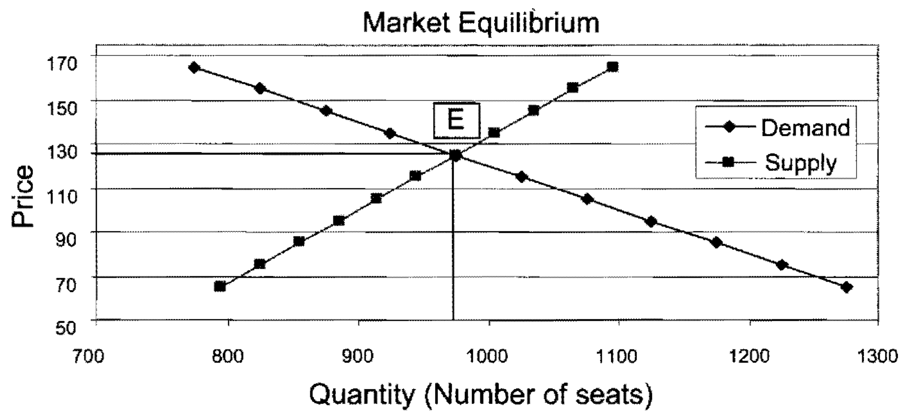

Figure 3.10 Supply and demand for DirectJet's flights from New York to Seattle

12 However, in the airline industry, the supply is generally not a smooth upward-sloping curve. An aircraft can accommodate only a fixed number of passengers. This would create the situation whereby the supply curve for a single flight would move up step-wise according to the number of aircraft in the airline's fleet.

SUMMARY

Building on the previous chapter, this chapter presented a more in-depth analysis of the airline demand and supply. A hypothetical demand schedule was derived and analyzed with concrete examples of the phenomenon of price discrimination. The specific topic of elasticities and what they mean and how to calculate them were discussed in great detail. Finally, the topic of airline supply was discussed, with appropriate definitions.

APPENDIX: ADVANCED PRICING AND ELASTICITY APPLICATION

Since the ultimate goal of any pricing policy is to maximize total revenue, and total revenue is known to be maximized where price elasticity is unitary elastic, formulas can be derived to determine the optimum price. In order to understand the formula derivation, consider the inverse demand function with the general form of:

$$P = a - bQ$$

The basic, general formula for point price elasticity can be slightly modified to reflect an inverse demand function. The derivative of the demand function is simply the inverse of the derivative for the inverse demand function. Since the derivative, or slope, of the inverse demand function is (*b*), the price point elasticity formula can be rewritten as:

$$\varepsilon_p = \frac{dQ}{dP} \times \frac{P}{Q}$$

$$\varepsilon_p = \frac{1}{slope} \times \frac{P}{Q}$$

$$\varepsilon_p = \frac{1}{-b} \times \frac{P}{Q}$$

In addition, the price variable (P) can be replaced by the general form of the inverse demand function. Therefore, the elasticity formula would be:

$$\varepsilon_p = \frac{1}{-b} \times \frac{a - bQ}{Q}$$

Since the goal is to determine the optimum pricing point where total revenue is maximized, the elasticity formula needs to be set equal to -1, or where unitary elasticity is achieved. Once the equation is set equal to negative one, the optimum quantity where total revenue is maximized can be determined. The optimum quantity is:

$$\frac{1}{-b} \times \frac{a - bQ}{Q} = -1$$

$$\frac{a - bQ}{Q} = b$$

$$Q = \frac{a}{2b}$$

While this formula provides the optimum quantity demanded where total revenue is maximized, it needs to be placed back into the inverse demand function to obtain the optimum price level. Total revenue is maximized where:

$$P = a - b(\frac{a}{2b})$$

$$P = \frac{a}{2}$$

Based on the derivations of the inverse demand function and the price point of elasticity formula, the total revenue maximizing price point can be found. Using the previous example, the revenue maximizing price and quantity would be:

$$Q^D = 1000 - 2P_x$$
$$P = 500 - 0.5Q$$

$$Q^* = \frac{500}{2(0.5)} = \frac{500}{1} = 500$$

$$P^* = \frac{500}{2} = \$250$$

Therefore, the revenue maximizing price is \$250, which creates a demand of 500 seats. At this point, the point price elasticity of demand is negative one, or unitary elastic. These formulae can aid decision-makers in setting the correct price for a product and service, and is the building block for yield management, which is covered in more detail in Chapter 11.

REFERENCES

American Cleared Again on Predatory Pricing Issue (2003). *Airline Business.* Retrieved on 1 February 2007 from Air Transport Intelligence at: http://www.rati.com/iata/oapd.

Anderson, L., McLellan, R., Overton, P., and Wolfram, L. (1997). Price Elasticity of Demand. *Mackinac Center for Public Policy.* Retrieved on 29 January 2007 from: http://www.mackinac.org/article.aspx?ID=1247.

Boeing (2007). The Boeing 747 Family. Retrieved on 1 February 2007 from: http://www.boeing.com/commercial/747family/background.html.

Knibb, D. (2000). Play by the Rules. *Airline Business.* Retrieved on 1 February 2007 from Air Transport Intelligence at: http://www.rati.com/iata/oapd.

4

Cost and Production Analysis: The General Concepts

Today, the situation is exacerbated with costs exceeding revenues at four times the pre-September 11 rate. Today, we are literally hemorrhaging money. Clearly this bleeding has to be stopped—and soon—or United will perish sometime next year.

James Goodwin, chairman and CEO of United Airlines

The theory of supply and production is one of the principal theories of economics and is the foundation upon which many other economic theories are based. This chapter adopts a modern approach to the principles and practice of airline supply, production and cost analysis. It is rigorous and clear in its presentation of the techniques involved in cost accounting, and goes one step further by demonstrating how this technical knowledge can be applied across a whole range of management decisions. Although this chapter has been prepared primarily for students, it may also be useful to the industry and government agencies.

In this chapter we discuss a number of topics related to production and cost theory. We are specifically interested in applying this theory to the airline and aviation industries. The chapter provides a body of knowledge that will allow readers to understand how the costs of production are minimized through either the mix of inputs or the reallocation of resources across multiple plants. It also illustrates which costs are relevant for specific airline decisions. Finally, the chapter discusses returns to scale, scope, and density. The topics covered are as follows:

- Cost classifications, including:
  - Cost classification: historical, current, future, replacement and sunk costs
  - Cost division: total, fixed, variable, average, mixed, and marginal costs
  - Explicit/implicit/opportunity costs
  - Accounting/economic costs
- Cost functions, including:
  - Using cost functions for managerial decisions
- Economics of scale, scope, and density
- Airline industry cost structure, including:
  - Fuel costs
  - Flight/cabin crew expenses

  - Maintenance costs
  - Other operating costs
- Airline economies of scale, scope, and density
- Airline break-even analysis
- Operating leverage
- Airline operating leverage.

Profitability can be improved by increasing revenues or decreasing costs. More precisely, the amount of profitability is the difference, or contribution margin, between unit revenues and unit costs. As a concrete example of these concepts, Table 4.1 displays the revenue per air seat mile (RASM), the cost per air seat mile (CASM), and the operating margin for the major US airlines in 2005. Southwest Airlines (WN) had the widest spread of US airlines, and this was largely as a result of a low operating CASM. On the other hand, the airline with the second highest contribution margin, Continental Airlines (CO), achieved this largely as a result of a high RASM.

Continuing with this comparison, Table 4.2 presents the RASM and CASM from 2006. From this table we see that both Continental (CO) and Northwest Airlines (NW) experienced a significant increase in their RASM minus CASM, with increases of 0.99 cents and 1.31 cents respectively. Southwest (WN) remained among the top airlines, with a marginal increase between the two years. However, it is important to note that consistency in profitability is the long-run goal of any airline, and Southwest has achieved this goal. JetBlue (B6) was the airline with the lowest cost in both years but also the lowest revenue—probably as a result of a limited route structure.

More generally, airlines have been successful at increasing revenues by opening up new routes, dropping unprofitable destinations, selecting the most efficient aircraft type, and implementing advanced revenue management practices. Up until 2001 airlines largely ignored a detailed focus on cost reductions, but the recent downturn in aviation demand has shown that reductions in costs are quite important. Through meal reductions, labor concessions, automation, and aircraft retirement, legacy carriers have made drastic changes to lower their cost structure. Low-cost carriers have also been successful due to their ongoing focus on managing costs. The lesson from the data is clear: it is vitally important to understand the cost structure of any airline and the variables that impact on profitability. The remainder of this chapter is devoted to these issues.

Table 4.1 Average cost and revenue for US airlines in 2005

US Airline RASM and CASM Comparison (2005)												
	AA	AS	B6	CO	DL	FL	F9	HP	NW	UA	US	WN
RASM	9.37	9.66	6.84	9.72	8.58	9.10	9.16	8.20	10.40	9.15	9.85	8.34
Operating CASM	5.96	6.12	3.98	5.96	5.83	6.30	6.02	5.82	6.85	5.50	6.16	4.30
RASM - CASM	3.41	3.55	2.85	3.76	2.74	2.80	3.14	2.38	3.55	3.65	3.69	4.04

Source: Compiled by the authors using Back Aviation Form41 data.

Note: Airline codes are defined in the appendix to Chapter 12.

Table 4.2 Average cost and revenue for US airlines in 2006

US Airline RASM and CASM Comparison (2006)												
	AA	AS	B6	CO	DL	FL	F9	HP	NW	UA	US	WN
RASM	10.33	10.52	7.66	10.60	9.78	9.82	9.49	9.34	11.74	10.18	10.96	9.33
Operating CASM	6.46	6.36	4.78	5.85	6.03	6.80	6.37	6.38	6.88	6.07	7.00	5.18
RASM - CASM	3.86	4.16	2.88	4.75	3.75	3.02	3.13	2.96	4.86	4.12	3.96	4.15

Source: Compiled by the authors using Back Aviation Form41 data.
Note: Airline codes are defined in the appendix to Chapter 12.

COST CLASSIFICATIONS

If the Wright brothers were alive today Wilbur would have to fire Orville to reduce costs.

Herb Kelleher, Southwest Airlines[1]

The economic definition of a cost is the foregone alternative use incurred in the exchange, transformation, use, and production of goods or resources. For example, for a manufacturer, the costs could be the alternative uses of raw materials, labor, buildings, and general overhead supplies. In the airline industry, costs incurred include labor (of various types), fuel, maintenance, aircraft, catering, and airport landing/usage fees—to name just a few. Typically, fuel, labor, maintenance, and aircraft ownership costs are the four largest costs for any airline. All these costs represent a foregone alternative to the organization. Costs are generally separated into different categories, and these classifications are discussed below.

Cost Classification: Historical, Current, Future, Replacement, and Sunk Costs

Costs can be classified depending on time. Historical costs are the costs actually incurred to acquire an asset, whereas current costs are the costs incurred under prevailing market conditions. For example, assume that DirectJet bought an aircraft in 2007 for $25,000,000, but today the aircraft is worth $20,000,000 in the open market. In this example, the historical cost for the aircraft would be $25,000,000, while the current cost would be $20,000,000. Although historical costs are used in accounting to value assets, current costs provide the best representation of the present situation. And, quite obviously, future costs are the expected costs that may be incurred sometime in the future; these will be affected by such macroeconomic variables as the (uncertain) rate of inflation and the rate of interest. Replacement costs are the costs required to duplicate the productive capabilities using current technology. For example, in the airline industry, an older-generation 737 could be replaced by a modern 737NG for the cost of the 737NG.

1 *USA Today*, 8 June 1994.

The final, and perhaps most important, costs are sunk costs—that is, costs that have been incurred in the past and are not recoverable. Since these costs have already occurred and they are not recoverable, they should never influence business decisions. Unfortunately, they commonly do. As an example of such misguided decision-making, suppose an airline has already purchased eight additional lavatory units and is considering installing these units in new aircraft. However, additional detailed analysis shows that these lavatory units would increase fuel costs, decrease passenger revenue (by removing seats), and increase maintenance costs. Suppose, further, that there is no secondary market for these lavatories. Clearly, installing the aircraft lavatories would be costly to the airline, but managers may want to install the lavatories anyway, since "we already have them." In this example, the pre-purchase of the lavatories has now become a sunk cost, and therefore should not affect the final decision whether or not to install them.

Cost Division: Total, Fixed, Variable, Average, Mixed, and Marginal Costs

While costs categorized by time are important in accounting, in economics the most common and practical method of classifying costs is by their relation to output. Costs that remain fixed in the short run, regardless of the level of output, are termed fixed costs (FC). Costs that directly vary with changes in production are termed variable costs (VC). When fixed and variable costs are added together, the result is total cost (TC).In simple terms the formula below provides the equation for determining total cost:

$$TC = FC + VC$$

As pointed out above, the timeframe is important when categorizing costs based on their relation to output. In the short run, some costs are fixed, since they have already been incurred and they are difficult or impossible to change. In the long run, however, all costs are assumed to be variable, since over time a company is fully able to change fixed costs (by selling and /or altering the fixed cost asset). An example of this could be long-term aircraft leasing contracts. Since airlines are legally and contractually obligated to pay aircraft lease payments to the lessee, in the short run the airline would be unable to avoid these payments no matter how it adjusts output. Therefore, lease payments are a fixed cost in the short run, but in the long run they are variable because, eventually, the contractual obligations will expire. At this point the airline can cut the aircraft from the fleet in order to adjust capacity. There is usually no set timeframe when fixed costs will turn into variable costs; however, the specific situation will dictate when all costs become variable.

Firms and industries have different ratios of fixed to variable costs. The airline industry tends to have high fixed costs, which increases barriers to entering the industry. The ratio of fixed to variable costs is called operating leverage, and this topic will be investigated later in the chapter.

If we divide the total cost function by output, we obtain another important category of costs, called the average costs (Boyes, 2004). By definition, then, average total cost is the total amount of costs (both fixed and variable) per unit of output. Average fixed cost is the fixed costs per unit of output, and average variable cost is the total variable costs per unit of output. Because fixed costs remain constant regardless of output, average fixed costs will decrease with increases in output, since the fixed costs are spread out over a greater range of outputs. This provides an incentive for a firm to increase output if it has high levels of fixed costs. Average variable costs tend to decrease over the first units of production (since so few units are initially produced) and then increase as production increases. Finally, average total costs will behave like average variable costs since they are made up of fixed and variable costs. The formula below describes the relationship between average total costs, average variable costs, and average fixed costs (Maurice and Thomas, 2005):

$$ATC = AFC + AVC$$

Another type of cost that is related to changes in outputs is mixed cost, which exhibits characteristics of both fixed and variable costs. That is, mixed costs are fixed for a certain range of outputs and then increase for a different range of outputs. While fixed costs are fixed in the short run for all outputs, mixed costs have smaller bands of fixed costs and fluctuate to a greater degree. For example, labor can be a mixed cost since it may be difficult to adjust staffing levels to rapidly changing levels of output, but eventually labor costs will have to be adjusted upward or downward to accommodate higher or lower levels of production. Therefore, a mixed cost appears like a step function with various levels.

A final and extremely important type of cost relating to output is marginal cost, which can be defined as the change in total costs resulting from an increase in one additional unit of output (Maurice and Thomas, 2005):

$$MC = \frac{dTC}{dQ} = \frac{\Delta TC}{\Delta Q}$$

Marginal costs are simply calculated by the change in total costs divided by the change in output. For example, it might cost a company $1,000 to produce 20 units and $1,100 to produce 21 units. In this example the marginal cost would simply be $100. In addition, it might cost $1,250 to produce the 22nd unit, so that the marginal cost of this unit would be $150 and so forth. It is important to note that marginal costs generally do not remain constant throughout; therefore, every marginal cost value is unique to that particular change in output.

Explicit/Implicit/Opportunity Costs

Another way of categorizing costs is by explicit and implicit costs. Explicit costs are the costs represented by actual out-of-pocket expenditures, while implicit costs are generally non-cash expenditures. Implicit costs are best thought of in terms of opportunity costs. As mentioned above, because the pursuit of any economic activity represents a choice between two options, the opportunity cost is the value of the next-best option. Therefore, the chosen economic activity must provide a better rate of return than the next-best alternative, otherwise the company would have been better off pursuing the other alternative.

In order to better illustrate opportunity costs, consider the situation an airline faces when it is looking to purchase either a Boeing or Airbus aircraft. Assume that both aircraft have the exact same seating capacity and have similar performance characteristics. The Boeing aircraft costs $70 million, while the Airbus aircraft costs $72 million. The fuel costs for each aircraft type have been estimated to be $1,350 per block hour for the Boeing and $1,370 per block hour for the Airbus. On the basis of these specifications the airline purchases one Boeing aircraft. Here, the explicit costs would be the purchase price of $70 million and the hourly fuel consumption rate of $1,350. The opportunity costs in this example would be $2 million of savings in the aircraft purchase price and $20 additional fuel costs per block hour, since the Airbus aircraft was the next-best alternative for the airline.

Accounting/Economic Costs

The final major method of classifying costs is either accounting or economic costs. Accounting costs generally recognize only explicit costs, while economic costs include both explicit and implicit costs. Therefore, accounting costs do not take into consideration opportunity costs. The two costs also deal with depreciation differently. Accounting costs calculate depreciation based on a predetermined historical usage rate applied against the cost incurred to acquire the asset. On the other hand, economic depreciation takes into consideration changes in technology that may render the useful life of an asset shorter or, conversely, changes in market conditions that may increase the asset's value. Both accounting and economic costs are useful in their context, with accounting costs used primarily for financial accounting purposes and economic costs for the managerial decision-making process.

COST FUNCTIONS

A cost function is a mathematical relationship between total cost and units of quantity produced. While every company has a unique cost function, there are several general forms of cost functions.

The linear cost function can be represented by the formula:

$$TC = a + bQ$$

where: a and b are both constants.

In a linear cost function, the constant, a, would represent the fixed cost, while the constant, b, would represent the marginal cost component, and bQ the total variable cost component.[2] The underlying assumption of a linear cost function is that there are constant returns to scale. That is, successive units of output can be produced for the same cost (in this case the constant, b).While linear cost functions are never attainable for all levels of production in the long run, in the short run they may be approximated within a given production range. A linear cost function generally applies best to a mechanized and

2 For this simple cost function the average variable (bQ/Q) and marginal costs are equal.

automated production line where unit costs are approximately the same. For example, if the production capacity of an assembly line is 300 cars per day, a linear cost function can approximate this process up to that point. For units above 300 per day, a new assembly line would have to be built, and the cost function would shift accordingly.

Consider a linear cost function of:

$$TC = 50 + 5Q$$

Assume that the capacity of the production line is 15 units. Table 4.3 breaks down the cost function into the various classifications.

As is clear from the table, fixed costs remain the same for all levels of output. As is also clear from the table, variable cost rises at a constant five units per unit of output. The total cost function is similar to the variable cost curve, but it includes the fixed cost. Finally, Table 4.3 displays the average and marginal cost functions for the linear example, with the marginal cost being equal to the average variable cost[3] and declining average total and fixed costs for increases in unit output.

To summarize, linear costs occur when the company experiences constant marginal costs of input charges. While the linear cost functions are not typical of most industries, they may be useful approximations over a given range of outputs.

Table 4.3 Cost classifications

Units	TC	FC	VC	AFC	AVC	ATC	MC
1	55	50	5.0	50.0	5.0	55.0	
2	60	50	10.0	25.0	5.0	30.0	5.0
3	65	50	15.0	16.7	5.0	21.7	5.0
4	70	50	20.0	12.5	5.0	17.5	5.0
5	75	50	25.0	10.0	5.0	15.0	5.0
6	80	50	30.0	8.3	5.0	13.3	5.0
7	85	50	35.0	7.1	5.0	12.1	5.0
8	90	50	40.0	6.3	5.0	11.3	5.0
9	95	50	45.0	5.6	5.0	10.6	5.0
10	100	50	50.0	5.0	5.0	10.0	5.0
11	105	50	55.0	4.5	5.0	9.5	5.0
12	110	50	60.0	4.2	5.0	9.2	5.0
13	115	50	65.0	3.8	5.0	8.8	5.0
14	120	50	70.0	3.6	5.0	8.6	5.0
15	125	50	75.0	3.3	5.0	8.3	5.0

3 Recall that the marginal cost is equal to the change in total cost for a one-unit change in output. In the linear case this change is simply equal to the variable cost because there is no change in the variable cost as the output increases.

The second major cost function is the cubic cost function; this has a general form of (Maurice and Thomas, 2005):

$$TC = a + bQ + cQ^2 + dQ^3$$

A cubic cost function represents the normal theoretical cost function. This is a common cost function for many industries since it first exhibits increasing marginal returns and then diminishing returns to scale. The average cost function is typically "u-shaped." Companies need to identify the cost function to determine the extent of economies and diseconomies of scale in order to select the optimal plant size. Consider a company whose total cost function is:

$$TC = 100 + 40Q - 8Q^2 + (2/3)Q^3$$

By plugging this formula into an Excel spreadsheet, Table 4.4 provides the numerical values of the various cost classifications. Also, from this equation we can apply the cost definitions covered earlier to review the various categories of cost (Maurice and Thomas, 2005):

FC = constant a of the general form (in this case, 100)

$$VC = 40Q - 8Q^2 + (2/3)Q^3 \text{ (costs that change with output)}$$

$$AFC = \frac{100}{Q} \quad \text{(fixed costs divided by output)}$$

Table 4.4 Cost classifications

Units	TC	FC	VC	AFC	AVC	ATC	MC
1	133	100	32.7	100.0	32.7	132.7	-
2	153	100	53.3	50.0	26.7	76.7	20.7
3	166	100	66.0	33.3	22.0	55.3	12.7
4	175	100	74.7	25.0	18.7	43.7	8.7
5	183	100	83.3	20.0	16.7	36.7	8.7
6	196	100	96.0	16.7	16.0	32.7	12.7
7	217	100	116.7	14.3	16.7	31.0	20.7
8	249	100	149.3	12.5	18.7	31.2	32.7
9	298	100	198.0	11.1	22.0	33.1	48.7
10	367	100	266.7	10.0	26.7	36.7	68.7
11	459	100	359.3	9.1	32.7	41.8	92.7
12	580	100	480.0	8.3	40.0	48.3	120.7

$AVC = 40 - 8Q + (2/3)Q^2$ (variable costs divided by output)

$$MC = \frac{\Delta TC}{\Delta Q} = \frac{TC_n - TC_{n-1}}{Q_n - Q_{n-1}}$$ (the change in total costs for a one-unit change in output)

In this scenario the company has fixed costs of $100, which remain constant regardless of the level of output. Since the fixed costs remain constant throughout, the average fixed cost declines for every unit increase in output because we are dividing a fixed amount by an increasingly larger quantity of output. This cost function also contains a variable cost function that may include costs such as labor and raw materials. Figure 4.1 displays the fixed cost function for this example, while Figure 4.2 displays the typical variable cost function curve, and Figure 4.3 provides a vertical summation of the cost curves, displaying the total cost function and the fixed cost line.

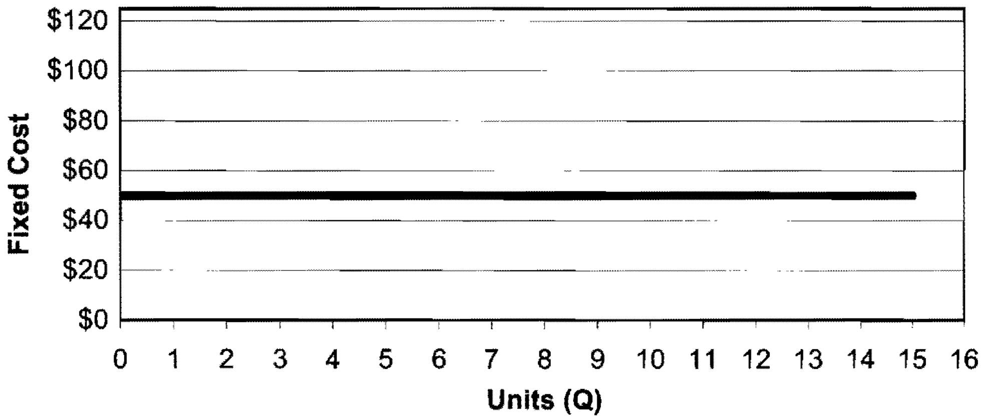

Figure 4.1 Fixed cost function

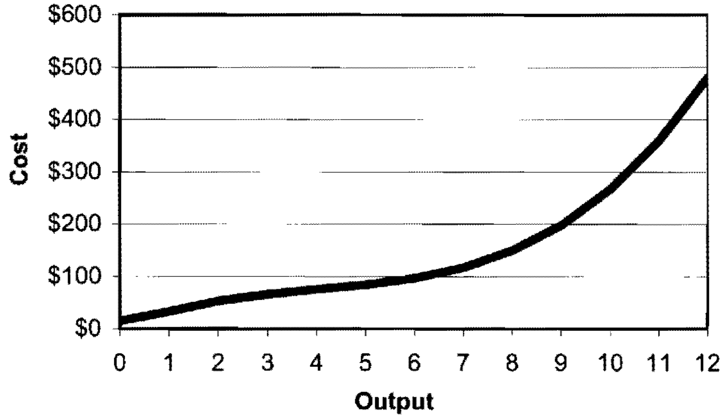

Figure 4.2 Variable cost function

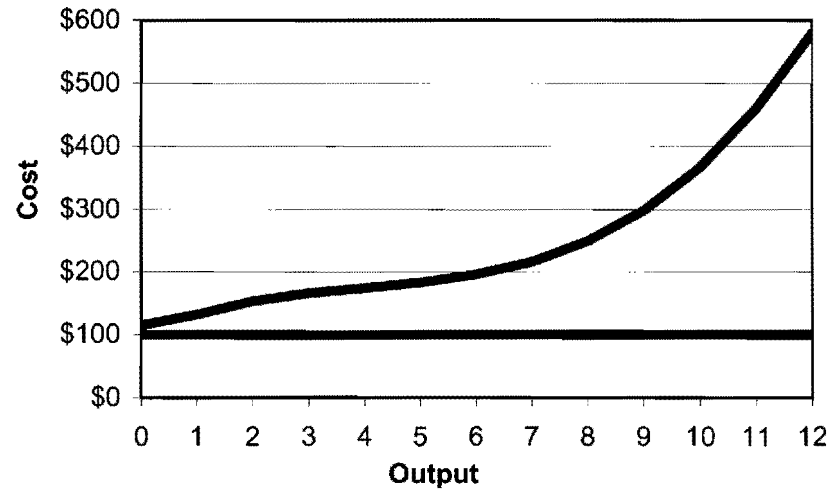

Figure 4.3 Total cost function

Figures 4.1 through 4.3 display the total, fixed, and variable cost functions for the production facility. Since fixed costs are the same for all levels of production, the fixed cost line remains horizontal throughout. The variable cost curve changes slightly for different rates of output, with costs escalating dramatically for higher levels of output. Finally, the total cost curve is simply the variable cost curve shifted up above the fixed cost line.

Figure 4.4 displays the marginal cost curve and the average cost curves for all three cost classifications. From this figure we see that the average fixed cost curve will always slope downwards at a declining rate where it will eventually be asymptotic with the x-axis. The average total cost curve is u-shaped which is similar to the average variable cost, since they differ by only the average fixed costs. The marginal cost curve crosses both the average variable cost and average total cost curves at their minimum point and continues above them as output rises. The intuitive reason for this is that, at first, we are adding unit costs that are less than the average—causing the average costs to fall. Then, as these additional unit costs increase, they eventually become equal to the average (at the minimum of the average). And, finally, the additional unit costs are above the average—causing the average costs to rise.

Using Cost Functions for Managerial Decisions

While it is instructive to review the cost categories, it is even more important to understand how they affect managerial decision-making. In this regard the cost curves provide the following rules for managerial decision-making:

- Marginal cost tells us where (how much) to produce.
- Average total cost tells us whether we should produce.
- Average variable cost tells us when we should cease production.

The intuitive explanation for this follows. Since marginal cost is the cost for extra units of output, it is obvious that, as long as the revenue from the extra unit exceeds its

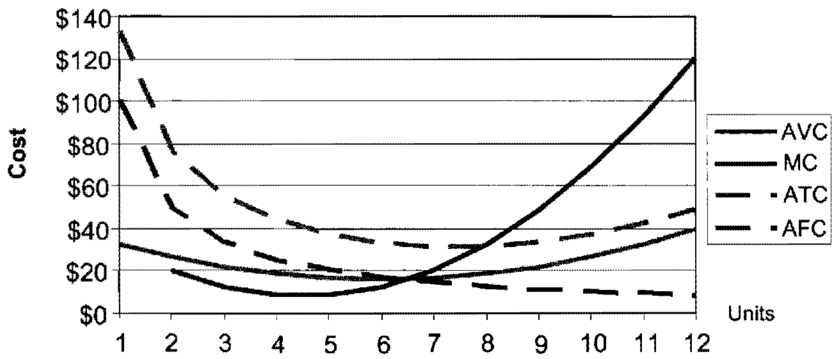

Figure 4.4 Average and marginal cost functions

cost, then we will want to produce more. Conversely, if the cost for that unit exceeds the revenue, then we will not want to produce that unit. Hence, we will continue production up to the point where marginal cost and marginal revenue are equal, but not beyond that point, because then the cost of producing that unit would exceed the revenue that could be achieved by selling it.

Average total cost can be multiplied by output to give total costs. As long as total revenues exceed total costs, we will, of course, continue production. In the event that total cost falls below total revenue but is still above variable costs, then we would still continue production (in the short run) since we are still making something to help pay our fixed costs. Recall that fixed cost must be paid in the short run whether we are producing or not. Hence, average total cost tells us whether we should continue production.

$$ATC = \frac{TFC + TVC}{Q} = \frac{TC}{Q}$$

$$TC = ATC \times Q$$

Finally, we have average variable cost which can be translated into total variable cost by multiplying by output. In this case, if our total revenues do not cover our total variable costs, then we should cease production since we will be losing more money on every unit we produce. A third type of cost function typically prevalent in the industry is the quadratic function, which has a general form of:

$$TC = a + bQ + cQ^2$$

This can exhibit decreasing returns to scale for the range of production possibilities since the curve is upward-sloping at an increasing rate. The total cost curve then becomes asymptotic at the capacity level of the production range.

The total cost function can be derived to determine both the average cost function and the marginal cost function for all three types of cost function. In order to determine the average cost function, simply divide the total cost function by Q. The marginal cost function is found by taking the first order derivative of the total cost function.

Consider the linear cost function:

$$TC = 50 + 5Q$$

By dividing the function by Q, the average cost function for a linear cost structure is:

$$ATC = \frac{50}{Q} + 5$$

Taking the first-order derivative of a linear function,[4] the MC = 5, which is also the constant b and the variable cost of the linear function.

For a cubic cost function, both the average and marginal cost functions are found to be quadratic. From the equation:

$$TC = 100 + 40Q - 8Q^2 + \frac{2}{3}Q^3$$

the average cost function is

$$ATC = \frac{100}{Q} + 40 - 8Q + \frac{2}{3}Q^2$$

The marginal cost function for the cubic cost function is:

$$MC = 40 - 16Q + 2Q^2$$

which is a quadratic cost function. The second-order derivative of the total cost function (or first-order derivative of the marginal cost curve) determines if the u-shape is a minimum or maximum and at which point the marginal costs are minimized. For this example, the second-order derivative would be:

$$\frac{dMC}{dQ} = -16 + 4Q$$

This point represents the minimum point of the marginal cost curve, which corresponds to the data in Table 4.4, and, since the value is positive, we know that the point is a minimum value. Conversely, if the second-order derivative yielded a negative value, then the point would be a maximum value.

Finally, the derivations of the quadratic total cost function would have the form of:

$$\frac{dTC}{dQ} = \frac{a}{Q} + b + cQ$$

4 Recall that the derivative is simply the change in the dependent variable for a one-unit change in the independent variable. In this case, a one-unit change in Q results in a five-unit change in TC.

for the average total cost, and

$$MC = b + 2cQ$$

for the marginal cost curve. This marginal cost curve is linear in nature with a constant slope, indicating that, for every increase in output, costs will increase at a greater rate. This is the basic definition of decreasing returns to scale.

ECONOMIES OF SCALE, SCOPE, AND DENSITY

Having covered the cost curves in some detail, we can now use these concepts to define economies of scale, scope, and density. These factors can play a crucial role in management decision-making and the cost structure of not only the company, but also the industry as a whole.

Economies of scale are advantages gained when average unit costs decrease with an increase in the quantity being produced. Economies of scale are common in highly capital-intensive industries with very high fixed cost, such as aircraft manufacturing, airline industry, railroads, and the steel industry.[5] The following are the sources of economies of scale:

- lower cost and higher productivity due to the division of labor
- increase in labor productivity due to concentration on a fewer number of tasks and more experience on the job (learning curve).

The reverse is diseconomies of scale, where average unit costs increase with an increase in production quantity. There are different causes for diseconomies of scale:

- an inability to efficiently coordinate material flows and manage employees due to larger facilities
- a slow decision ladder
- workers and management becoming more segregated and communication becoming less effective
- inflexibility
- capacity limitations on entrepreneurial skills (hiring qualified employees).

Depending on the firm's cost function, economies of scale usually do not exist for every level of production. Companies will have levels of quantity where economies of scale are present and levels where diseconomies of scale exist. This is exactly what occurs for cubic cost functions. Referring back to Figure 4.4, the average total cost function is u-shaped, indicating economies of scale for production quantities of 1 through 7 and diseconomies of scale for production quantities 8 and above.

Economies of scope refer to the situation in which the company can reduce its unit costs by leveraging efficiencies through sharing resources for multiple projects or production

5 Because of very high fixed cost only a few companies can stay in the market. As a result, there are only two aircraft manufacturing companies in the world (Boeing and Airbus) that manufacture large commercial aircraft.

lines. Put more simply, multiple projects/processes can be more cost-efficient when they are done together rather than individually. The presence of economies of scope provides benefit by allowing a company to house activities together and concurrently. Possible synergies achieved through economies of scope could include shared labor, shared knowledge, and shared capital equipment. For example, Boeing houses three production lines (747, 767, and 777) at its large Everett production facility (Boeing, 2006). By operating three production lines in the same building, Boeing is able to share resources such as labor and equipment between all three lines to maximize resource efficiency. If all three production lines were individually located throughout the country, Boeing would not be as cost-effective in manufacturing aircraft since resources could not be shared among all three production lines. Also, Boeing is able to leverage capital knowledge by reducing R&D expenses through utilizing technology developed in other projects. These savings can be considerable when developing new aircraft such as the Boeing 787. Another example of economies of scope would be Aer Lingus, which in 1970 began to seek new sources of revenue by offering engineer training, maintenance services, computer consulting, and data processing services to other airlines (Harvard Business School, 2000).

Economies of density are achieved through the consolidation of operations. The airline industry is a good example of this as it has developed the so-called hub-and-spoke system for air travel. That is, airlines have found it more cost-effective to consolidate operations at a single airport rather than operate a point-to-point service. For example, consider five airports that could all be connected together either by using one airport as a hub (Figure 4.5), or by flying between each city (Figure 4.6). Using a hub, all the airports can be connected to each other with a minimum of four flights, while the point-to-point service would require ten flights. The difference in the number of required flights is a cost saving as the airline can provide service to all the cities with less resources. Of course there are other benefits to flying point-to-point services, such as being able to offer nonstop flights for the passengers. In addition, the advent of the regional jet has enabled airlines to offer an increased amount of point-to-point flying. However, the hub-and-spoke network is still the dominant flying structure in the United States, and will probably continue to be so for the foreseeable future.

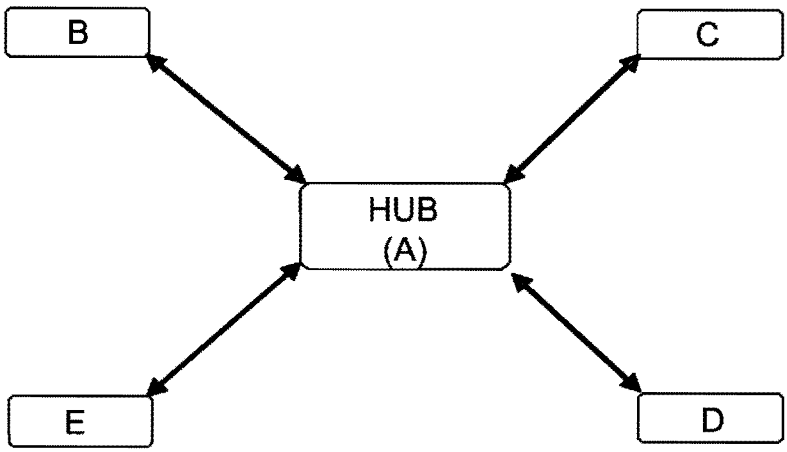

Figure 4.5 Hub-and-spoke route network

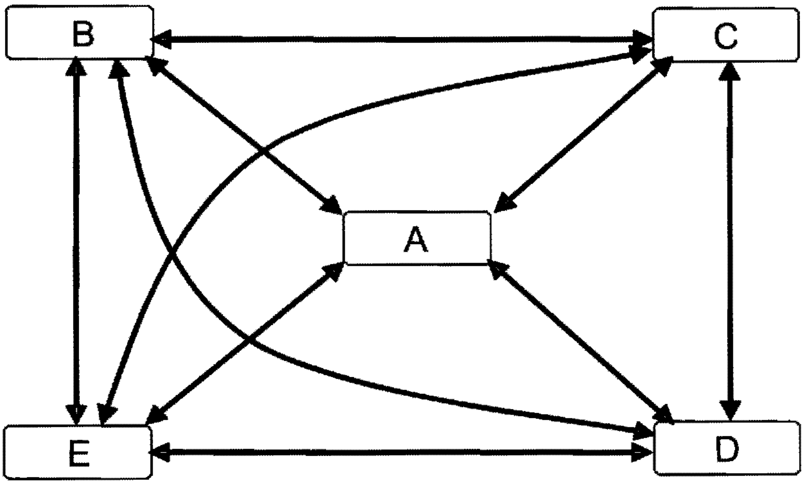

Figure 4.6 Point-to-point route network

The economies of scope and density of a hub-and-spoke network can also be stated mathematically. Figure 4.5 displays a typical hub-and-spoke diagram for a network utilizing five airports, while Figure 4.6 displays a point-to-point route network also for the five airports. Clearly, by simply looking at the diagrams, the hub-and-spoke network has significantly fewer flights. The number of flights in a hub-and-spoke network can be calculated by using the formula:

$$\text{Number of Flights} = n - 1$$

where: n is the number of airports.

Using this formula, the number of flights required to connect five airports in a hub-and-spoke network is 4 (5 − 1). These four flights are displayed in Figure 4.5.

The mathematical number of flights required in a point-to-point system can also be determined by using the formula:

$$\text{Number of Flights} = \frac{n \times (n-1)}{2}$$

Using the above formula, the number of flights required in a point-to-point network for five airports is ten.[6] This reduction in flights is the major reason why the hub-and-spoke network has been adopted by almost all carriers in the United States in the post-deregulation period. Southwest Airlines is the sole major US carrier that has decided to adopt the point-to-point route network, even though the majority of Southwest passengers make connections between flights.

6 $5(5-1)/2 = 10$.

AIRLINE INDUSTRY COST STRUCTURE

Since every airline's operation is unique, it can be difficult to compare airline operating costs from airline to airline. The most common metric used to standardize airline costs are CASMs (costs per available seat miles). An available seat mile (ASM) is one aircraft seat, flown one mile, regardless of whether it is carrying a revenue passenger. Costs per ASM, or CASM, are the cost of flying one aircraft seat for one mile. CASMs can be created for a variety of costs, such as operating costs, total operating costs, or simply crew costs. Table 4.5 provides a breakdown of various CASMs for 12 major US airlines in 2005, with the total CASM representing the direct operating costs of fuel, labor, maintenance, and other operating and non-operating costs (typically called overhead costs). The four direct operating costs can be considered as variable costs, while the non-operating costs can be considered as fixed costs. Figure 4.7 displays this information graphically.

Figure 4.7 displays the composition of the airlines' various costs. For all the airlines, the non-operating costs comprise the largest share of the airline's total costs. In many cases, this share is over 50 per cent of the airline's total costs. In United's case, non-operating costs represent 80 per cent of the airlines' total cost function, but this is largely a result of bankruptcy costs incurred by the airline in 2005. Similar situations also occurred with Delta and Northwest in 2005. Since non-operating costs are fixed with output, this chart confirms that airlines have large fixed costs. Acquisitions, such as aircraft, and investment in airport infrastructure are large capital expenditures; these in turn create an industry with high barriers to entry. Recently, however, these capital requirements have been lowered by the advent of attractive aircraft leasing options, thereby enabling easier entry into the market.

Direct operating costs (DOC) are those costs that are directly attributable to the airline's operations. The principal categories of direct operating costs, which will be discussed in more detail in the following paragraphs, are listed below:

- fuel costs
- flight and cabin crew expenses

Table 4.5 US airline CASM breakdown, 2005

US Airline CASM Comparison (2005)											
	AA	AS	B6	CO	DL	FL	HP	NW	UA	US	WN
Fuel Cost per ASM	2.8382	2.6562	2.0535	2.5988	2.6911	2.6716	2.5249	3.0846	2.6242	2.9009	1.7182
Maintenance Costs per ASM	1.1693	1.0132	0.5764	1.0567	0.8374	0.7925	1.1252	1.1537	1.2209	1.1029	0.7912
Crew Costs per ASM	1.1359	1.3161	0.5807	1.0759	1.2328	0.9009	0.8698	1.4777	0.8261	0.9627	1.107
Other Operating Costs per ASM	0.8146	1.1302	0.7725	1.2285	1.0711	1.93	1.2965	1.1318	0.8274	1.1949	0.6827
Total Operating Costs per ASM	5.958	6.1157	3.9831	5.9599	5.8324	6.295	5.8164	6.8478	5.4986	6.1614	4.2991
Non Operating Costs per ASM	6.3453	4.4448	3.2906	10.3861	10.2186	3.1305	6.0628	11.4121	23.3449	8.7326	3.5615
Total Costs per ASM (CASM)	12.3033	10.5605	7.2737	16.346	16.051	9.4255	11.8792	18.2599	28.8435	14.894	7.8606

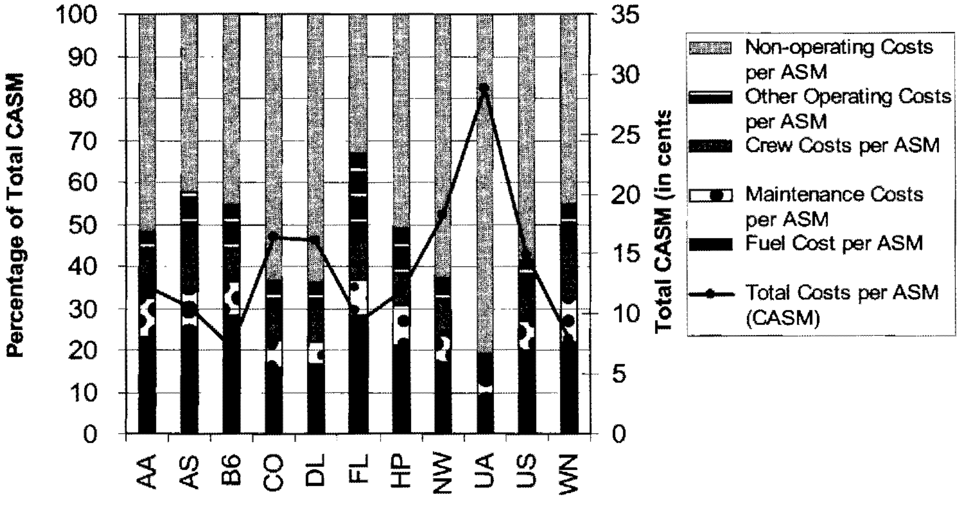

Figure 4.7 US airline CASM breakdown, 2005

Source: Compiled by the authors from Back Aviation Form41 data.

- direct maintenance expenditures
- other operating costs, including landing fees and capital equipment charges.

Fuel Costs

In terms of direct operating costs, fuel costs in 2005 represented the greatest share of an airline's direct operating costs. This was due to the record high prices of oil that were experienced in that year. The general formula for airline fuel costs is:

$$\text{Fuel Costs} = \text{ASM} \times \frac{\text{Fuel Price per Gallon}}{\dfrac{\text{ASM/Block Hour}}{\text{Gallons/Block Hour}}}$$

An airline's fuel cost per available seat mile is a result of two factors: the price of fuel and fuel efficiency. Although the price of fuel is generally beyond the airline's control, the airline can lessen the impact of this cost by using more complex investment strategies. For example, Figure 4.7 shows that Southwest's fuel costs were approximately 22 per cent of the airline's total costs. And, even though Southwest's fuel cost per ASM of 1.72 cents was the lowest of the 12 airlines compared in Table 4.5, this was not the direct result of Southwest's fuel-hedging strategy. In fact, Southwest purchased options for the price of oil and not for the price of aviation fuel—thus purchasing aviation fuel at the market bearing rate—and offset this with investment gains from oil options. Investment strategies such as these help offset increases in the cost of fuel. However, these gains are recorded as investment gains in the airline's consolidated financial statements and not as fuel cost benefits. Therefore, Southwest's advantage in fuel costs is a result of effective fuel management strategies.

The second way in which airlines can adjust their fuel costs is by being more fuel-efficient. The simplest way to do this is to operate new, fuel-efficient aircraft rather than older, less fuel-efficient aircraft. Although the capital expenditures required to purchase new aircraft are considerable, they may outweigh the fuel costs associated with operating older aircraft. Of all the airlines listed in Table 4.5, Northwest had the highest fuel cost per ASM at 3.08 cents. This is largely due to Northwest operating older DC-9 and DC-10 aircraft. The reverse is true of airlines such as Southwest and JetBlue that operate mostly newer-generation aircraft that are more fuel-efficient.

Other fuel efficiency methods center on technological advances—for example, the installation of blended winglets. Aviation Partners Boeing, the joint-venture company that manufacturers blended winglets, estimates that the winglets reduce fuel burn by 3.5–4.0 per cent on flights greater than 1,000 nautical miles for Boeing 737NGs (APB, 2006). The winglet technology does not provide substantial savings on short flights as the fuel-burn advantage is offset by the increased weight. Winglets were originally offered on just 737NG aircraft, but their success has led to their installation on a number of different aircraft types.[7]

While the methods highlighted above usually require substantial capital investments to reduce fuel costs, more subtle fuel management strategies by airlines can also increase fuel efficiency. One strategy commonly employed by airlines is to use only one engine during normal taxiing procedures, thereby reducing the fuel costs associated with operating an engine during these maneuvers. The more congested the airport is, the greater the amount of taxiing, and the bigger savings such a program can provide. In addition, airlines can selectively shut down an engine(s) during ground delays when the aircraft is forced to sit idle.

Flight planning plays a significant role in fuel efficiency in that it can help plan flights for minimum fuel-burn routes and altitudes. Flight planning can also be enhanced by measuring onboard weight more accurately in order to avoid carrying extra fuel around.

Altering the location where fuel is purchased allows airlines to take advantage of lower fuel prices in certain regions. Employing this strategy involves a cost–benefit analysis as the fuel cost savings from lower prices needs to be compared with the additional fuel burn generated from the additional weight involved in tankering the extra fuel to the desired region. Airlines can also pool resources when purchasing fuel in order to achieve bulk discounts. This is a strategy a few Star Alliance carriers have explored. It has been estimated that joint fuel purchasing could provide around $50 million in savings for alliance members over three years (Ionides, 2004).[8]

At airports, airlines can employ self-imposed ground delays to reduce airborne holding, and can also redesign hubs and schedules to reduce congestion.

All these factors can and do contribute to airlines being more fuel-efficient. Figure 4.8 provides a comparison of fuel efficiency for 12 US airlines in 2006, in terms of domestic ASMs per gallon. In Figure 4.8 the metric is ASMs per gallons of fuel used; this means that a higher value is more desirable since airlines want more ASMs to be flown on one gallon of fuel. The figure also gives the average of the carriers to aid comparison of individual airlines against the industry.

Figures 4.8 and 4.9 show that JetBlue is the most fuel-efficient airline; this is largely due to the fact that its average stage length is the longest of all the airlines. Stage length is important since longer flights burn less fuel per ASM. Because aircraft takeoff and landing use the most fuel per ASM, it follows that the longer the flight, the more fuel-efficient ASMs there are to dilute the less efficient takeoff and landing phases. Furthermore, JetBlue

7 Winglets are being installed by airlines on 727s, 757s, and 737 Classics. Airlines such as Continental and Southwest have heavily promoted the use of winglets, which have resulted in decreased fuel costs per ASM.

8 This initiative was being employed at four major airports worldwide: Los Angeles, London Heathrow, Paris Charles de Gaulle, and San Francisco

operates a new fuel-efficient fleet of Airbus A320 aircraft, and this significantly increases its fuel efficiency.

Figure 4. 10 represents the historical data: using 2003–2006 historical information yields an extremely high correlation coefficient of 0.8552 between fuel efficiency and average stage length. An anomaly to this correlation was Southwest Airlines which has relatively high fuel efficiency, but the lowest average stage length of all the airlines. Southwest has managed to overcome its short stage length with operating procedures that conserve fuel; these include using only one engine when taxiing and opting to fly out of less congested airports.

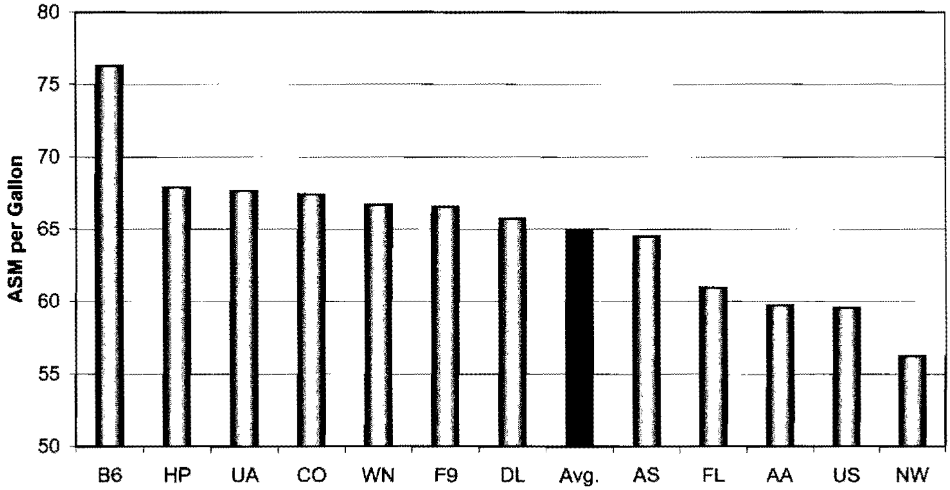

Figure 4.8 US Airline fuel efficiency, 2006

Source: Compiled by the authors from Back Aviation Form41 data.

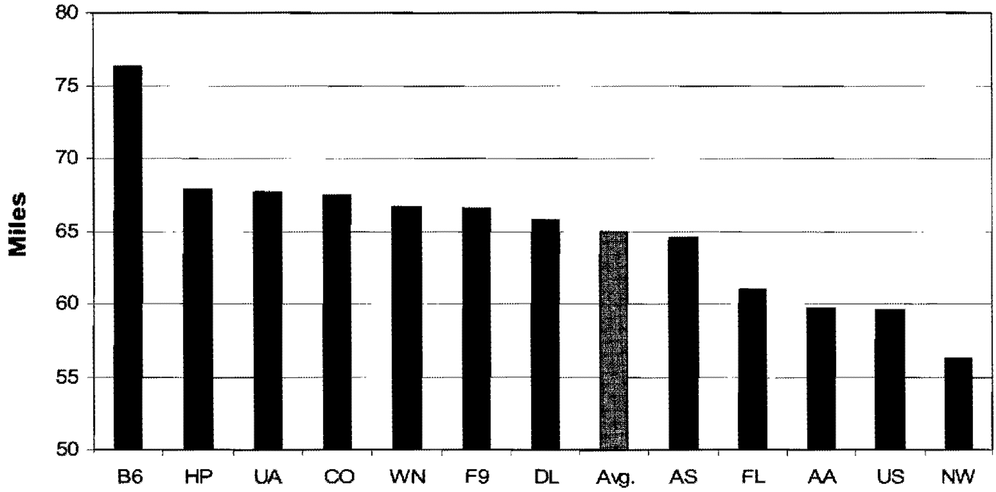

Figure 4.9 US airline average stage length, 2006

Source: Compiled by the authors from Back Aviation Form41 data.

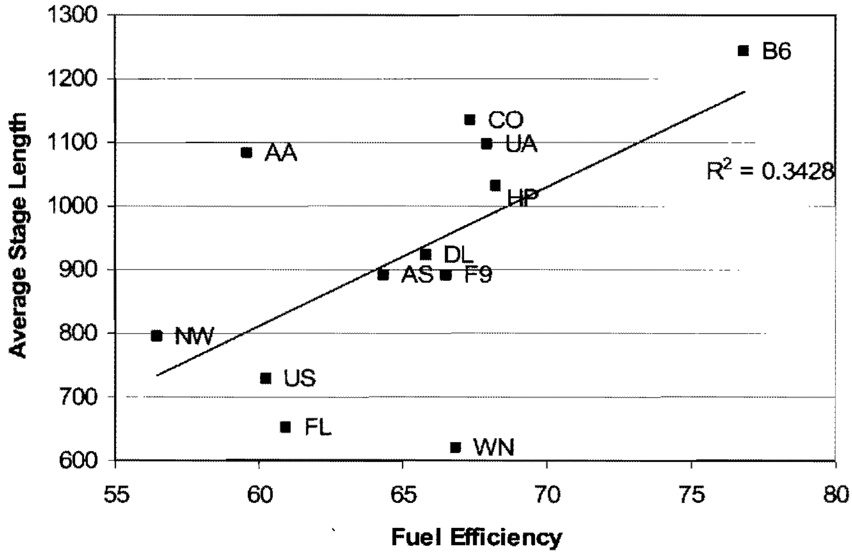

Figure 4.10 Correlation between fuel efficiency and average stage length, 2006

Source: Compiled by the authors using Back Aviation Form41 data

Flight/Cabin Crew Expenses

The next greatest direct operating cost for US airlines is crew expenses. Prior to the recent downturn in the health of the aviation industry, crew costs constituted a much greater proportion of an airline's direct operating costs; however, mainly through bankruptcies, airlines have recently been able to dramatically reduce their labor costs. Since most airlines deal with a heavily unionized labor force, it can be difficult for airlines to adjust labor input to output, causing the crew costs to resemble mixed costs. In other words, contractual agreements with labor groups make it difficult for the airline to furlough employees, causing long lag times for airlines to respond to decreasing travel demand and output. Productivity gains may also be limited by unions and government regulations concerning work rules. The general formula for flight personnel costs is:

$$\text{Flight Personnel Costs} = \text{ASM} \times \frac{\text{Labor Rate}}{\text{ASM/Block Hour}}$$

Figure 4.11 provides a comparison of crew costs (both flight deck and cabin crew) for 12 major US airlines. The data in Figure 4.11 closely resemble the crew costs per ASM in Table 4.5, but block hours are the most common measurement metric of crew costs in the airline industry.

Three low-cost carriers, AirTran (FL), Frontier (F9), and JetBlue (B6) had the lowest crew costs per block hour in 2006. The general assumption that low-cost carriers always pay the least is somewhat inaccurate, however, since Southwest Airlines (WN) is in the middle in terms of crew costs. Southwest is effective through a more efficient use of its employees, who are expected to perform many different tasks in addition to their primary duties. These productivity gains help Southwest offset its higher pay rate. AirTran, Frontier, and

JetBlue benefit from being relatively young companies, and therefore have a relatively younger workforce on a lower pay rate. Airlines such as American (AA) and Continental (CO) have been around so long that many crew members are quite senior and command a higher pay rate than junior members. This factor is enhanced when pension and Medicare issues are included in the crew cost calculations.

In 2006 Delta (DL) had the highest crew costs in the industry; however, Delta now believes that bankruptcy proceedings will enable it to reduce their crew costs per block hour (by renegotiating abrogated crew contracts).[9] While Delta's costs still remain high, they have come down considerably from 2003 and 2004 levels thanks to the bankruptcy protection. Another example of the benefits of bankruptcy protection is United Airlines (UA) which at one time had the highest crew costs in the industry, but now has crew costs per block hour below those of Southwest Airlines.

The recent trend in the aviation industry has been toward significant reductions in crew costs as US airlines posted record losses. However, Southwest's crew costs per block hour have actually increased in the last three years and it remains to be seen if these costs are sustainable.

Maintenance Costs

Another considerable cost for airlines is maintenance costs. However, since safety is the number one priority for every airline, maintenance costs are usually not under as much

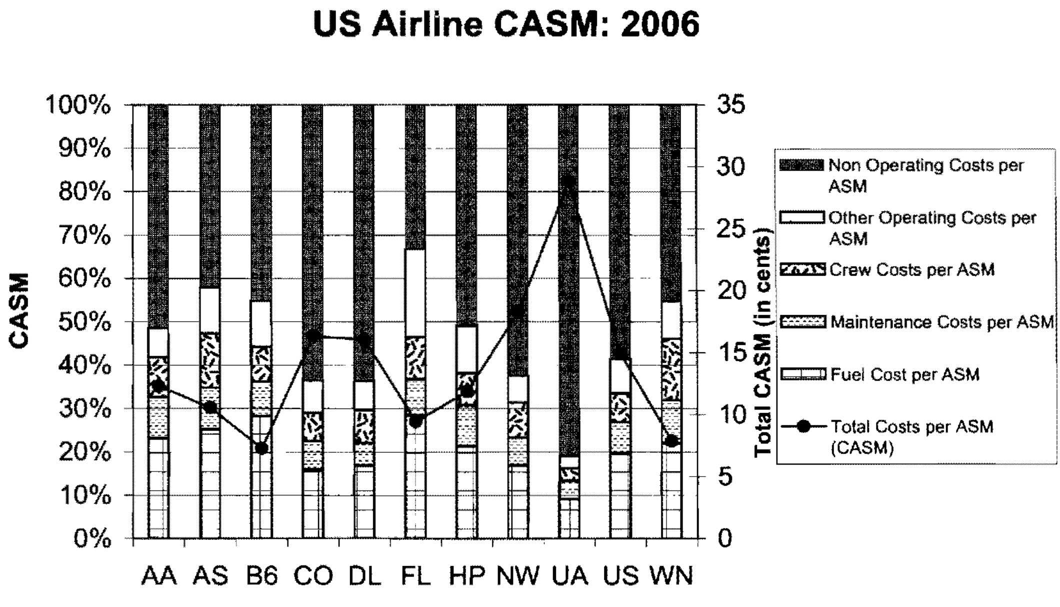

Figure 4.11 US crew costs per block hour, 2006

Source: Compiled by the authors from Back Aviation Form41 data.

9 Delta Air Lines emerged from bankruptcy in April 2007, following a 19-month reorganization and a fight against a hostile takeover.

cost-saving scrutiny. That said, airlines must still cost-effectively manage their maintenance operations and staffing levels, while also being safety-conscious. In order to accomplish this, a major innovation in the maintenance area has been the outsourcing of maintenance activities to third-party vendors, especially for aircraft heavy checks. Maintenance costs can be calculated by using:

$$\text{Maintenance Costs} = \text{ASM} \times \frac{\text{Maintenance Labor and Materials/Block Hour}}{\text{ASM/Block Hour}}$$

Each airline is unique in the amount of outsourcing that it does. American Airlines, for example, does the majority of its maintenance internally, and also does contract maintenance work for other airlines. On the other hand, Continental Airlines does all its wide-body heavy maintenance externally, but has internalized some maintenance operations with the opening of a 757 heavy check line. At the opposite end of the spectrum, JetBlue externally sources almost all its heavy maintenance and performs only line maintenance internally.

Figure 4.12 displays the maintenance costs for twelve US airlines, including outside labor costs; these costs are standardized per flight hour since flight hours are the primary driver of an aircraft's maintenance cycle. The data in Figure 4.12 show that in 2006 Northwest Airlines (NW) had the highest maintenance costs, while Frontier (F9) had the lowest. One possible explanation for the large difference in maintenance costs is that Frontier's fleet is quite new whereas Northwest's is fairly old. Airlines enjoy a maintenance honeymoon on new aircraft because costly heavy checks are not required for a few years, but older aircraft have more frequent and costlier heavy checks. Thus there is some relation between the airline's average aircraft age and maintenance costs. JetBlue's maintenance costs from 2003 to 2006 are a good example of the effect of aircraft coming off their maintenance honeymoon; the airline's costs have increased substantially as a result of its aging fleet.

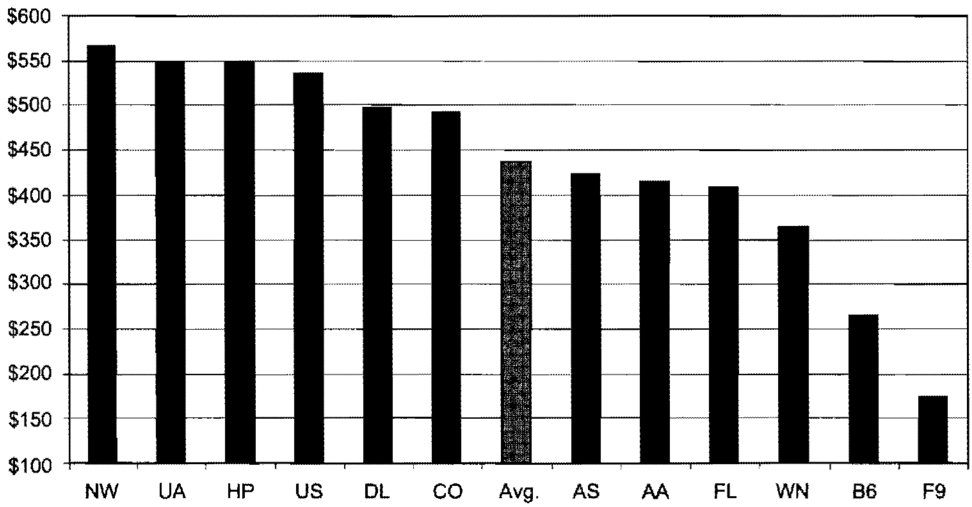

Figure 4.12 US airline maintenance costs per flight hour, 2006

Source: Compiled by the authors from Back Aviation Form41 data.

Related to maintenance costs are maintenance checks; as these are more expensive for larger aircraft, airlines such as Frontier benefit from operating only narrow-body aircraft, unlike Delta, United, and Northwest which operate extensive wide-body fleets. Finally, aircraft commonality plays an important role in the maintenance costs for airlines. Airlines with diverse aircraft fleets usually need a large inventory of spare parts on hand for each aircraft type. One way round this problem is to segment aircraft markets. This is the primary reason why American Airlines decided to operate only certain aircraft types out of certain hubs. For example, MD-80s do not operate out of the Miami hub, 738s do not operate out of Chicago, and the A300s do not fly to either Dallas or Chicago. This rationalization means that the airline does not have to stock spare parts for every aircraft type at every major hub.

Other Operating Costs

The final category of operating costs included in Table 4.5 was other operating expenditures. These costs can include a variety of things, such as airport-related expenditures (that is, landing fees, gate agents, and baggage handlers) and in-flight catering costs. Being easy areas for immediate cost-cutting, these expenditures have shown dramatic reductions, particularly catering. The majority of airlines now use third parties to supply their airport and catering services. From Table 4.5, America West (HP), along with Continental Airlines (CO) had the highest other operating costs per ASM. Continental's costs can possibly be explained by the fact that it still owns its own catering kitchen. If airlines once again begin to offer more elaborate services, other operating expenses should be expected to increase. Other operating expenditures, or indirect costs, can also be described using the simple formula:

$$\text{Indirect Costs} = \text{ASM} \times \frac{\text{Indirect Costs}}{\text{ASM}}$$

Indirect operating costs can also be grouped as the following:

- distribution cost (sales and promotion)
- station cost
- ground expenses
- passenger services
- administrative expenses.

AIRLINE ECONOMIES OF SCALE, SCOPE, AND DENSITY

The airline industry is affected by economies of scale, scope, and density. Economies of scale play a significant role in the industry since, as we have seen earlier, fixed costs are extremely high. These high fixed costs and marketing requirements encourage expansion in the industry. This is a major reason why there are very few small airlines in the industry and why there is continual consolidation. From an operational standpoint also, economies of scale play a significant role. Because of the need to train pilots and maintain a stock of spare aircraft parts, it is less costly to simplify aircraft fleets and focus on just a few aircraft types. Therefore, airlines generally prefer to operate a minimum number of aircraft types where possible. The optimum number of aircraft required to achieve economies of scale is unknown, but JetBlue believes that it achieved all economies of scale from its 80-strong Airbus 320 fleet. The airline decided to order an additional fleet of 100 Embraer 190 aircraft in 2006 (JetBlue Airways, 2006). Thus, although the airline has decided to use two types of aircraft, it believes that both fleets are of sufficient size to maximize economies of scale.

Hub airports also contribute significantly to economies of scale, in addition to the economies of density mentioned earlier. Hubs are extremely costly operations, and the costs that they generate, such as multiple labor shifts, terminal leases, and ground equipment, are fixed costs in the short term. Therefore, in order to spread the costs over more units of output (air seat miles), airlines have a strong incentive to use these assets as intensively as possible. While most airlines operate banked hubs[10] to provide shorter connection times for their passengers, airlines such as American and Delta have experimented with rolling hubs in order to better utilize hub assets. With banked hubs, assets sometimes remain unused for extended periods of time between banks. On the other hand, rolling hubs use hub assets throughout the day, making the airline's use of assets more efficient and thereby achieving greater economies of scale.

Economies of scope also play an important role in the aviation industry. As mentioned earlier, aircraft manufacturers such as Boeing capitalize on economies of scope when producing aircraft. Airlines achieve economies of scope by operating various ancillary programs/services, such as frequent-flyer plans, maintenance activities, catering, and ground handling. Ultimately, the amount of outsourcing an airline does depends on how many synergies exist between the organizations that are creating economies of scope. For example, frequent-flyer programs are usually more effectively run by the airline itself, since a frequent-flyer program's main cost is inventorying reward seats, and economies of scope are achieved when this is done internally (because the airline already has a staff to schedule seats). One frequent-flyer plan, Air Canada's Aeroplan, is bucking this trend by allowing their loyalty program[11] to become a separate operation. Airlines have also experimented with economies of scope by becoming involved in other related industries (such as cruise lines and hotels), but history has shown that most of these external activities have failed (for example, Sabena hotels, Carnival Airlines), and the economies of scope that were envisioned either did not exist or were not achieved.

Economies of density exist in the airline industry through the use of hubs and the consequent reduction of flights. Similarly, aircraft size exhibits both economies of scale and density. Economies of scale are realized when airlines can put more seats into the aircraft to reduce unit costs, but economies of density are also achieved by using larger aircraft. For example, suppose an airline could use a 100-seat aircraft or two 50-seat aircraft to service a route. On shorter domestic flights the airline may opt for a higher frequency of flights, while on longer domestic and international flights it will usually select the larger aircraft in order to capture the economies of density (that is, the extra costs of pilots, gate agents, landing fees, and baggage-handling for the high-frequency decision). While airlines typically have moved toward smaller aircraft to provide increased frequency on a route, this strategy is less effective if the flights have to depart within minutes of each other. In this case, the single larger aircraft will always be cheaper to operate than the two smaller aircraft on account of

10 A banked hub is one that brings the aircraft into the hub airport in as short a time as is feasible so that passenger connections are as short as possible. On the other hand, a rolling hub stretches the arrival over a longer time period so as to avoid the airport and airspace congestion that may arise from a banked hub.

11 A similar type of situation would be catering companies, but with airlines largely ignoring food services on flights, this no longer becomes an area where airlines can achieve significant economies of scope.

the various fixed costs required to operate a flight (that is, pilots, gate agents, landing fees, and baggage-handling). This trade-off benefit is another example of economies of density and is common in the airline industry, especially on very long-range flights, where (due to time zones) most flights depart at around the same time.

AIRLINE BREAK-EVEN ANALYSIS

An important measurement of any company's cost structure is its break-even analysis. Break-even analysis is the number of units or revenue required in order for the firm's costs to be recovered. In manufacturing, the break-even point is represented in product units. For example, the revised break-even forecast for the Airbus A380 program is 420 aircraft (Shannon, 2006). Airbus will have to sell 420 aircraft to simply recoup the fixed costs related directly to the A380 program.

The basic formula for break-even in units is:

$$Q_{B\text{-}E} = \frac{\text{Fixed Costs}}{(\text{Price - Variable Costs})}$$

Since the difference between the price and variable costs of a good is also called the contribution margin, the break-even formula can be rewritten as:

$$Q_{B\text{-}E} = \frac{FC}{P-V} = \frac{FC}{Contribution\ Margin}$$

Break-even in the airline industry is usually expressed as a percentage of total ASMs. This provides a break-even load factor, or a load factor which the airline must meet to recover all its fixed costs. The general formula for airline break-even is the following:

$$\text{RPM} \times \text{RRPM} - \text{ASM} \times \text{CASM} = 0$$

where:
- RPM = revenue passenger miles
- RRPM = revenue per revenue passenger mile
- ASM = available seat miles
- CASM = cost per available seat mile.

From this basic formula, two passenger load factors can be found—actual load factor and break-even load factor. The formula derivations to achieve the two ratios are:

$$\text{RPM} \times \text{RRPM} = \text{ASM} \times \text{CASM}$$

$$\frac{\text{RPM}}{\text{ASM}} = \text{Load Factor}$$

$$\text{Breakeven Load Factor} = \frac{\text{CASM}}{\text{RRPM}}$$

$$LF_{B-E} = \frac{\text{CASM}}{\text{RRPM}}$$

Obviously, if the actual load factor is greater than the break-even load factor, then the airline is making enough money to cover its fixed costs. However, if the actual load factor is less than the break-even load factor, then the airline is losing money. Figure 4.13 shows the actual load factor and the break-even load factor for 12 major US airlines in 2006.

From Figure 4.13, two airlines, Delta (DL) and Northwest (NW), have break-even load factors greater than 100 per cent. This implies that both these airlines are currently incapable of recovering their fixed costs, and it is therefore not surprising that both carriers are in bankruptcy protection. US Airways (US), AirTran (FL), and Southwest (WN) all have actual load factors that exceed their break-even load factor. This means that these airlines are presently earning enough revenue to cover their fixed costs. The figure also shows that these airlines also have the lowest break-even load factors, indicating that their cost structures are very efficient. Not surprisingly, Southwest leads the industry with a break-even load factor of roughly 65 per cent.

OPERATING LEVERAGE

Operating leverage is a powerful metric that highlights the ratio between operating profit growth and sales growth. More directly, the degree of operating leverage is an elasticity of the overall company's financial health with respect to sales growth. Operating leverage can also provide an indication of the company's cost structure, especially with respect to fixed costs. The general formula for the degree of operating leverage is:

$$DOL = \frac{\%\Delta \text{ } in\, EBIT}{\%\Delta \text{ } in\, Sales}$$

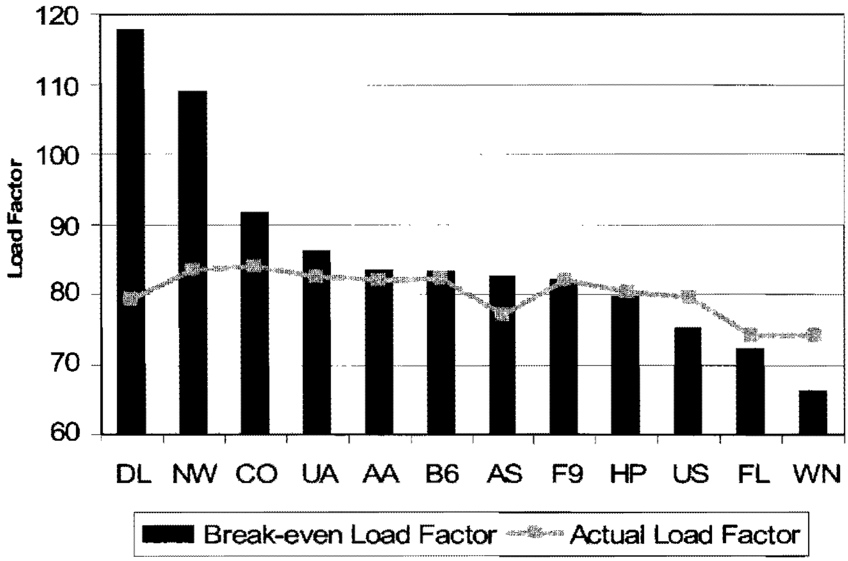

Figure 4.13 Comparison of actual and break-even load factors for US airlines, 2006

Source: Compiled by the authors using Back Aviation (2006) Form41 data.

The explicit formulas are as follows:

$$DOL = \frac{Q(P-V)}{Q(P-V)-FC}$$

$$DOL = \frac{S-VC}{S-VC-FC}$$

where: S is initial sales in dollars
EBIT is earnings before interest and taxes and is a measure of the profitability of the company.

The degree of operating leverage can vary considerably among companies in the same industry. Companies with a high degree of operating leverage are much more reactive to changes in output. This is partly a result of the company having sizeable fixed costs, which can either be leveraged effectively during times of increasing sales (because of decreasing average costs), or become a burden to the company during decreasing sales. As a concrete numerical example, if the degree of operating leverage is 2, this means that a 1 per cent increase in ticket sales will result in a 2 per cent increase in the airline's operating profit. In this case there would be increasing returns to scale. Unfortunately, the reverse is also true during downswings in the economy, with the company experiencing a greater operating loss. Therefore, companies with a high degree of operating leverage will experience greater volatility of operating profit than companies with smaller degrees of operating leverage.

Operating leverage can also be negative, which indicates that the company is counter-cyclical to the industry. For example, a company with a degree of operating leverage of –1 will experience a 100 per cent increase in operating profit growth when the sales figure declines by 100 per cent.

This scenario indicates decreasing returns to scale, probably due to the fact that the company has grown too quickly and is not effectively leveraging its fixed costs. Therefore, in order to improve its degree of operating leverage, the company should reduce its fixed costs while keeping the same level of output, or increase its output with the same fixed cost infrastructure. Furthermore, changes to the company's contribution margin will also dramatically affect the company's degree of operating leverage. An airline may choose between a high or low level of fixed assets. For example, an airline may substitute self-service check-in machines at different airports for check-in agents. If labor is not replaced with check-in machines, fixed costs are held lower, and variable costs are higher. With a lower level of operating leverage, the airline shows less growth in profits as sales rise, but faces less risk of loss as sales decline.

This overall understanding of an airline's cost structure is critical in a company's strategic planning phase. It provides the company with a comparison between sales and operating profit, and, depending on the value, can help guide strategic direction. It should be noted that operating leverage should not be confused with financial leverage, which deals with the amount of borrowed money that a company is using to finance its activities.

AIRLINE OPERATING LEVERAGE

As has been explained, operating leverage is the percentage of fixed costs in an airline's cost structure. Generally, the higher the operating leverage, the more a company's income is affected by fluctuation in sales volume. Table 4.6 provides the degree of operating leverage for US airlines based on 2005 data (except for US Airways proper). The contribution margin was found using operating revenues and expenses, while fixed costs were assumed to be contractual obligations for the airline in the next year. These contractual obligations include firm aircraft purchases, regional capacity purchase agreements, and long-term debt. A breakdown of these contractual obligations is a mandatory requirement in SEC 10K forms. Finally, the output used was enplaned passengers.

As Table 4.6 shows, the degree of operating leverage for US airlines varies considerably among carriers. The carriers can be grouped into three categories based on operating leverage: major network or "legacy" airlines, expanding low-cost carriers, and Southwest Airlines. All the legacy carriers (if you consider Alaska (AS) and America West HP legacy carriers) have positive degrees of operating leverage, but are less than 1. This indicates that all these carriers' profits will increase at a smaller rate than their increases in sales. In addition, all these carriers have a negative contribution margin, where operating expenses exceeded operating revenues. Within this range you have bankrupt Delta (DL) and Northwest (NW) at the high end, while stable carriers such as Alaska (AS) and Continental (CO) have degrees of operating leverage close to zero. Carriers with low degrees of operating leverage are relatively stable as changes in sales do not significantly impact on their operating margin; this means that Alaska and Continental will probably be relatively better off during downturns in the aviation industry and worse off during booms in the economy. This essentially follows recent trends as both these carriers did fairly well and have not applied for bankruptcy protection. It will be interesting to see how these carriers perform comparatively when the aviation economy begins to pick up again. Conversely, carriers such as Delta, Northwest, and United (all three who entered bankruptcy protection) are more highly leveraged. These carriers would experience greater swings in their operating margin than carriers such as Alaska and Continental.

The second group of carriers has negative degrees of operating leverage. This group includes expanding low-cost carriers such as JetBlue ((B6) and AirTran (F9). In both cases they had small contribution margins that were overcome by the long-term fixed costs for

Table 4.6 US airline operating leverage, 2005

	AA	AS	B6	CO	DL	FL	HP	NW	UA	US*	WN
Operating Revenues (millions)	20,712	2,975	1,701	11,208	16,191	1,450	5,457	12,286	17,379	7,073	7,584
Operating Expenses (millions)	20,805	2,983	1,653	11,247	18,192	1,437	5,599	13,205	17,598	7,421	6,764
Contractual Obligations (millions)	4,016	861	1,809	4,245	5,263	421	1,133	761	3,300	2,165	1,765
Passengers (millions)	79	17	15	45	78	17	22	57	67	42	78
Degree of Operating Leverage	0.647	0.126	-0.642	0.292	0.967	-1.057	0.735	0.986	0.816	0.870	1.028

Source: Compiled by the authors from SEC 10K filings.

Note: US Airways data is from 2004 as a result of the merger.

the company Therefore, these airlines will need to do a better job of translating fixed costs into sales and ultimately profits.

Southwest Airlines stands out by itself. With a degree of operating leverage of 1.028, it is the only US carrier with an operating leverage greater than 1. This means that for every increase/decrease in sales, the airlines operating profits will increase/decrease by 1.028 per cent. Southwest is certainly in an enviable position compared to other carriers as its contribution margin is significantly larger than its long-term obligations. Based on the degree of operating leverage, Southwest can expand and increase its sales to successfully increase its operating profit.

SUMMARY

This chapter explored the topic of costs. The various classifications of costs were explained and illustrated with hypothetical examples from the industry. The use of costs for managerial decision-making was then explained, as was the intuitive rationale behind these decision rules. In addition, various costs were compared across the spectrum of airlines. Finally, economies of scale, scope, and density were explained and discussed in detail.

REFERENCES

Aviation Partners Boeing (2006). Fuel savings. Retrieved on 9 November 2006 from http://www.aviationpartnersboeing.com/.

Boeing (2006). Commercial sites. Retrieved on 9 November 2006 from: http://www.boeing.com/commercial/overview/overview5.html.

Boyes, W. (2004). The New Managerial Economics (1st edn). Houghton Mifflin.

Harvard Business School (2000). *Dogfight Over Europe*, HBS9-700-15.

JetBlue Airways (2006). *Annual Report*. Available at: http://library.corporate-ir.net/library/13/131/131045/items/238155/JetBlue_2006AnnualReport.pdf.

Maurice, S. and Thomas, C. (2005). *Managerial Economics* (8th edn). New York: McGraw-Hill.

Ionides, N. (2004). Star Sees Benefits from Joint Fuel Purchasing. *Air Transport Intelligence News*. 4 April. Retrieved on 1 April 2007 from Air Transport Intelligence at: http://www.rati.com/iata/oapd.

Shannon, D. (2006). Airbus increases A380 breakeven level 55 per cent to 420 aircraft. *Air Transport Intelligence News*. 19 October. Retrieved on 5 March 2007 from Air Transport Intelligence at: http://www.rati.com/iata/oapd.

5

Aviation Infrastructure: Operations and Ownership

> So here we are. One of the worst summers on record for delays is headed for aviation's history books. Total delays are up 19 percent from where they were last summer ... Yet despite the progress we're making, our air traffic system is still not even close to what it needs to be—what the flying public demands it should be. The system's too old and it's not nimble enough for today's activity. I'll underscore that this is why we need to move to a system for the next generation, NextGen. This is the modernization step we need to take.
>
> Marion Blakey, FAA Administrator

One of the most unique features of the aviation industry is the unprecedented amount of regulatory and operational control that the industry is subject to. These controls are manifest in the form of the many government agencies that regulate and control everything from the direct ownership of airports, the control of aircraft in the air, and the certification of aircraft production on the ground. While virtually all industries are regulated in varying degrees, very few have their day-to-day operations under the direct control of a government agency. Such control presents many unique challenges to the aviation industry.

Using the theoretical constructs of supply and demand that were introduced in the previous chapters, this chapter analyzes the situation and discusses alternative arrangements that have been suggested and, in some countries, have actually been implemented. The chapter specifically focuses on the operational infrastructure of the industry, namely, air traffic control (ATC) and airports. We begin with a brief history of air traffic control, followed by an economic analysis of the existing system, and, finally, the prospects for reform. The last part of the chapter is devoted to an economic analysis of airport ownership and the likely outcomes when public or private ownership is considered.

This chapter covers the following topics:

- The air traffic control system,
- International problems in US air traffic control
- Air traffic control in a government corporation
- Political obstacles to reform
- Airport ownership and management, including:
  - Trends in airport privatization
  - Reasons for privatization

- Types of privatization
- Privatization in the US airport industry.

THE AIR TRAFFIC CONTROL SYSTEM

Early control of air traffic began in the 1920s and was mainly concerned with navigation rather than control per se. This rudimentary system involved the use of flags, lights, and bonfires to locate and identify airports and runways, and to communicate with pilots. In the early 1930s this system was replaced by a more formal set-up that consisted of a series of light towers. The system was called the Transcontinental Lighted Airway and, at one time, consisted of over 1,500 beacons and 18,000 miles of airways. Thus, the early technology forced aircraft to rely on point-to-point navigation over predetermined routes, rather than the more direct routing that the aircraft were capable of. Gradually, as radios became more technically advanced, they replaced the earlier systems. As traffic between major metropolitan areas began to increase, it became apparent that a more centralized system was needed not only to provide separation, but also to facilitate navigation.

Accordingly, in the mid-1930s an airline consortium established the first three centers to pool information on specific flights so as to provide better separation. These centers were taken over by the Bureau of Air Commerce within the Department of Commerce when it assumed responsibility for air traffic in the United States. Separation was accomplished mainly through flight scheduling over the already-established prescribed routes. Again, the existing technology was forcing the aircraft to fly from ground-based fix to ground-based fix over the predetermined routes. At the end of the 1930s Congress passed the Civil Aeronautics Act which transferred civil aviation responsibilities from the Department of Commerce to a newly created agency called the Civil Aeronautics Authority. In 1940 President Roosevelt separated the Authority into two agencies, the Civil Aeronautics Administration (CAA) and the Civil Aeronautics Board (CAB). The newly created CAA was tasked with certification, safety, airway development, and air traffic control (Kent, 1980). At about the same time the Civil Aeronautics Board was split off from the CAA and charged with the economic regulation of the transport industry.[1] The Federal Aviation Agency was then created by the Federal Aviation Act of 1958 and tasked with the old responsibilities of the CAA and the new responsibility of safety rule-making. Finally, in 1967 the Agency's name was changed to the Federal Aviation Administration, and it was placed under the Department of Transportation.

Throughout the 1970s the FAA installed new radars, computers, and radio communications to upgrade and enhance the air traffic control system. However, the problem with this development was the fact that it was created under the same premise as the older systems—that is, aircraft were expected to travel from ground-based navigation point to ground-based navigation point along the predetermined and pre-existing airway route structure. This had the effect of lining up the traffic in a linear fashion along the various routes. Controllers would then use various techniques to maintain predetermined separation standards. As one would expect, the flow of traffic would be metered by the slowest aircraft and/or the largest separation distances. This, of course, is exactly analogous to a ground-based highway system—it was often called the 'highway in the sky' system—and it ignores the capability of an aircraft to travel in three dimensions and

1 The effect of this regulation is covered in greater detail in later chapters.

directly through the airspace. Nonetheless, the system worked reasonably efficiently with the then existing volume of traffic and level of technology.

However, as traffic increased and technology advanced, it became increasingly evident that the system was becoming outdated and that there was an urgent need to modernize and update it.[2] Unfortunately, bureaucratic tendencies and various political considerations now took over, and meaningful reform and modernization became increasingly difficult, if not impossible.[3] Some expensive efforts were undertaken to improve the system, but none of these was very effective. It should be pointed out that one of the principal reasons why the system could function with the older equipment and procedures was the undeniable fact that it could operate reasonably well under good weather conditions. In good weather pilots could see and be seen, with the consequence that separation standards could be reduced, and, since good weather conditions are generally the norm in the United States, problems with the system and procedures were generally not evident. But, by the end of the 1990s, and with the advent of the twenty-first century, the volume of traffic had begun to overwhelm the system. The tragic events of 9/11 slowed this growth in volume, but air traffic had returned to, and exceeded, earlier levels by 2006. Hence Marion Blakey's comments at the beginning of this chapter.

One of the principal advantages of air travel is the speed with which an individual can arrive at his or her destination. Therefore, factors that contribute to delay in the system certainly reduce its attractiveness. As these factors increase in magnitude, they increasingly reduce the quantity of air traffic demanded. The economic effects of delay can be analyzed using the supply and demand models introduced in the earlier chapters. Consider Figure 5.1 below.

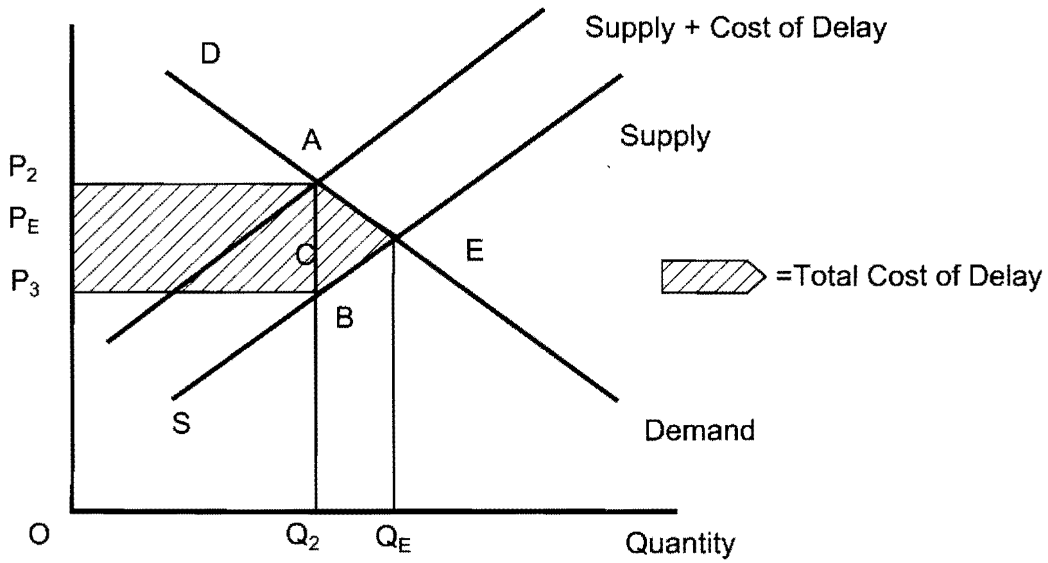

Figure 5. 1 The short-term economic effects of air traffic delay

2 Principal among these developments were precise methods of navigation and the ability to accurately locate the position of any aircraft in the sky.

3 Among these considerations were the location of facilities and the question of union job loss.

As explained in the earlier chapters, there is an equilibrium price and quantity supplied of air travel. This equilibrium is represented by Pe and Qe and is the market equilibrium based on supply and demand. As you may recall, any price above the equilibrium price will result in a surplus of air travel supplied—that is, too many empty seats. In this situation, competition between airlines will lower the price to fill the seats and return to the equilibrium position. Any price below the equilibrium will result in a shortage and consumers will bid up the price to obtain the seating. Now suppose that the air traffic control system imposes repeated and prolonged delays in the form of ground holds or airborne holding. This can come about if the system is not capable of handling the volume of traffic either because of separation standards or the inability of the human controllers to handle the volume of traffic. This situation can be thought of as an externally imposed cost to both the consumers and the producers of air travel. To the consumers it is an unanticipated delay that can be monetized as the cost of the consumer's time or, in the case of the business travelers, as forgone opportunities. To the producer the cost is also real, and can be measured, among other costs, in terms of higher crew wages, more fuel burn, and the loss of the ability to utilize the airplane for extra flights. Therefore, the delay that is imposed on the system can be thought of as a tax on both the consumers and suppliers of air traffic.

In Figure 5.1 these costs are shown as the straight line AB that joins the demand and supply curves. We can think of the costs as a parallel shift in the supply curve, so that the new supply curve intersects the demand curve at point A. However, since these are extra costs, the new equilibrium quantity and price will be defined by the intersection of the new supply curve with the old demand curve, or P2 and Q2. On the other hand, the actual price that the producers receive is determined from the old supply curve, net of the cost introduced by the externally imposed delay, or P3. We can also see that the price that the consumers must now pay is the new equilibrium price P2 which is clearly higher than the original equilibrium price. Therefore, the cost of the delay is shared between the consumers and the producers. The consumers bear part of the cost in the form of a new and higher ticket price, while the producers bear part of the cost in the form of a new and lower ticket price. The total amount of the cost is equal to the rectangle ABP3P2 plus the triangle ABE. And, as we can see from the diagram, the total cost is borne by both the consumers and the producers. For consumers, the costs are equal to the rectangle P2ACPe plus the triangle ACE. For producers, the costs are equal to rectangle PeCBP3 plus the triangle CEB.

The cost of delay is analogous to the impact of a conventional tax on air travel, though congestion costs are worse. Suppose that a tax, equal to congestion cost, had been imposed. In conventional demand and supply tax analysis, the entire rectangle ABP3P2 is the amount that the taxing authority receives from the tax, while the triangle ABE is the dead-weight loss that results from the imposition of the tax. This dead-weight loss can best be thought of as the transactions between buyers and sellers that do not take place because of the imposition of the tax—that is, the consumers who would have purchased tickets absent the presence of the tax and, likewise, the producers who would have sold them the tickets absent the tax. In the case of delay that is imposed by the air traffic control system on both consumers and producers, the situation is, in reality, worse than a tax, because the entire area—that is, the rectangle ABP2P3 plus the triangle ABE—is a dead-weight loss. In other words, there are no tax proceeds being transferred to government, consumers and producers suffer a loss of wealth (wasted time, wasted fuel, higher maintenance costs, and so on), but government gets nothing out of that. Instead, all of the costs are borne by the consumers and producers in the form of a dead-weight loss.

The question of who bears a larger amount of the costs is more complex, but can still be addressed using supply and demand analysis. As may be clear from Figure 5.1, the question of who bears a greater amount of the cost depends on the slopes of the demand and supply curves. As the supply curve becomes more and more inelastic—that is, as the slope gets higher and higher—it becomes increasingly difficult for the producers to shift the burden of the cost to the consumers. Figure 5.2 shows the situation in more detail. We can imagine a limiting situation where supply is perfectly inelastic at some given time. In this case, the supply is represented by a perfectly vertical line and is fixed regardless of price. Since supply is fixed by definition, it cannot shift when the extra cost is introduced. Instead, we can show the cost of delay as a downward shift in the demand curve. In this case, the ticket price remains the same for the consumers, and the entire cost of the delay is borne by the producers and is equal to the rectangle PeEAP2. We have presented the limiting case where the producers bear all of the costs since it is highly likely that, in the aviation industry, supply is relatively inelastic, at least in the short to intermediate term, in comparison with demand. Therefore, we can expect that the greater part of the cost of delay in the industry will be borne by the producers. And, as we shall see in the latter part of the chapter, this appears to be true when we observe the preferences of the market participants.

But all of this is not the end of the story. Unfortunately, there are other longer-term supply and demand economic effects that must be considered. On the supply side, if the cost of delay is persistent and longer-lasting, as it appears to be for the foreseeable future, then the value of the specialized resources presently in use in the aviation industry will be diminished accordingly. These resources include, among other inputs, the production of aircraft, the supply of spare parts, the manufacture of avionics, airport-related concessions, air travel-related accommodations, automobile rental concessions, income for pilots, flight attendants, mechanics, and a host of other related factors. As these specialized resources wear out, they will not be replaced at the same rate as previously, and this will further reduce output in the

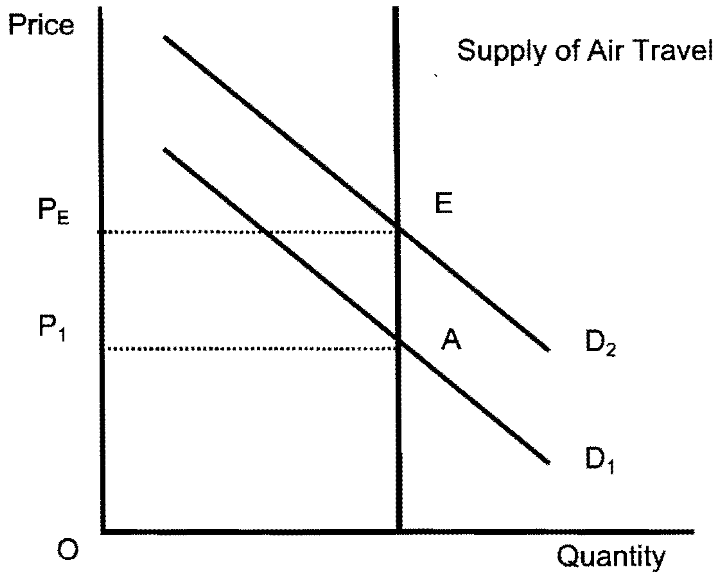

Figure 5.2 The cost of delay with inelastic supply

industry. This process is illustrated in Figure 5.3 where the initial supply decrease occurs, and then a further decrease takes place as these specialized resources wear out or exit the industry. As we can see from the figure, the long-term effect further lowers quantity and raises price.

On the demand side, we can expect a shift in demand away from air travel to alternative modes of travel where this is feasible. Although air travel is clearly the fastest mode of travel when considering relative speed, it is the total trip time that is of primary interest to the traveler. As things stand now, with delays and long security lines at some airports at critical times of the day, as well as persistent air traffic system delays, the inherent speed advantage of air travel is severely compromised. If this situation continues, then more and more consumers will opt for surface travel, and this will most likely be by private automobile. Since automobile traffic is inherently more risky than air travel, there will be a concurrent rise in accidents and fatalities.

This process is illustrated in Figure 5.4 with both an initial demand decrease and then a longer-term additional decrease. In other words, demand decreases immediately with delay since the quality and speed of air travel is reduced, but this initial reaction is multiplied over time, reducing demand further, as air travelers have more time to adjust their behavior, work out other travel arrangements, utilize modern communications to cut down on number of trips, and so on.[4] The demand decreases reinforce the decline in total air travel caused by the supply decreases; the industry is substantially smaller and less efficient than it would be without these delays. However, the impact on price becomes theoretically ambiguous because falling demand tends to reduce price whereas supply decreases tend to raise it.

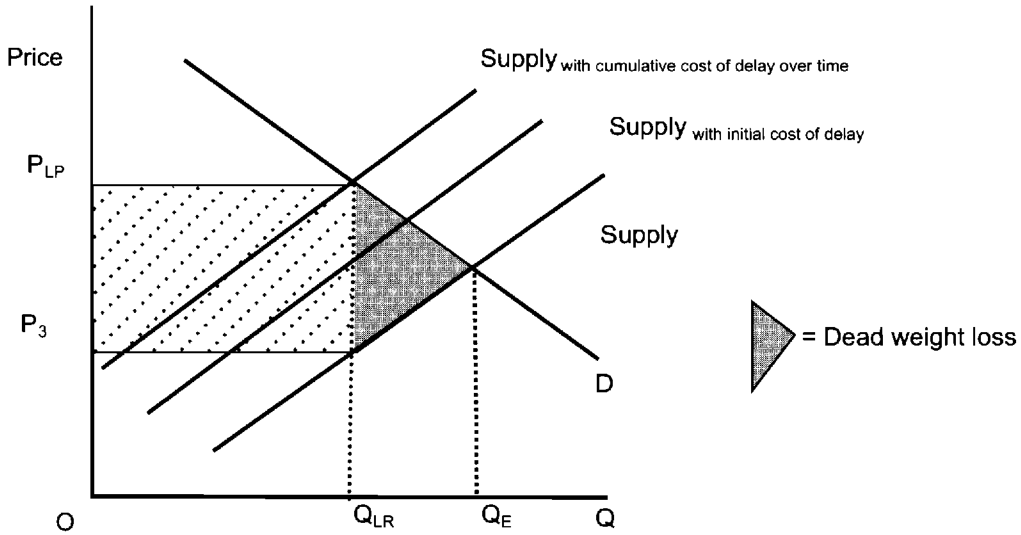

Figure 5.3 Costs of delay over time

4 In essence, this is merely a way of expressing the standard principle that demand (and supply) in the long run is more elastic than in the short run. For simplicity, we have omitted the long-run curves and focused on the additional shifts in the short-run curves.

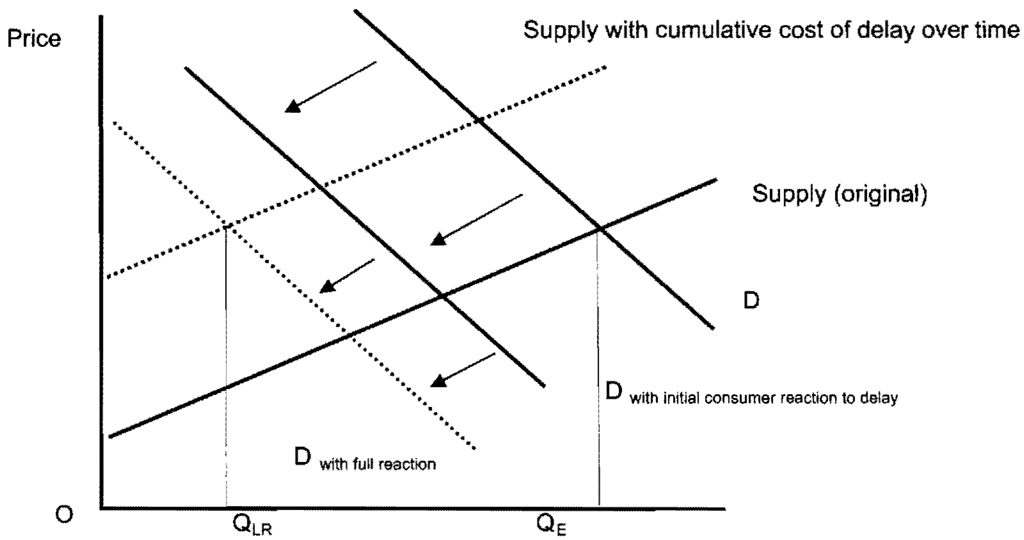

Figure 5.4 Market reaction to changes in supply

The preceding analysis used the theoretical tools of supply and demand to illuminate a very real and pressing problem in the aviation industry. The fact is that the final product of the industry is under the direct control of an outside government agency (the FAA) that does not have the same incentives or goals as the industry, especially those concerning profitability. This can and has led to large external costs that are increasingly imposed on the industry. Moreover, the FAA itself has come under increasing criticism as to whether it is the appropriate agency for operational air traffic control. The next sections discuss these criticisms in more detail.

INSTITUTIONAL PROBLEMS IN US AIR TRAFFIC CONTROL

A number of studies have concluded that US ATC management is inherently flawed and in need of major reform. These include:

- the Aviation Safety Commission in 1988
- the National Commission to Ensure a Strong Competitive Airline Industry in 1993
- the National Performance Review in 1993
- the Secretary of Transportation's Executive Oversight Group in 1994
- the National Civil Aviation Review Commission (Mineta Commission) in 1997.

The fact that ATC funding flows from an unpredictable revenue stream subject to the federal budget process is a commonly raised issue. Sometimes there also seems to be an inability, stemming in part from limitations in the civil service system, to attract and retain needed managers and engineers who are skilled at implementing complex technology projects.

The lack of both mission focus and clear accountability, because authority is shared by Congress and FAA management in a confusing bureaucracy, is at the root of much of the criticism. The future implementation of NextGen, for example, will be complicated by the fact that numerous outdated FAA operations around the United States will need to be closed. But Congressional Representatives are reluctant to vote in favor of shutting down any operation that "provides jobs in their district."

A possible solution to these sorts of problems, according to some economists, might be to move to a private, non-profit corporation, as Canada did in 1996 with its NavCanada Corporation. However, although it is commonly believed that this is not politically feasible, there is broad support for establishing a government corporation to address the concerns that have stymied previous attempts at major ATC reform.

ATC IN A GOVERNMENT CORPORATION

A key feature of a government corporation is non-political funding as user fees replace taxes and Congressional budgeting. The existence of an independent revenue stream allows access to private capital markets to fund modernization. In turn, the elimination of tax funding creates an exemption from government procurement rules that have previously tended to impede the acquisition of new technology. Likewise, independent funding allows exemption from civil service regulations that can make it difficult to attract, manage and maintain the appropriately skilled workforce.

New Zealand converted its ATC operation from a government division to a self-supporting government corporation in 1987. As of 2007, over 40 countries have implemented similar commercialization reforms, including Australia, France, Germany, Switzerland, the United Kingdom, the Benelux countries, and Scandinavia. Only a few of these are privatized in the sense of being outside of government; most are government corporations. All of these commercialized, self-supporting air navigation service providers (ANSPs) belong to the Civil Air Navigation Services Organization (CANSO), which has become a key participant in international aviation policy debates. All ANSPs are subject to safety regulation and some form of economic regulation because of their monopoly on ATC services.

In 2005 the Government Accountability Office (GAO) conducted a large-scale evaluation of the performance of commercialized ANSPs. The GAO collected extensive data from, and visited, five major ANSPs in Australia, Canada, Germany, New Zealand, and the United Kingdom (UK). It found that, since the commercialization of ATC, safety had either been unaffected or even improved. It also found that all five of the systems studied had taken significant steps to invest in new technology and equipment, and had taken meaningful steps to reduce operating costs. Similarly, a 2005 FAA study found that commercialized systems were more cost-effective in airspace with equivalent traffic density (Federal Aviation Administration, 2005). In short, commercialized ATC has become the norm for most of the industrialized world, and apparently has a solid record of improved efficiency with no decline, and even some improvements, in safety.

POLITICAL OBSTACLES TO REFORM

Most opposition to ATC commercialization centers on user fees. In general, the aviation industry has been particularly fearful of user-fee impact. However, it is very feasible, and probably politically necessary, to exempt piston aircraft flying under visual flight rules (VFR) from user fees. For those that sometimes fly instrument flight rules (IFR) reasonable accommodations can be made. In Canada, for example, they pay a modest annual fee based on aircraft weight. Moreover, the envisioned board of directors for a US government corporate ATC system would include representatives of general aviation, who would have to approve any changes in user fees.

Business jets would have to pay some user fees, but experience shows that these can be reasonable. Both Canada and Europe have experienced strong growth in business aviation after ATC commercialization (Poole, 2006).

One might expect public employees to be potentially strong opponents of commercialization. However, The National Air Traffic Controllers Association, the main FAA union, supported the Clinton Administration's proposal to divest ATC to a government corporation, structured along the lines discussed here. Moreover, commercialization could readily include 'no-layoff' guarantees for all current controllers and technicians.

Government reform often proceeds at a glacial pace—no matter how inefficient existing institutions are there are always interest groups which perceive a vested interest in maintaining the status quo. However, in the case of ATC perhaps the United States has reached a stage where the problems are so severe that the political log-jam blocking major reform may soon be broken.

Recently, the industry itself has mounted a strong lobbying campaign to change the situation. This effort has been led by the Air Transport Association and has involved a political campaign to influence Congress to change the air traffic control system. Basically, the aim is to replace the existing ground-based radar and voice communication system with a more technologically advanced system based on a much more accurate surveillance technology—namely, the Automatic Dependent Surveillance Broadcast or ADS-B. This system will also provide a more open standards-based architecture to replace the existing national airspace system software and will also include airborne collision avoidance and shared intent information within the cockpit. The present radio communication system is supposed to be replaced by a data link system that will allow direct exchange of messages between controllers and pilots. Finally, new navigation systems will allow direct routing and replace the current airway system.

The FAA expects flight operations to increase by over 50 per cent over the next decade, and this increases the cost of delay discussed above. Whether or not meaningful reform will actually take place is problematical, since it is difficult to believe that an organization with a set of incentives and goals that are fundamentally different from the aviation industry will be able or willing to implement any of these changes. The following quote is taken from the Airport Transport Association (ATA) Smart Skies initiative:

> Without dramatic change in the way our airspace is managed, congestion and resulting delays will be overwhelming for passengers, shippers, consumers and businesses. Failure to meet future airspace demand could cost the U.S. economy $40 billion annually by 2020.

www.smartskies.org

AIRPORT OWNERSHIP AND MANAGEMENT

This final part of the chapter analyzes and discusses the appropriate ownership of airports and the use of pricing mechanisms to allocate and improve the scarce resources at the airports. According to Airports Council International (ACI), in 2006, world airports handled a record 4.4 billion passengers—an increase of 4.8 per cent from the previous year. There are approximately 1,670 commercial airports serving more than 900 airlines worldwide. According to the FAA there are 546 commercial airports. Of these, 422 have more than 10,000 enplanements and are grouped as commercial service airports in the United States. Furthermore, there are 33 large hubs, with 464,486,847 enplanements; 35 medium hubs, with 115,177,169 enplanements; 68 small hubs, with 50,202,980 enplanements; and 410 non-hub airports. At present, almost all the airports in the United States are under some type of government control. One of the byproducts of this control is typically a pricing system (landing fees) that is fixed over the entire day. Consider Figure 5.5 below. If the airport authority sets a price below equilibrium for a particularly advantageous period of the day, then it is easy to see that the quantity demanded will exceed the available supply by the amount of Qe – Q1, and some other rationing device must be found. This typically takes the form of delay for some or all the aircraft and/or, in some cases, a rationing of the available landing times (slots) for the airport in question. Both these solutions have inherent efficiency problems when considered from an economic point of view. And, in the case of slot controls, the airlines that are awarded the slots benefit from an economic rent.[5] Therefore, the issue becomes the appropriate ownership of airports. Many economists would maintain that private ownership of the airports would provide a better set of incentives for the long-term viability of the industry. The next few sections discuss this question in more detail.

Trends in Airport Privatization

Increasingly, many countries around the world are rethinking the appropriate role for government in the operation and ownership structure of aviation infrastructure. Many

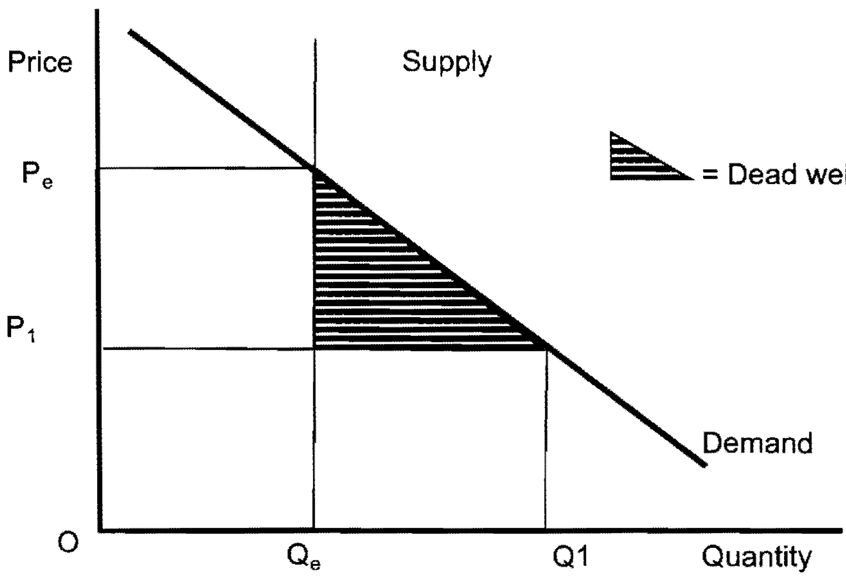

Figure 5.5 Demand for airport services

5 An economic rent, also known as Ricardian theory of rent, is the payment over and above that needed to keep a factor of production in its current use.

countries have privatized airports, concluding that the private sector can run airports more efficiently just as they run airlines more efficiently or produce better aircraft than would a state-owned manufacturer.[6] In 1987 the British government led the way when it completely privatized its seven major airports, selling the British Airport Authority (BAA) to the public for $2.5 billion. BAA owns and runs seven airports in the UK, including Heathrow, Gatwick, Stansted, Southampton, Glasgow, Edinburgh and Luton. In 2006 the company was purchased by the Spanish firm Ferrovial with a market value of $18 billion. Following the apparent success of BAA, many countries have followed suit. BAA, led by the Ferrovial Group, also operates Indianapolis International Airport and has retail-management agreements at Baltimore,Washington, Boston Logan, and Pittsburgh international airports. As of early 2007 over 100 major airports have been privatized worldwide, including those at Belfast, Brussels, Budapest, Copenhagen, Düsseldorf, Frankfurt, Hamburg, Rome, South Africa, Argentina, Chile, Colombia, Mexico, Auckland, Brisbane, Melbourne, Sydney, and many others, with Hong Kong and Tokyo in the works.[7] Roughly a dozen global airport companies are in the business of running airports. The financial company, Macquarie, has created a privatized airports mutual fund for global investors (see Poole, 2007). Figure 5.6 summarizes the number of airport privatization transactions from 2001 to 2005. From the figure, it can be determined that Budapest was the most expensive while Hamburg (HAM) was the cheapest. This was determined by the value of the airport divided by EBITDA (earnings before interest, tax, deprecation, and amortization). The figure also illustrates that airports of all economic levels are privatizing. The average cost was about $15 million, but, as you can determine, there is a large degree of variation between the different airports.

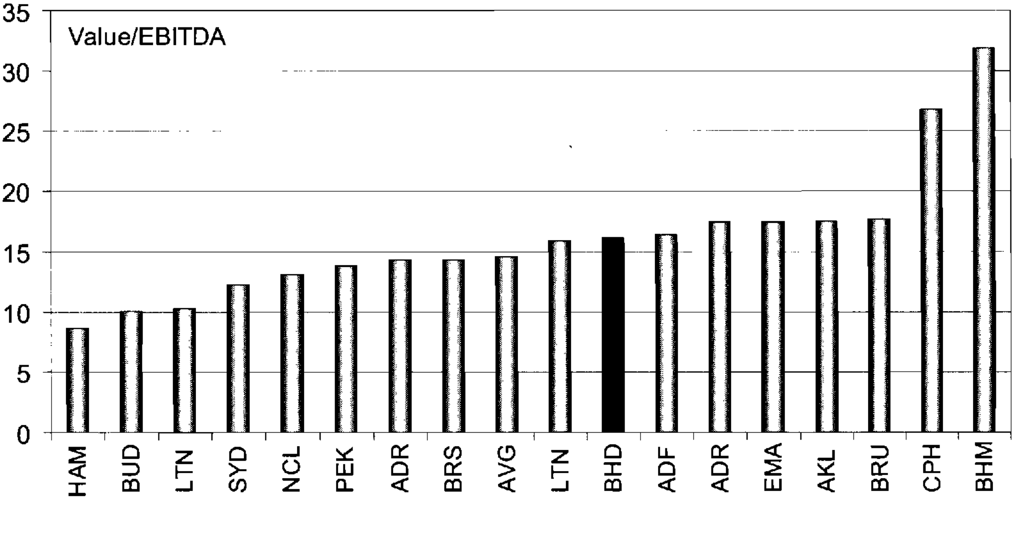

Figure 5.6 Recent examples of airport privatizations, 2000–2006

Note: See the appendix for the airport codes.

6 See Chapter 2 for a discussion of why efficiency tends to be enhanced by reducing government's role.

7 Since its privatization in 1987, BAA acquired partial ownership stakes in the several other airports including Budapest and Sydney.

Reasons for Privatization

Proponents of selling or long-term leasing of airports to the private sector perceive three principal advantages:

- greater efficiency of operations, particularly in developing the non-aviation side of the airport.
- capital infusion, which opens up non-traditional sources of capital
- the conversion of a private airport into a tax-paying corporate entity.

In addition to being generally more efficient, private companies can readily raise funds for needed airport projects without becoming entangled in the political problems and delays that often plague government airports looking for grants to expand or renovate. Moreover, these companies can engage in equity financing, while government is only able to issue debt. Of course, there may be potential problems with privatization (see, for example, Vasigh and Gorjidooz, 2006). Some worry that, even with continued government regulation, a private company may not be motivated to properly maintain infrastructure. However, two decades of experience seem to indicate few problems in this regard. Most economists would probably argue that profit motives provide strong incentives for the proper maintenance of airport infrastructure. Since consumers tend to be hypersensitive to safety concerns in air travel, any hint of corner-cutting in this regard is likely to depress demand and sink profits. Also, private airports, particularly in the United States, are more likely to be held accountable by liability laws, since it is generally easier to sue a private party for damages than to sue the government.[8] In essence, a private airport appears to have the same regulatory incentive for safety as a government airport, since there is no change in these regulations with privatization, plus the added incentives of stricter legal liability and of a profit motive. Indeed, these added safety incentives for private companies may explain why one seems sometimes more likely to encounter dangerous infrastructure failures in government levees, bridges, and roads than on private roads, parking lots and other structures.

Another fear that is sometimes voiced is that private airports might occasionally go bankrupt. Although this is a distinct possibility, bankruptcy is more a financial disaster for stockholders than an operational problem for air travelers. Just as airlines have (routinely, unfortunately) continued to operate normally in bankruptcy so, too, would viable airports. Furthermore, if management is markedly at fault, then a bankruptcy judge might well eject it. Thus, bankruptcy provides a new channel—one not available in the case of mismanaged government airports—for eliminating poor management.

Another common objection to airport privatization is that airports have monopoly power and, if private, will raise landing fees to extremely high levels. However, it is not certain that this is good either for the industry or air travelers. Recall from the economic analysis in the introduction to this section that artificially low landing fees can result in high costs from dead-weight losses that are worse than the cost of a higher landing fee. Very low prices for airline operations result in excess fuel burn, maintenance costs, and time wasted for both crew and travelers. Moreover, any business that is unable to choose

---

8 Of course, overly harsh liability laws can be an impediment to economic efficiency and consumer well-being, and many would argue that the US could benefit from tort reform that would render liability laws less harsh. However, tort reform is not a necessary precondition to airport privatization—many companies are clearly interested in buying US airports under the existing tort system, whatever its faults.

its price is fundamentally hampered. In the case of airports, major benefits from airport expansion, for instance, may not be affordable if price is held artificially low.

Also, the danger of excessively high monopoly pricing may not be as great as it first appears. In the end, airports can only charge airlines higher fees if airlines are able to pass on those higher costs to customers. It is probably safe to say that the leisure traveler will not bear such costs in that they will probably either travel by other modes of transportation or fly via more distant airports if prices go up substantially at their home airport. Business travelers are, perhaps, more likely to pay higher prices, but even in their case, the monopolist airport must consider competition from other transportation modes, secondary airports, modern telecommunications, corporate jets, and the developing "air taxi" competition from the very light jets that are able to operate out of smaller airports. Given these considerations, some experimentation with free pricing of landing fees may be warranted.

Although the theoretical grounds for supporting airport privatizations seem solid to most economists, actual experience may be the most persuasive evidence. The very fact that so many different governments are abandoning control of their airports is a strong statement in favor of privatizing. Following some initial skepticism there now seems strong support for the argument that divestiture can enhance the efficiency of airport operations (Truitt and Michael, 1996). Vasigh and Hamzaee (1998) emphasize the benefits of privatization of airports in Western Europe, Latin America, and Asia that should inspire officials in search of new economic opportunities in transforming airports from publicly run to private businesses (see Vasigh, Yoo, and Owens, 2004).

Types of Privatization

The techniques used to privatize airports vary in terms of the scope of responsibility and, in some cases, the degree of ownership transferred to the private sector. A traditional privatization tool involves the contracting of selected services (restaurants, parking, security services, cargo, baggage-handling, and/or fueling services) to the private sector, while the government retains overall operating responsibility for the airport. More comprehensively, under the contract management approach, the government transfers all responsibility for all airport operations and implementation of strategy to the private sector, while retaining the ownership and investment responsibilities. Several US airports are currently operated under management contracts. These include Westchester County airport, Albany County airport in New York, and Burbank airport, which is owned jointly by the cities of Burbank, Glendale, and Pasadena, California. The Burbank airport has been managed by Lockheed Air Terminal, Inc. since 1978. Lockheed receives a fixed management fee, plus expenses for the services it provides for the airport, and the airport authority is responsible for capital improvements. Burbank airport, which ranks 59th in size among US airports (as measured by annual passenger enplanements) is often held up as a viable model of public–private partnerships in airport operations (Ashford and Moore, 1992).

A long-term lease approach allows the government to legally (and politically) retain ownership but to transfer investment, operational, and managerial responsibilities to a private tenant, with the lease long enough to motivate them to more or less behave as an owner. This method may be used to allow the financing of the construction of the airport or associated project by the private sector, which must then relinquish control at the end of the lease term. Several examples of this type of public–private partnership already exist in the United States, including airports in Atlantic City and Morristown, New Jersey. Perhaps the best-known example of such a lease arrangement is Teterboro airport, in New Jersey. The lease to operate Teterboro was established in 1970, when Pan American World Airways (now known as Johnson Controls World Services) negotiated a lease with Teterboro's owner, the Port Authority of New York and New Jersey. Johnson Controls recognized that general aviation activity was causing congestion and threatening its commercial operations in New York. The company believed that it could relieve some congestion if general aviation aircraft could be lured away from the city. With this in mind, Johnson Controls secured a 30-year lease to operate Teterboro airport.

Finally, using a full divestiture/sale of shares, the government transfers full (or partial) ownership to the private sector—the airport is sold off, as in the BAA case. It should be noted that even where the airport is sold, the government still retains substantial regulatory control in many areas, including safety and, usually, the regulation of landing fees.

The United States is conspicuously absent from the long list of countries that have sold or long-term leased major airports to private companies. However, private companies have had substantial success in operating US government-owned airports under contract. For example, in October 1995, BAA took over the management of Indianapolis International Airport promising to raise non-airline revenues by $32 million within the ten-year period of the contract. The contract was renegotiated in 1998 and extended until 2008. Between 1995 and 1999, costs per passenger were reduced from $6.70 to $3.70 and despite a moderate passenger annual growth rate of 3.5 per cent, non-airline revenue per passenger more than doubled by 2003 (Vasigh and Haririan, 2003).

Privatization in the US Airport Industry[9]

In many countries around the world, except the United States, governments own and operate aviation infrastructure including gates, assigning them dynamically to airlines as needed (common-use gates). Under this system airlines paid landing fees and space rentals, at pre-set rates, based on how much of the facilities they used. This same model has continued under privatization, meaning very little change for the airlines. In contrast, the typical US approach has been one in which the anchor-tenant airlines signed long-term lease agreements with charges based on "residual cost"—that is, the cost of operating the airport would first be covered by revenue sources other than the airlines, such as parking, concessions, and so on. Whatever costs were not covered by these revenues would then be assessed on the airlines via landing fees and space rentals.

In effect, these signatory airlines became joint owners of the airport. In good years the airports would take in more of the non-aviation revenues and the airlines would enjoy lower fees. Of course, in bad years with fewer passengers and therefore lower airport revenues, the airlines would have to pay higher fees. Thus, bad years for the airlines became even worse.

One might expect the airlines to refuse an arrangement that made their profits even more intensely cyclical, but two considerations made the arrangement worthwhile for them. First, the federal funding system for airport expansion was often so cumbersome that joint airport ownership/funding by airlines was the only viable way for the timely

9 This section draws heavily from Poole (2007) and from Vasigh, Yoo, and Owens (2004).

expansion of the airport. Second, the airlines gained the ability to veto airport spending on terminal expansion that they saw as wasteful and might lead to more fees and charges. Thus, US legacy carriers tend to have a vested interest in maintaining the status quo, opposing airport (and, perhaps, as discussed below, ATC) privatization because it might indirectly open their markets to more competition.

As airport privatization gained momentum in the 1990s there was a call for legislation to eliminate federal regulations hostile toward private airports. A key problem was that the regulations were interpreted to imply that any local government that sold an airport would then have to repay all previous federal airport grants and, of course, that airport would be ineligible for any future grants. In an attempt to address these problems the 1996 Airport Privatization Pilot Program was passed; cities whose airports were accepted for the pilot program would not have to repay previous grants.

However, political pressure, largely from the airlines, resulted in a provision that rendered the program essentially useless. In order to make use of lease or sale proceeds, a city has to get the approval of 65 per cent of the airlines serving the airport. Otherwise, all profits must be reinvested in the airport, making the whole exercise, from the viewpoint of the airport owners, not worth the trouble.

The only airport actually privatized under the 1996 law—Stewart Airport in Newburgh, NY—failed to get the necessary airline approval, meaning that the city cannot gain from the profit. So, the sale or lease of US airports is likely to remain politically unfeasible for the foreseeable future unless airline opposition weakens. Contractual privatization, in part or in whole, remains the only viable alternative.

SUMMARY

This chapter has introduced some more practical aviation applications of the theoretical constructs of supply and demand that were introduced in the previous chapters. The chapter analyzed the costs of delay imposed by the regulatory agency as an external tax that has been levied on the industry through the failure of the regulatory agency to modernize and use effective technologies. These costs are large and likely to grow larger over the foreseeable future. The chapter also used supply and demand analysis to introduce the concept of airport privatization and the use of the market price system to allocate scarce resources at the airport.

APPENDIX: AIRPORT CODES

Airport Company	CODE
Bristol	BRS
Hamburg	HAM
Aeroporti di Roma	ADR
Beijing, Capital	PEK
Auckland	AKL
Birmingham	BHM

Newcastle	NCL
London, Luton	LTN
East Midlands	EMA
Aeroporti di Roma	ADR
Sydney	SYD
Aeroporto di Firenze	ADF
Belfast City Airport	BHD
Brussels	BRU
London Luton	LTN
Budapest	BUD
Copenhagen	CPH

REFERENCES

Ashford, N. and Moore, C. (1992). *Airport Finance*. Kluwer Academic Publishers.

Federal Aviation Administration (2005). *International Terminal Air Traffic Control Benchmark Pilot Study*, February.

Kent, R. (1980). *Safe, Separated, and Soaring: A History of Federal Civil Aviation Policy, 1961–1972*. Washington: DOT/FAA.

Poole, R. (2007). Will Midway Lease Re-Start U.S. Airport Privatization? *Public Works Financing*, Reason Foundation, January.

Poole, Robert (2007). Bizjet Sales Cut by Commercialization? *ATC Reform News*, No. 31, Reason Foundation, January. Available at: www.reason.org/atcreform31.shtml. Accessed 11 February 2007.

Truitt, L. and Michael, J. (1996). Airport privatization: Full divestiture and its alternatives. *Policy Studies Journal*, 24, pp. 100–24.

Vasigh, B. and Gorjidooz, J. (2006). Productivity Analysis of Public and Private Airports: A Causal Investigation. *Journal of Air Transportation*, 11(3), pp. 142–62.

Vasigh, B. and Hamzaee, R. (1998). A Comparative Analysis of Economic Performance of U.S. Commercial Airports. *Journal of Air Transport Management*, 4(4), pp. 209–16.

Vasigh, B. and Haririan, M. (2003). An Empirical Investigation of Financial and Operational Efficiency of Private versus Public Airport. *Journal of Air Transportation*, 8(1), pp. 91–110.

Vasigh, B., Yoo, K. and Owens, J. (2004). A Price Forecasting Model for Predicting Value of Commercial Airports: A Case of Three Korean Airports. *International Journal of Transport Management*, 1(4), pp. 225–36.

6

International Economics and Aviation

... when I cock my ear toward Mexico, I still hear that "giant sucking sound" of American jobs headed south of the Rio Grande.

Ross Perot, the Reform Party's 1996 presidential candidate

Although other opinions are addressed, the thrust of this chapter is to explain why economists largely agree that international trade, for the most part, benefits the economy. Aviation is generally not an exception to that rule, though it has some special complications. Anti-trade arguments are discussed and, with some slight exception, demonstrated to be mainly fallacious. This chapter lays the foundation for the next, which details international aviation agreements and industry alliances. Allowing international competition in aviation and elsewhere produces a net gain to the economy rather than a destructive "concession," as politicians often suggest. This chapter will cover the following topics:

- International economics and trade, including:
  - Arguments for and against free trade
  - National security concerns
  - Aircraft manufacturing and government subsidies
  - Trade deficit and surplus
- The logic of production possibility, absolute advantage, comparative advantage, and free trade, including:
  - Absolute advantage
  - Comparative advantage
  - International trade policy in air travel—optimality versus political realities
- Trade protections and trade barriers
- Foreign currency and exchange rates, including:
  - Exchange rate quotes
  - Exchange rate regimes: fixed, floating, pegged, and the gold standard.

INTERNATIONAL ECONOMICS AND TRADE

Economists in general agree on most of the key policy issues relating to international economics. However, as is often the case in economics, this consensus has failed to break through many popular misconceptions, and therefore public policy in international trade deviates substantially from the ideal. All countries, regardless of their size, depend to some degree on other economies, and are affected by trade and transactions outside their borders. The globalization of the international economy has occurred in almost in every country and has led to the development of many regional trade agreements. International trade allows countries to take advantage of other countries' resources through the theory comparative advantage (explained later in the chapter). It has been successfully argued and empirically demonstrated that overall world production increases through trade and partnership

In 2006 the volume of world economy and trade increased by 8 per cent. Countries such as Germany, the United States and China have emerged as important exporters in world merchandise trade (Table 6.1). The United States is the world's largest market for exporting countries. In 2005 it imported more than $1732 billion worth of merchandise (Table 6.2). Japan exported more than $594 billion worth of products to the rest of the world—4.8 per cent of total exports—and imported $514 billion during the same period. Among the ten leading exporters, the five most dynamic economies are Germany (9.3 per cent), United States (8.7 per cent), China (7.3 per cent), Japan (5.7 per cent) and France (4.4 per cent). The five leading importers comprise the same countries as the group of the top five leading exporters, except that the United Kingdom replaces France in fifth place with 4.7 per cent (World Trade Organization, 2006).

Globally, the Middle East was the region with the highest exports in 2005 (see Figure 6.1). This means that all the Middle Eastern countries collectively beat all the other regions

Table 6.1 Top ten exporting countries in world merchandise trade, 2005

Rank	Exporters	Value (billion dollars)	Percentage share	Annual percentage change
1	Germany	969.9	9.3	7
2	United States	904.4	8.7	10
3	China	762.0	7.3	28
4	Japan	594.9	5.7	5
5	France	460.2	4.4	2
6	Netherlands	402.4	3.9	13
7	United Kingdom	382.8	3.7	10
8	Italy	367.2	3.5	4
9	Canada	359.4	3.4	14
10	Belgium	334.3	3.2	9

Source: World Trade Organization (2006).

Table 6.2 Top ten importing countries in world merchandise trade, 2005

Rank	Importers	Value (billion dollars)	Percentage share	Annual percentage change
1	United States	1732.4	16.1	14
2	Germany	773.8	7.2	8
3	China	660.0	6.1	18
4	Japan	514.9	4.8	13
5	United Kingdom	510.2	4.7	8
6	France	497.9	4.6	6
7	Italy	379.8	3.5	7
8	Netherlands	359.1	3.3	12
9	Canada	319.7	3	15
10	Belgium	318.7	3	12

Source: World Trade Organization (2006).

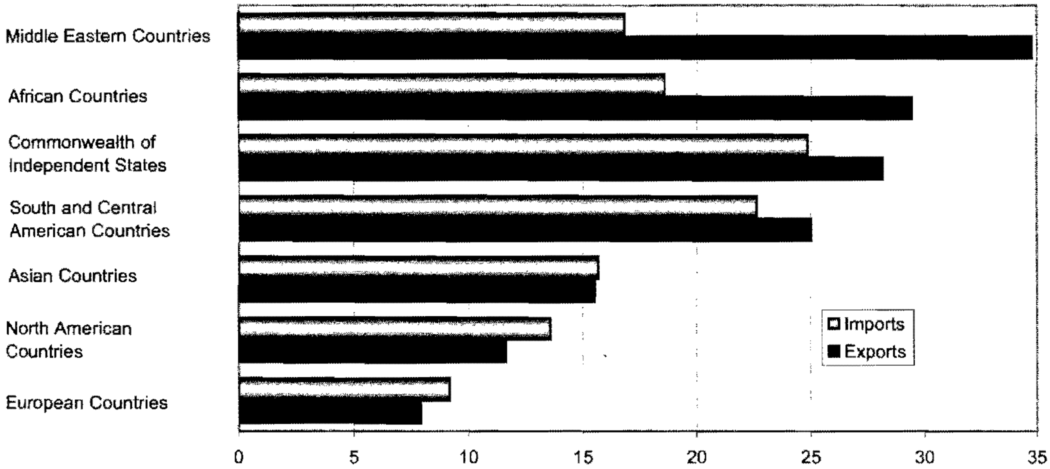

Figure 6.1 World merchandise trade distributed among regions

Source: World Trade Organization (2006).

Note: The results do not reflect any one country's specific trade, but all countries in their respective regions collectively.

even though no individual Middle Eastern country made it to the top ten exporting countries list.

Similarly, the Commonwealth of Independent States (CIS) was the region with the highest imports in 2005.[1] None of its member countries made it to the top ten importing countries list. Table 6.3 lists the top US trading partners for 2006. (US Census Bureau: Foreign Trade Division, 2007).

1 The CIS was created in December 1991, and at present includes Azerbaijan, Armenia, Belarus, Georgia, Kazakhstan, Kyrgyzstan, Moldova, Russia, Tajikistan, Turkmenistan, Uzbekistan, and Ukraine.

Table 6.3 Top ten US trading partners, 2006

Overall Trade Rank	Country	Exports (Year-to-Date)	Imports (Year-to-Date)	Total, All Trade	Percentage of Total Trade
1	Canada	230.6	303.4	534	18.50
2	China	55.2	287.8	343	11.90
3	Mexico	134.2	198.3	332.4	11.50
4	Japan	59.6	148.1	207.7	7.20
5	Germany	41.3	89.1	130.4	4.50
6	United Kingdom	45.4	53.4	98.8	3.40
7	South Korea	32.5	45.8	78.3	2.70
8	France	24.2	37.1	61.4	2.10
9	Taiwan	23	38.2	61.2	2.10
10	Malaysia	12.6	36.5	49.1	1.70
	Total, All Countries	**1,037.30**	**1,855.40**	**2,892.70**	**100.00**

Source: US Census Bureau: Foreign Trade Division (2007).

ARGUMENTS FOR AND AGAINST FREE TRADE

> Trade is the natural enemy of all violent passions. Trade loves moderation, delights in compromise, and is most careful to avoid anger. It is patient, supple, and insinuating, only resorting to extreme measures in cases of absolute necessity. Trade makes men independent of one another and gives them a high idea of their personal importance: it leads them to want to manage their own affairs and teaches them to succeed therein. Hence it makes them inclined to liberty but disinclined to revolution.
>
> Alexis de Tocqueville

There are many arguments for and against free trade. Today, most countries around the world impose some form of trade restrictions such as tariffs, taxes, and subsidies. The arguments most often heard against free trade are:

- It is important to keep jobs in the country.
- Imports should be limited to keep money in the country.
- Free trade may be a threat to national security.
- Other nations don't treat their workers fairly.
- A nation may become too specialized and dependent on other nations.

Economists generally argue that the world at large will benefit from free trade and that trade liberalization can promote development. Free traders claim that trade protection always harms all the trading partners. The main political objections to free trade center on fears that cheaper imports will destroy jobs and reduce income. Labor unions and management may oppose free trade if they believe that the competition from free trade will cause them to lose jobs and/or bankrupt the firm or industry. In the simple two- good case illustrated later in the chapter it is easy to see that, for the economy as a whole, this fear has no basis in reality. Although free trade may indeed render some jobs obsolete (or lost), other jobs are created. The new jobs are more productive, so average wages and overall income rises. Again, this is easy to see in the two-good case (where the country lost jobs making one product, but gained more productive jobs making another).

Proponents of free trade cite the following advantages:

- It reduces the price of every item sold in the market.
- It increases the supply of products in other markets and results in lower prices for those products.
- It encourages other nations to trade more freely with their trading partners, which helps the global economy.
- It increases the number and variety of products for consumers to choose.
- It is a driving force behind a high standard of living.

Even though free trade increases average wealth, some individuals will probably be made worse off. This, after all, is true of any advance in technology. For example, discovering a cure for cancer would destroy some jobs. In an advancing economy most people will be able to adapt to new job opportunities with little or no serious problems, although in extreme cases some individuals may require substantial assistance to survive and adjust to technological advances. But with trade, as with improving technology, overall wealth is increasing in society, so such assistance is more readily affordable. Low trade barriers inherently create hundreds of billions of dollars more in benefits than they impose in costs (Pugel, 2007).

The same comparative advantage principle holds when there is trade with many goods, but the impact is not so obvious. Suppose, for instance, a country dramatically increases purchases of imported automobiles. Everyone will be able to see the resulting layoffs in the domestic automobile sector—the downside of this international trade. However, the new jobs created by this trade are widely dispersed and not at all obvious. When consumers are able to buy cheaper automobiles some will then be able to spend more on computers, some will enjoy more air travel, some will buy new clothes, eat out at restaurants more, buy a nicer home, and so on. Thus, cheap automobile imports allow us to produce more of other products and create additional job to enable this new production. However, the job gains will be widely dispersed throughout many, seemingly unrelated, industries. The individual who obtains an airline job made possible by the availability of cheaper automobiles in the economy is unlikely to see the connection. In other words, even though the benefits of trade far outweigh the costs, free trade is often controversial because the costs are concentrated, visible to even the least discerning citizens, and therefore easy to politically exploit. Hufbauer and Elliott (1994) state that the average protected job costs $170,000. So, rather than using trade barriers to "protect jobs," it would be cheaper for consumers to allow free trade and then compensate all workers who lose jobs to more efficient foreign competition.

National Security Concerns

In a very few special cases it may be true that national security justifies a particular trade barrier. It would not be prudent, for example, to allow aircraft from hostile nations to have "open skies" access to our airspace. But, more typically, free international trade enhances national security by raising income and promoting friendly relations. This is particularly true for aviation, since free trade in air travel inherently "makes the world smaller," promoting more economic integration and social interaction between nations. When economies are strongly integrated, many of each country's citizens have a large stake in other countries, and therefore a strong vested interest in avoiding the massive destruction and attendant loss of wealth associated with war. This is why wars between major trading partners are relatively unusual in world history. Indeed, it was the desire to promote peaceful interaction, even more so than economic development, which initially motivated European leaders to form the European Union (van den Berg, 2004).

Some try to argue that allowing foreign airlines, even those from friendly nations who are staunch allies, to compete freely in domestic markets somehow inherently jeopardizes national security. They insist that airlines must be domestically owned just in case the government needs to use civilian aircraft in some emergency to, say, move military troops within the country. Most economists are very skeptical of such claims. All governments reserve the right, for example, to confiscate private property (hopefully, with just compensation being paid at the appropriate time) in emergencies, regardless of who owns the needed property. Instances where government would need to confiscate or commandeer civilian aircraft are rare, possibly even non-existent in many cases, but if the situation ever arises, the government has the power to take what it needs.

In the United States, the government does have a contract in place with some airlines to provide troop transport if ever needed (US Air Force: Air Mobility Command, 2006). Given the relative efficiency of US airlines and the availability of military transport aircraft, even if the US domestic air travel market were completely open to foreign competition it is unlikely in the near term that foreign carriers would so dominate the market that there would not be enough US aircraft to move troops. If, in the distant future, foreign carriers begin to achieve such dominance there seems to be no reason why the United States could not arrange the same sort of contract with a foreign carrier; it could even be required as a condition of the carriers being allowed to compete in the US market. Other countries could do likewise. National security does not provide a reasonable argument against free international trade and competition in air travel.

Aircraft Manufacturing and Governmental Subsidies

It is commonly alleged that Airbus and Boeing are unfairly subsidized by their respective governments, allowing them to charge less than full price for their aircraft at a disadvantage to other manufacturers. For example, in the case of Boeing, partisans argue that the US government should impose some sort of tariff on Airbus to offset the subsidy and afford Boeing a level playing field. Airbus disputes this analysis and its conclusion, but economists point out that, even if the subsidy exists as alleged, it would be harmful to the United States to impose a tariff. In 2005 the governments of both the United States and Europe agreed to stop subsidizing Boeing and Airbus for a short period of time while they try to resolve a decades-old dispute over billions in subsidies to the aircraft-makers.

If people in the United States can buy aircraft more cheaply, they will have more resources to produce and consume additional aircraft and whatever else they might desire—total wealth clearly increases. The impact would generally be the same if European taxpayers mailed checks directly to US consumers rather than giving money to Airbus. If there is a loser in this case, then it would be the European taxpayers who see some of their wealth transferred to the United States

Boeing does suffer some lost sales from the subsidy, but Boeing's loss is less than the combined gain to airlines, air travelers, and the economy in general. Notice that the general effects of trade are not significantly affected by the existence or lack of a subsidy. Boeing loses less than the general economy gains if Airbus provides a better aircraft for the price. Regardless of whether that better price comes from a subsidy, hard work, luck, better technology, or whatever, the impact is the same.

Trade Deficit and Surplus

It is useful to point out first that the term "trade deficit" is completely arbitrary and might just as easily be called a "trade surplus." If imports are greater than exports that is, more goods and services are flowing into a country than flowing out—a deficit exists. Conversely, if the difference in the value of a nation's exports over imports is positive, the country enjoys a trade surplus. To illustrate this point, suppose someone came to your house and brought a number of products to you. Although they took some of your goods in return, the value of what they gave you was greater than the value of what they took. Would it not be more natural to say you enjoy a surplus of trade with this individual, rather than following convention which would term this situation a trade deficit?

But is it harmful when money "leaves the economy?" First, remember that the world's leading currencies are no longer backed by gold or any other real assets. Dollars, yen, and euros are pieces of paper, valuable in exchange, but very inexpensive to print in virtually limitless quantities. Of course, printing too much currency results in the devaluation of that currency—in other words, inflation—but printing currency simply to replace that which leaves the country will not be inflationary at all (as long as foreigners just hold the dollars). So, suppose that the United States were to experience the ultimate trade deficit—foreigners acquired dollars and simply collected them, refusing to buy anything from the Unites States. If the government did nothing to offset the effect of dollars pouring out of the United States, then the country would experience deflation—pervasive falling prices. So, to keep the value of the currency stable, the government needs to create new currency and put it into circulation to replace the currency leaving the country.

To put money into the system the authorities (the Federal Reserve System in the United States) buy existing US Treasury debt. Because interest is paid on this debt, the government is now paying interest to itself; in effect this debt is retired and no longer a burden to taxpayers. Thus, the more money that leaves the country the better![2] The United States is able to trade currency, cheap pieces of paper, for real goods plus retire substantial portions of debt. For this reason the United States, and all governments, generally encourage other countries to use their currency. Another way of looking at this is to state the obvious fact that the principal place where dollars can be spent is in the United States. Obviously, this will increase the demand for US goods and services or, as is more likely to be the case,

2 Recall the original assumption that foreigners prefer to hold the dollars.

investment in productive resources in the United States. There are currently more dollars circulating outside the United States than inside—something the United States welcomes (Elwell, 2004).

But what if those outside dollars are suddenly returned and spent in the United States? Although this would not be catastrophic, it would be somewhat costly. Consumption would have to fall as foreigners traded their dollars for goods and services—in other words, rather than having pieces of paper leave and goods and services flow in, the United States would now see the reverse. To prevent the returning currency causing inflation, the United States would have to take some currency out of circulation. Some of that US Treasury debt would be resold, and taxpayers would now be charged to make the outside interest payments to the buyers of the US Treasury debt. The currency gained from reselling that debt would be held out of circulation.

Of course, foreigners willingly hold dollars because they believe that it is in their interest to do so, so they are very unlikely to suddenly change their minds en masse and flood the United States with returning currency. But it is comforting to know that, if such a thing should somehow happen, the negative impact is not at all overwhelming.

As indicated in Figure 6.2, the United States has generally been experiencing a perpetual, growing trade deficit since about 1980. A few trade surpluses have occurred, but only when the economy has weakened in recessions. Although some currency has flowed out of the country, the trade deficit is basically balanced by a surplus of capital inflows. The capital account surplus (capital inflows minus outflows), also shown in Figure 6.2, is essentially the mirror image of the trade deficit. In other words, foreigners who wish to invest in the United States have largely outbid foreigners who wish to consume US goods. When the United States "sends dollars out of the country," they largely "come back" in the form of capital flows—for instance, foreigners directly build factories in the United States

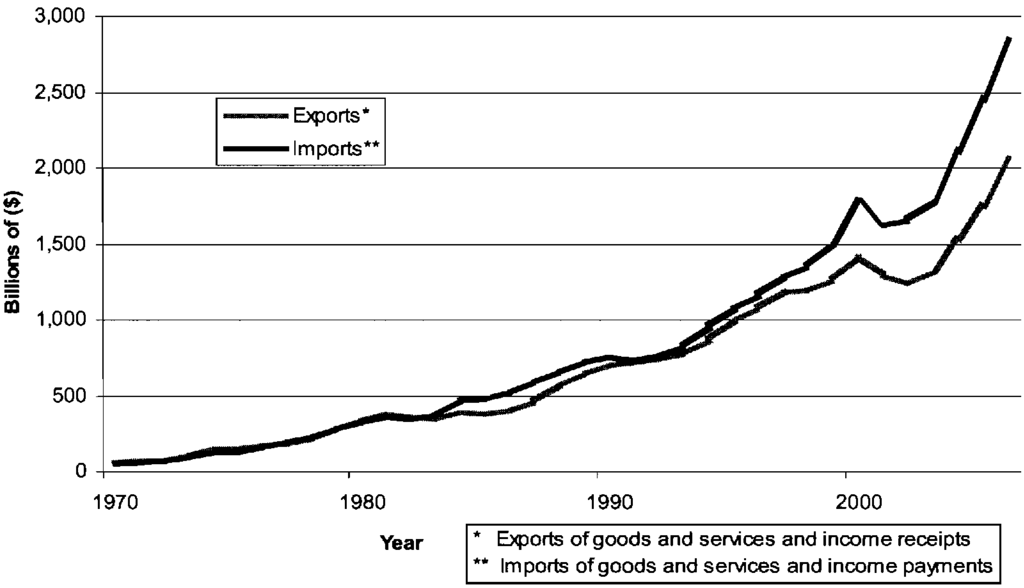

Figure 6.2 Imports and exports in the United States, 1970–2006

Source: US Department of Commerce: Bureau of Economic Analysis (2007).

or buy stocks and bonds from US companies which then expand in their home country (US Department of Commerce: Bureau of Economic Analysis, 2007). The United States is seen as an ideal place to invest: international investors are attracted to the world's largest economy, political stability, and a substantial degree of economic freedom. As long as this holds true, the United States is likely to continue to see capital surpluses/trade deficits for the foreseeable future.[3]

THE LOGIC OF PRODUCTION POSSIBILITY, ABSOLUTE ADVANTAGE, COMPARATIVE ADVANTAGE, AND FREE TRADE

Whenever the structure of production is addressed within an economy of any scale, it is helpful to understand the concept of a production possibility curve (PPC). A PPC shows all different combinations of output that an economy can produce by using all the available resources. The curve is plotted along a two-dimensional axis in which the y and the x axes each signify a quantity of a good, making it a "two-good" analysis. The curve itself is seen as a frontier outside of which production of the two goods is impossible. Any combination of outputs within the curve would prove to be inefficient, and maximum efficiency lies only along the curve. Therefore, any point within the production possibility curve represents inefficiency and any point outside the production possibility curve represents something unattainable, given available technology and other resources. The production possibility curve illustrated in Figure 6.3 shows the trade-off in production between, for example, "Food and Clothing" and any two combinations of these two products, such as A, B, C; alternatively, any others could be chosen.

For example, at point A, a quantity of 600 units of food can be produced along with 300 units of clothing, and at point B 200 units of food can be produced along with 700 units of clothing. Hence, the opportunity cost to create 400 more units of food than at point B can be said to be 400 units of clothing. Point M represents a level of production that can't be achieved with current levels of resources and technology; therefore it lies outside the production possibilities curve. Point N, lying inside the curve, represents a level of inefficient production.

The slope at any point on the curve describes the marginal rate of substitution or how much of one good must be sacrificed to produce one more unit of the other. For example, if the slope at point C is 0.75, then that is equal to the opportunity cost at point C: to produce one more unit of clothing, 0.75 units of food will be taken out of the production schedule, or, inversely, to produce one more unit of food, 1.33 units of clothing will not be produced.

3 Even though the so-called trade deficit is not harmful, one can occasionally observe negative reactions to it in financial markets. This sometimes happen because investors fear that rising trade deficits/capital surpluses may eventually trigger trade barriers that will harm the economy. Also, under certain circumstances, a rising trade deficit/capital surplus can be an early indicator of currency depreciation, which can also spook investors.

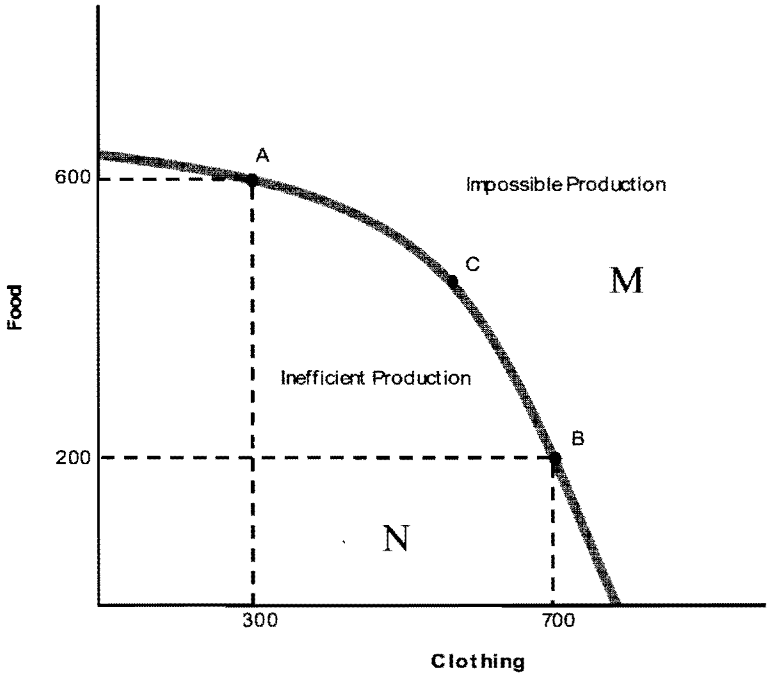

Figure 6.3 Production possibility curve

Absolute Advantage

A country is said to have an absolute advantage over another in the production of a good if it can produce the good with less resources. If a country can produce more output per unit of productive resources than its trading partner, then that country is said to have an absolute advantage in the terms of trade. Through specialization, different countries can produce and export goods where they have a natural or acquired absolute advantage and import those goods they don't specialize in. Table 6.4 shows an absolute advantage situation for the United States with respect to China in aircraft and China with respect to the United States in automobile production. Assume that, as shown in Table 6.4, China can produce 100 aircraft or 100,000 automobiles per one unit of productive resources, while the United States can produce 150 aircraft or 80,000 automobiles per unit of productive resources. Clearly, the United States has an absolute advantage in aircraft production whereas China has an absolute advantage in car production:

China: 1 aircraft = 1000 automobiles (100,000/100)
United States: 1 aircraft = 500 automobiles (80,000/160)

Using this information, we can calculate the opportunity cost of producing each product in each country. To produce one aircraft means that China must forego 1,000 automobiles, giving an opportunity cost of 1,000 automobiles. Furthermore, to produce one aircraft, the United States must forego 500 automobiles, giving an opportunity cost of 500 automobiles.

Table 6.4 Production possibilities

	Aircraft per unit of input	Automobiles per unit of input
China	100	100,000
US	160	80,000

Table 6.5 Opportunity cost

China	Opportunity cost
cost of producing 1 aircraft	1,000 automobiles
cost of producing 1 automobile	0.001 aircraft
US	Opportunity cost
cost of producing 1 aircraft	500 automobiles
cost of producing 1 automobile	0.002 aircraft

With an absolute advantage a country can charge a lower price than a competing trading country since more of the good with the absolute advantage can be produced with fewer resources. In the absence of free trade, in China one unit of aircraft will exchange for 1,000 automobiles, and in the United States one unit of aircraft will exchange for 500 automobiles. With the introduction of free trade, both the United States and China can gain benefits. If the United States can get more than 500 cars per unit of aircraft, then it will be better off than it would have been without trade. On the other hand, China has been giving up 1,000 cars to get one aircraft so if it can get an aircraft for less than 1,000 automobiles, then it will be better off. This means that the terms of trade should fall somewhere between 500 and 1,000 automobiles per unit of aircraft. In terms of the production possibilities curve introduced above, trade has effectively moved the curve to the right, thereby making available more of both goods.

Clearly, this simple example shows that there are significant gains to trade when each country has an absolute advantage in the production of one of the goods. However, what happens when one of the countries has an absolute advantage in the production of both goods? This situation is discussed in the next section.

Comparative Advantage

In 1817 David Ricardo outlined the theory of comparative advantage which shows how the gains from trade can still come about even if one country has an absolute advantage in the production of both goods. A country has a comparative advantage in the product with the lowest opportunity cost of production, so China should specialize in automobiles and the United States should specialize in aircraft. To demonstrate this, let us alter the numbers in the example we presented above to provide a simple proof of the benefits of international trade when one country as an absolute advantage in both goods. Suppose now that the United States is a superior producer of both automobiles and aircraft.

Table 6.6 shows that the United States has an absolute advantage in both products. Nonetheless, Ricardo argued that both countries can still benefit from trade. In the United States the opportunity cost of 1,000 cars is one aircraft. This includes the opportunity cost of production, meaning that, if labor, energy, material and other resources are reallocated away from automobile production to aircraft, then US firms can produce another aircraft but will lose 1,000 automobiles in its output. Likewise, if we reverse the reallocation and devote more resources to making automobiles, then US firms can produce another automobile, but will have to sacrifice 1/1000th of an aircraft. Meanwhile, in China the opportunity cost of an aircraft is 1,600 automobiles, with the opportunity cost of an automobile, equal to 1/1600th of an aircraft. Thus, even though the United States has an absolute advantage in the production of both automobiles and aircraft, it has a comparative advantage only in the production of aircraft, since each aircraft costs 1,000 automobiles compared to a cost of 1,600 automobiles in China. On the other hand, China's firms have a comparative advantage in the production of automobiles since their cost is only 1/1600th of an aircraft compared to 1/1000th of an aircraft opportunity cost in the United States.

If unrestricted trade is now permitted, then firms in each country will naturally shift production into the product with the comparative advantage. Given the numbers in Table 6.7, it is clear that the terms of trade will fall somewhere between 1,000 and 1,600 automobiles for each aircraft. Thus, US firms will produce aircraft and "convert" each aircraft into (say) 1,300 automobiles through trade. Note that, without trade, the Unites States could gain only 1,000 cars for each aircraft given up. By trading aircraft for automobiles the United States is able to acquire more automobiles and US citizens can afford to consume both more cars and more aircraft.

China's wealth also increases since it can now trade 1,300 automobiles for an aircraft rather than having to give up 1,600 automobiles for each aircraft it directly produces (see Table 6.8). This simple mathematical proof confirms common sense—when a society produces goods at the lowest possible cost of resources, it is possible to produce more. Allowing free international trade unambiguously increases overall wealth.

Likewise, it follows that international trade barriers—such as tariffs (special taxes on imports), import quotas, discriminatory regulation, or outright import bans—reduce wealth (see Figure 6.4). In a 2004 study, economists estimated that trade barriers reduced world GDP by about $500 billion, about $30 billion in the United States alone (Bradford and Lawrence, 2004).

Table 6.6 Comparative advantage

	Aircraft per unit of input	Automobiles per unit of input
China	50	80,000
US	100	100,000

Table 6.7 Production costs without trade

US	China
1 aircraft costs 1,000 automobiles	1 aircraft costs 1,600 cars
1 car costs 1/1000th of an aircraft	1 automobile costs 1/1600th of an aircraft

Table 6.8 Production costs with trade

US	China
1 automobile costs 1/1300thh of an aircraft (as opposed to 1/1000th without trade)	1aircraft costs 1,300 automobiles (as opposed to 1,600 without trade)

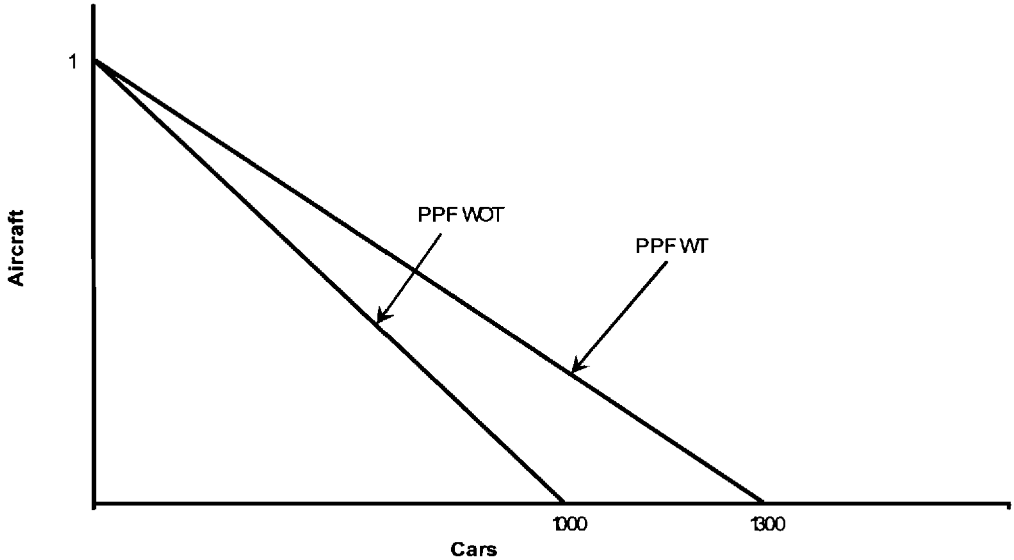

Figure 6.4 Production possibility frontier (with and without trade)

International Trade Policy in Air Travel—Optimality versus Political Realities

Economists generally agree that free trade is the best policy, in aviation and almost everything else. Ideally, foreign airlines from friendly nations should be allowed to freely compete; indeed, there is no reason for government policy to favor domestic airlines over foreign airlines. Implementation of this policy would maximize competition and efficiency in air travel—prices would be lower, there would be more variety of service, and consumers would have more choices. Overall wealth would increase because air travel would be more efficient in its use of resources, and the improved transportation system would help virtually all industries to be more efficient. A more efficient air travel industry has impacts analogous to a more efficient road system—it allows firms to expand into more output markets, gather resources from more input markets, and, as appropriate, take more advantage of economies of scale in production.

We mentioned earlier that it has been estimated that establishing free international trade in 2004 would have increased annual world income by about a half trillion dollars. Formal estimates for the impact of free competition in air travel alone are not available, but income would certainly increase by many billions.

Given the huge benefits of free trade in air travel, why have politicians in most nations failed to implement it? Think back to the public-choice principles discussed in Chapter 2. Most citizens are "rationally ignorant," and are unaware of the benefits of free international trade; there is no strong consumer movement clamoring for free trade in air travel or anything else. However, all firms, including airlines, like to avoid increased competition as much as possible. Consequently, governments have a tendency to act in the interests of airlines rather than in the interests of the nation as a whole. In other words, absent a well-informed public, politicians tend to give into the special interests of the domestic airline industry.

This is reflected in the typical language of trade politics. Whenever foreign firms are allowed to compete in a market—that is, reduce prices and increase the importing nation's wealth—politicians refer to this as a "trade concession," something they appear to reluctantly agree to in exchange for rights for domestic firms to compete in the other country. The entire attitude of politicians is that increasing import competition is an awful result that must be tolerated in order to negotiate export rights for domestic firms. This, understandably, is the perspective of the domestic firms these politicians seek to please: access to more markets is welcomed, more competition and lower prices are not welcomed, and even though national wealth increases, the wealth of companies facing more competition usually does not. But it is ironic that politicians who increase national well-being by reducing trade barriers feel compelled to cloak these good steps by calling them concessions.

The optimal policy favored by economists would be for a country, of course, to unilaterally open its own market to foreign competition. We would want to see increased competition and lower prices as soon as possible, with no negotiations necessary.

It is useful to summarize the likely results of such a policy in the US market. If the United States allowed cabotage—foreign carriers handling domestic traffic within the country—the immediate results, though beneficial, are unlikely to be spectacular. US airlines have been deregulated for a long time and are quite efficient by world standards—the market is not an easy one to make a profit in. Thus, foreign airlines would not be anxiously pouring into this market. Some, Virgin America Inc. for instance, would enter in a major way. Others might enter a few markets they already have some link to. For example, foreign carriers are already flying some blind routes where they are currently prohibited from handling domestic traffic. In other words, an airline might already fly, for example, from London to New York to Los Angeles, but is currently prohibited from picking up passengers in New York and dropping them off in Los Angeles. With the prohibition lifted they could freely market that segment.

From the perspective of US domestic carriers' self-interest, the disadvantage of increased competition would be at least partially offset by the injection of foreign capital. In abolishing the laws prohibiting cabotage the United States would inherently also be abolishing the laws prohibiting foreign controlling investment in US airlines. Some cash-strapped carriers would welcome some sort of partnership, perhaps even a formal merger, with a wealthier foreign airline. Likewise, the best strategy for a foreign airline looking to break into the US market might be to team up with a US partner. Current airline alliances achieve only a very limited amount of this sort of cooperation.

Since this would result in lower prices and greater efficiency there would be more airline passengers and therefore higher demand for labor in the airline industry; consequently, employment in the US airline industry would rise. Some of the new jobs might go to foreign workers brought in by foreign-based airlines, but net airline employment for US workers is still likely to rise since airlines, like most service industries, prefer to hire locally in order to promote better customer relations.[4]

The impact on average wages in the airlines is more complex. In competitive labor markets, rising labor demand would normally bid up the wages. However, much of the US airline industry is dominated by unusually powerful labor unions. Essentially these unions band workers together to bargain as a labor monopoly and thereby raise wages above competitive levels. Since increased competition tends to erode union monopoly power, the effect from cabotage could theoretically reduce average wages. In other words, rising labor demand tends to raise wages, while increasing competition tends to reduce union power and union wages, so the overall wage effects of cabotage are not immediately clear.

It is beneficial to review the US experience since airline deregulation occurred in 1978. In this case, the effect of increased labor demand swamped the effect of eroded union monopoly power, and wages generally rose after deregulation. Even after major union concessions following the 9/11 terrorist attacks, US airline employees were compensated at a level almost twice the average for all US industries (Ben-Yosef, 2005, p. 251).

Since the move from regulation to deregulation probably impacted on airline competition far more dramatically than cabotage would, it seems reasonable to conclude that airline wages would not be driven down following the opening of the US market. Of course, the effects on the US economy overall are unambiguously beneficial, though, as mentioned earlier, these effects are unlikely to be dramatic (at least in the short run) given the relatively efficient state of the US airline industry and the likelihood of only moderate initial new entry.

The same general impact, naturally, would occur in any country that opened its airline market, though the effects would often be more intense. In a number of cases it is quite possible that efficient foreign carriers would drive prices so low that flag carriers could not survive. To most economists, this loss would not be at all tragic, since the gains for the broader economy deriving from a more efficient air transport system would easily exceed the sentimental regret at the loss of an inefficient flag carrier—there is no more reason to insist that air travel be supplied internally than there is to insist on locally grown pickles.

But, does "national pride" justify preserving an inefficient flag carrier? One response is to note that, if the market is opened, the flag carrier is free to market itself to travelers on the basis of national pride—if consumers feel it is important to support the flag carrier they can do so. Naturally there are likely to be limits on how much of premium consumers would be willing to pay—a flag carrier that is vastly less efficient than the competition is probably doomed. But, if so, this clearly implies that people don't value "national pride" that much; in some sense, this argument in defense of the flag carrier is inherently invalid if consumers won't freely support it.

It may also be possible to compromise with, and overcome, the politics of protectionism in this sort of case by requiring foreign carriers to exclusively employ native-born employees, and maybe use aircraft painted in the home nation's colors with appropriate insignia. This would preserve much of the feel of having a flag carrier while still enjoying at least some of the benefits of open competition.

4 Although most economists would prefer to avoid added regulations, if employment politics are an obstacle to establishing cabotage then government regulation could stipulate a certain level of "native employment."

It is ironic that government interference has rendered air travel—an industry that should naturally be more global than most—far less global in its operations than virtually any other large industry. The pace of government reform is mostly very slow, but the direction of the trend is at least somewhat encouraging. Deregulation of air travel within the European Union has been impressive. The United States and Europe seem to be slowly moving toward establishing a 'Trans-Atlantic Common Aviation Area' (European Commission: Directorate General, Energy and Transport, 2004). Perhaps, open skies in lieu of cabotage will eventually become routine and allow a true flourishing of the air travel industry; one day, it may truly be a very small world. Foreign ownership of US airlines is another restriction that prevents a full open market for airlines in the US airline industry: US law limits the amount of foreign ownership in its domestic airlines to a maximum of 49 per cent, with a maximum of 25 per cent control. Nonetheless, many other countries protect their domestic markets in a similar fashion.

TRADE PROTECTIONS AND TRADE BARRIERS

> Underlying most arguments against the free market is a lack of belief in freedom itself.
>
> Milton Friedman

Trade barriers are attempts by the government to regulate or restrict international trade. They all work on the same common principle of imposing an additional cost on the imported good that will result in an increased price for that good. A country can protect domestic industry by imposing a trade *tariff*, a *quota* or a trade *subsidy*.

A tariff, a common trade barrier, is an additional tax on imported goods. Some tariffs are intended to protect local industries from cheaper foreign goods, while others are an attempt by government to generate revenue from the imported good. Tariffs are the easiest trade barrier to impose and can successfully reduce free trade (Husted and Melvin, 2007). Assume that Figure 6.5 shows supply (S) and demand (D) for automobiles in the United States In the absence of trade restriction, automobiles will be imported at the prevailing market price of (P_{FT}). In this example:

OQ_4 = total consumption in US
OQ_1 = total domestic production
Q_1Q_4 = total import
$OQ_4 = OQ_1 + Q_1Q_4$

P_{NT} and Q_{NT} are the unit price and quantity available in conditions where there is no international trade and hence no international competition. P_{FT} and Q_4 are the unit price and quantity available in conditions where there is free international trade and no international trade barriers.

Suppose the United States imposes a tariff, equal to M per aircraft, on imported automobiles. As a result of this, the price of automobiles will rise by the amount of the tariff, to P_T. In the absence of any retaliation by other countries, the increase in price of the automobile reduces consumption and increases domestic production. This would change the values to:

OQ_3 = total consumption in the United States

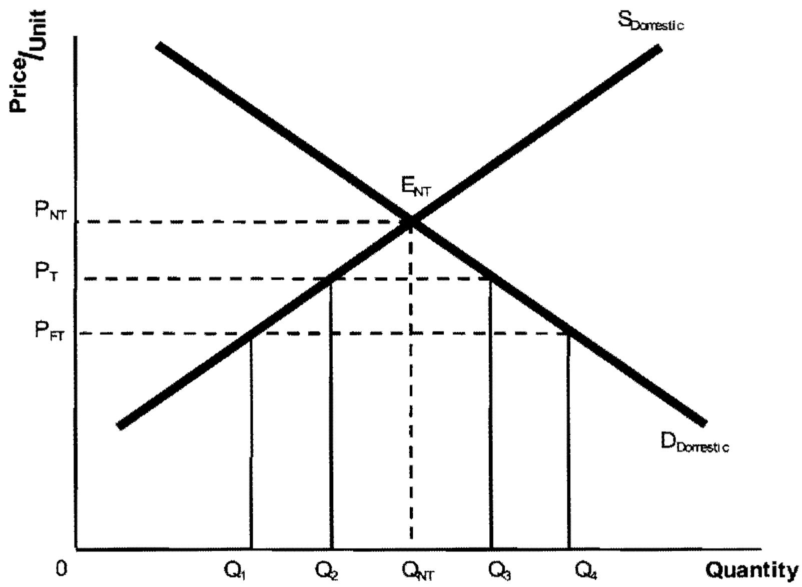

Figure 6.5 The effects of international trade barriers to a domestic market

OQ_2 = total domestic production
Q_2Q_3 = total import
$OQ_3 = OQ_2 + Q_2Q_3$

Another case is a subsidy—that is, when a government gives financial aid to a company to help produce or purchase a product. Subsidies have been used to aid a new or failing industry if that industry cannot generate enough revenue to maintain itself and it is in the interest of the general public (Carbaugh, 2007). Amtrak, for example, has received government subsidies in the form of loans to keep its service running. In 2007 the Bush Administration agreed to pay $900 million of Amtrak's estimated $3.1 billion budget.

One of the more controversial subsidies is that which allegedly exists between Airbus and its associated countries. As the World Trade Organization (WTO) began a hearing on the plane-maker's dispute with Boeing in March, 2007, the United States accused Airbus of benefiting from more than $100 billion in illegal aid. (US Delegation, 2004). According to WTO rules, government protection of domestic industries is illegal if another member can prove that the subsidy has harmed one of its companies or industries. The European Union (EU) has brought the same charges about the US aid given to Boeing, claiming that the plane-maker's 250-seat 787, to be introduced in 2008, has benefited from $5 billion in assistance (EU Delegation, 2004). The EU claim rests on the fact that Boeing has received government contracts for other aircraft, and the EU asserts that this amounts to government funding. However, viewing these contracts as subsidies is a novel way of looking at things, since the US government expects a product in return, whereas the governments contributing to Airbus simply expect a return on their original monetary investment.

The objective of an import quota, like a tariff, is to protect domestic producers from outside competition. Import quotas limit the quantity of various commodities that can be imported into a country during a specified period of time. In Figure 6.6 import quotas

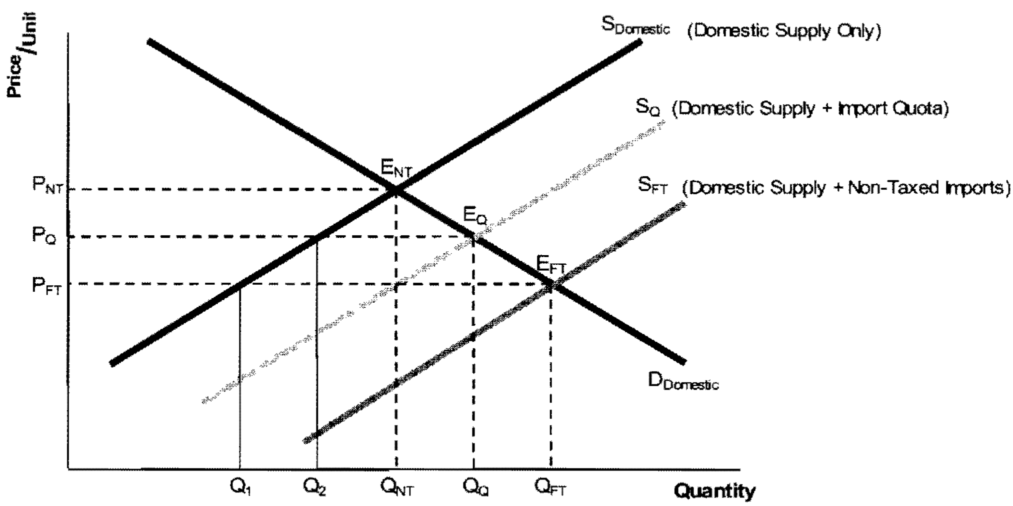

Figure 6.6 The effects of trade quotas on a domestic market

on automobiles will induce domestic manufacturers to expand production from Q_1 to Q_2, Q_Q being the total supply including both domestic production and the quota. US dairy products are a good example of this as they are subject to annual import quotas administered by the Department of Agriculture.

Another, more controversial trade practice is so-called "dumping." Dumping supposedly occurs when a manufacturer in one country exports its product and sells it in another country at an unreasonably low price, usually claimed to be below the cost of production. It is then usually alleged that workers in the second country may become unemployed because of this "unfair competition." However, free-market advocates see dumping, if it actually occurs, as beneficial to consumers since it obviously lowers the price of the product in question. According to WTO regulations a government may act against dumping when material injury to the domestic industry has occurred. One can argue that this directive is intentionally vague in order to allow each government to decide on its own to what extent dumping will be permissible.

FOREIGN CURRENCY AND EXCHANGE RATES

International trade—imports and exports—require foreign currency in order to complete the transactions. When goods and services are bought in a country, they are bought using that country's currency. To obtain foreign currency, one must trade in one's own local currency via the currency exchange rate. The exchange rate, or the price of one nation's currency in terms of another nation's, is a central concept in international finance. Exchange rates are influenced by a wide range of different factors, and the importance of each differs from country to country. For example, one factor affecting the exchange rate between currencies is the rate of inflation. As a general rule, the currency from countries with lower inflation rates rises in value, while the currency from countries with higher inflation rates falls in value. Consequently, the products from countries with low inflation rates become more attractive than the products from countries with higher inflation rates.

Another factor affecting the exchange rate between currencies is interest rates. Everything being equal, a higher interest rate on US securities (compared to, say, Canadian securities) would make investment in US securities more attractive. Therefore, increases in the US interest rate increase the flow of Canadian dollars into US securities, and decrease the outflow of American dollars into Canadian securities. This increased flow of funds into the US economy would increase the value of the US dollar and decrease the value of the Canadian dollar. Hence, the ratio of US dollar to Canadian dollar, as it is represented in the foreign exchange market, would decrease.

Finally, a country's balances of payments with the rest of the world influence that country's exchange rate. Demand for foreign currency arises from the import of foreign merchandise, payment for foreign services, or from the redemption of foreign capital obligations. The supply of foreign currency, on the other hand, comes from the export of goods and services, or from an inflow of foreign capital.

Exchange Rate Quotes

An exchange rate quotation is the value of one currency in terms of another. For example, a quotation of 1.06 CAD/USD signifies that 1.06 Canadian dollars will be needed to acquire 1 US dollar. In this quotation, the price currency is CAD (Canadian dollars) and the unit or base currency is USD (US dollars).

When the base currency is the home currency (the United States in this example), it is known as a direct quotation. Using direct quotation, the exchange rate decreases when the home currency appreciates (strengthens) and increases when the home currency depreciates.

Exports generally increase the exchange rate of the domestic currency (appreciation) since they cause more foreign currency to come into the domestic country. For example, if Japan Airlines bought large numbers of wide-body aircraft from the US-based Boeing Company, it would have to convert its Japanese yen to US dollars to complete the transaction. This would result in the US banks receiving the Japanese yen and exchanging them for the requisite amount of US paper dollars at the current exchange rate. This increased demand for US dollars acts like any increase in demand—that is, the price of dollars in yen is increased. This increase in price is illustrated in Figure 6.7.

Imports reverse this procedure—that is, US consumers demand yen to purchase Japanese products (demand for yen shifts to the right). This acts exactly the same as the demand for dollars in the previous example, except that now the vertical axis is US dollars for yen rather than yen for US dollars and it is clear that the price of yen in dollars is increased. In other words, fewer yen are required to purchase dollars and this results in the depreciation of the dollar relative to the yen.

Exchange Rate Regimes: Fixed, Floating, Pegged, and the Gold Standard

Under fixed exchange rates, the central bank tries to keep the exchange rate at a predetermined level by buying or selling foreign currencies in financial markets. During the Second World War, the United Nations held the "UN Monetary and Financial Conference" at the Mount Washington Hotel in Bretton Woods, New Hampshire, in July 1944. It was there that the

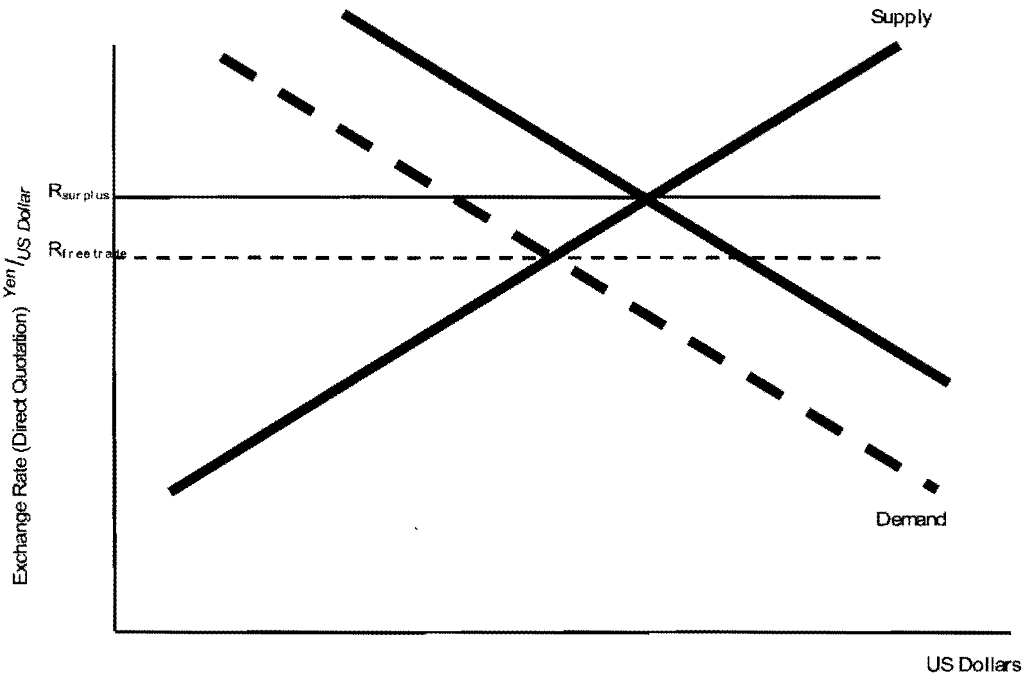

Figure 6.7 Demand for dollars

Allied nations, which represented most of the world leading industrial nations, signed the Bretton Woods Agreement. One of the main features of the Bretton Woods system was the attempt to maintain a fixed price of currencies in relation to gold. This agreement came to be known as the "gold standard" and was nothing more than an attempt to maintain fixed exchange rates by pegging the value of currencies to the price of gold.

Up until 1973 the longstanding Bretton Woods system was the major source of monetary management of the worlds leading industrial nations. When the price of gold and various global economic factors caused a devaluation of the US dollar, the gold standard was abandoned, and the US dollar became a floating currency (subject to complex factors that include open-market operations). Up until that time many countries had had their currencies pegged to the US dollar, since it was one of the few currencies that would readily convert to gold. After the collapse of the Bretton Woods system, relatively few countries remained pegged to the US dollar. Some private organizations are still pegged to the gold standard, but it has officially been abandoned by all nations.

Floating exchange rates are just what the name implies. That is, the value of the country's currency is allowed to fluctuate on the basis of the supply and demand for goods and services and/or capital investment. Over time, the exchange rate will fluctuate so that the real goods markets are always tending toward an equilibrium position.[5]

The appendix to this chapter contains a fairly detailed description of the various regional trade agreements that have been established worldwide. The extent and scope of these agreements provide conclusive proof that, regardless of what domestic politicians

5 That is, the real goods and services and capital investments that are purchased with the currency that is exchanged in the currency markets.

might say, the governments involved have concluded that reducing trade barriers is an effective means of increasing domestic wealth.

SUMMARY

This chapter has discussed the international aspect of the airline industry. Economists generally argue that the world at large will benefit from free trade and that trade liberalization can promote development. Trade protectionists claim that trade protection in most cases benefits all the domestic economies. The arguments for and against free trade have been explored, along with a presentation of the theories of absolute and comparative advantage. These theories were illustrated with quantitative examples. Next, the determination of foreign exchange rates was analyzed, along with historical methods for controlling exchange rates. Finally, the topic of free trade agreements was discussed.

APPENDIX: INTERNATIONAL FREE TRADE AGREEMENTS

Labor and regional trade agreements (RTAs) among different countries are the framework by which most of the world's economy is organized. Over the last few years the number of RTAs has significantly increased, but their effectiveness is a matter of controversy. Although the main purpose of many RTAs is to reduce trade barriers and encourage free trade, an increasing number of agreements also deal with other trade-related issues, such as the environment and labor. This appendix reviews and discusses several different regional trade agreements such as North American Free Trade Agreement, the Dominican Republic–Central America Free Trade Agreement, the US–Jordan Free Trade Agreement, and US–Australia Free Trade Agreements.

North American Free Trade Agreement (NAFTA)

NAFTA was originally signed on 1 January 1994 as a trade agreement that would link Canada, the United States, and Mexico. The immediate effects of NAFTA were an end to tariffs on goods (immediately for some, and over time for others). As a result of NAFTA, imports and exports from the United States to Canada and Mexico increased from one-quarter to one-third. One of the major fears with the implementation of NAFTA was the potential loss of jobs in the United States to Mexico. And, while manufacturing jobs in the United States did decrease, this loss was compensated for by the creation in the US economy of more than 2 million jobs per year from 1994 to 2000 in other industries.

Dominican Republic–Central America Free Trade Agreement (DR-CAFTA)

Originally this trade agreement was known as CAFTA with Costa Rica, El Salvador, Guatemala, Honduras, Nicaragua, and the United States in negotiation. However, in 2004 the Dominican Republic joined the negotiation, and the agreement was amended to the present name. CAFTA aims to create a trade free zone amongst its member nations like that of NAFTA. Eighty per cent of tariffs on US exports will be eliminated immediately, with the remaining 20 per cent to be phased out over time. This agreement has also been seen as a stepping stone for the FTAA (Free Trade Area of the Americas), which would lower tariffs amongst 34 countries in North and South America.

The US–Jordan Free Trade Agreement (JFTA)

The agreement between the United States and Jordan was implemented on 17 December 2001. This was the first agreement the United States made with an Arab nation, though it was the third free trade agreement implemented by the United States. The goal of this agreement is to eliminate tariff and non-tariff barriers on industrial goods and agricultural products over the following ten years of implementation.

US–Australia Free Trade Agreement (FTA)

The FTA is an agreement (modeled after NAFTA) between Australia and the United States implemented on 1 January 2005. This agreement, like others, sought to reduce barriers to trade; however, following its inception, trade from Australia to the United States declined, while trade from the United States to Australia has continued to increase.

European Union (EU)

The EU was established in 1993 by the Treaty on European Union (the Maastricht Treaty) and is the successor to the six-member European Economic Community (EEC), an organization established by the Treaty of Rome in 1957 between Belgium, France, Italy, Luxembourg, the Netherlands, and West Germany, and known informally as the Common Market (the Six). The EU is a confederation run by 25 member nations mostly located in continental Europe. Representatives of these nations make decisions partly by unanimity, partly by majority vote and partly by delegation to lesser bodies. It has its own flag, anthem, central bank, currency, elected parliament and Supreme Court as well as a common foreign and security policy.[6]

Citizens belonging to EU member states are also EU citizens. They are allowed to invest, live, travel, and work in all member states except for temporary restrictions on newly inducted member states. With a few exceptions, systematic border controls were mostly abolished by the Schengen Agreement in 1985. The EU economy relies on a complex web of multilateral trade agreements, international rules, and standards that cover products, markets, investment, health and environmental issues. There are still concerns about the nature of the union being intergovernmental (unanimous voting only) or supranationalist (majority votes imposed on all members); however, the EU has proved to be a mix of both. In the last five decades the EU has shown remarkable success in achieving economic prosperity and stability on a continental scale. It now accounts for about 30 per cent of global GDP and 20 per cent of global trade flows, and the euro has become an important

6 As of today, the European Union has 25 members including; Austria, Belgium, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden, and the United Kingdom.

international currency. The example of the EU is now considered a working model for regional integration.

Union of South American Nations (UNASUL)

Loosely modeled on the European Union, the Union of South American nations will combine the free trade organizations of MERCOSUR and the Andean community plus the three countries of Chile, Guyana, and Suriname. The Union's headquarters will be located in Quito, the capital of Ecuador. Formerly known as the South American Community of Nations, it was renamed at the First South American Energy Summit on 16 April 2007. The foundation of the Union was formally announced at the Third South American Summit, on 8 December 2004. Representatives from 12 South American nations signed the Cuzco Declaration, a two-page statement of intent. An important operating condition is the use of institutions belonging to the pre-existing trade blocs (MERCOSUR and Andean communities) to establish the union. So far, most of the countries within the union have waived visa requirements for travel, and there is an established consensus for a single South American currency.

AFTA (ASEAN Free Trade Area)

The Association of Southeast Asian Nations (ASEAN) is an organization of ten countries located in Southeast Asia. The ASEAN Free Trade Area (AFTA) agreement was signed by the member nations on 28 January 1992 in Singapore. When the AFTA agreement was originally signed, ASEAN had six members, namely Brunei, Indonesia, Malaysia, the Philippines, Singapore, and Thailand. New member nations were required to sign the agreement upon entry into ASEAN and were given timeframes within which to meet AFTA's tariff reduction obligations. Beginning in 1997, ASEAN began creating organizations within its framework with the intention of accelerating South-East Asian integration to include the People's Republic of China, Japan, South Korea, India, Australia, and New Zealand.

African Union (AU)

The African Union is an organization of 53 African states. It was created in 2001 from the amalgamation of various pre-existing regional blocs. The AU preserved the free trade areas established by these pre-existing blocs and will be combining and expanding them under the banner of the African Economic Community. The blocs are as follows:

- COMESA (Common Market for Eastern and Southern Africa): currently comprising 19 member nations. Officially established in 1994.
- EAC (East African Community): currently comprising five member nations. Established in 1967. Collapsed in 1977. Re-established in 1993. Common passport introduced in 1999.
- CEMAC (Economic and Monetary Community of Central Africa): currently comprising six member nations. Established in 1994.

- UEMOA (West African Economic and Monetary Union): currently comprising eight member nations. Established in 1994. Both customs and monetary union.
- SACU (South African Customs Union): currently comprising five member nations. Established in 1969.

The AU aims to achieve a single currency and a sustainable economy by bringing an end to intra-African conflict and creating an effective common market.

Greater Arab Free Trade Area (GAFTA)

The Greater Arab Free Trade Area came into existence on 1 January 2005. Similar to ASEAN, GAFTA was an agreement that was initially signed by 17 Arab League members; the agreement aimed at decreasing the customs on local production and the creation of an Arab Free Zone for exports and imports between members. The GAFTA rules involve member nations coordinating their tariff programs, maintaining common standards for specifications and restrictions on goods, promoting the private sector across all member countries, maintaining a base of communication, and decreasing customs duties. The members participate in 96 per cent of the total internal Arab trade, and 95 per cent with the rest of the world. Overall, the agreement would tie the pre-existing African-based AGADIR and Middle Eastern GCC organizations together to form one large free trade area.

South Asian Free Trade Area (SAFTA)

Born out of the efforts of the South Asian Association for Regional Cooperation (SAARC), the South Asian Free Trade Area was an agreement reached on 6 January 2004 for the creation of a free trade area involving India, Pakistan, Nepal, Sri Lanka, Bangladesh, Bhutan, the Maldives and Afghanistan. Involving almost 1.5 billion people, its influence is the largest of any regional organization in terms of population. The SAARC members have frequently expressed their unwillingness to sign free trade agreements. Though India has several trade pacts with the Maldives, Nepal, Bhutan, and Sri Lanka, similar trade agreements with Pakistan and Bangladesh have been stalled due to political and economic concerns on both sides. However, even with this slow progress, the foreign ministers of the member countries have signed a framework agreement to bring their duties down to 20 per cent by the end of 2007 and zero customs duty on the trade of almost all products in the region by the end of 2012.

Trans-Pacific Strategic Economic Partnership (TP SEP)

The Trans-Pacific Strategic Economic Partnership is a free trade agreement between Brunei, Chile, New Zealand, and Singapore, which was signed on 3 June 2005. The Trans-Pacific SEP was previously known as the Pacific Three Closer Economic Partnership (P3-CEP). Despite cultural and geographical differences, the four member countries share the similarities of being relatively small countries (in comparison with some of their trading partners) and are members of the Asia-Pacific Economic Cooperation (APEC). It aims to completely eliminate all trade tariffs by 2015. Because of an accession clause within the agreement, it has the potential to include other nations as well. Countries belonging to the 21-member APEC have shown some interest in this agreement.

Pacific Regional Trade Agreement (PARTA

The Pacific Regional Trade Agreement (PARTA) was founded in 1971 and is aimed at increasing trade between the island nations of the Pacific. Australia, New Zealand, New Caledonia, and French Polynesia are associate members of PARTA. Most of the member island countries are smaller in population and some are quite poor. Australia and New Zealand have much larger populations and are wealthier. Australia's population is around twice that of the other 15 members combined, and its economy is five times larger. Because of their position, the poorer countries are awarded concessional tariff deals to ease their exports.

Caribbean Community (CARICOM)

The Caribbean Community was originally called the Caribbean Community and Common Market and was established in 1973. Its membership has now grown to a total of 20 countries, the majority of which have joined the CARICOM Single Market and Economy (CSME) and the CARICOM Common Passport. Moreover, CARICOM is representing all its members as one single entity for bilateral agreements with the EU, members of NAFTA, and members of UNASUL. Twelve of the CARICOM countries have signed an oil alliance with Venezuela (Petrocaribe) which permits them to purchase oil on conditions of preferential payment.

Central American Common Market (CACM)

The Central American Common Market is an economic trade organization that was established in 1960 between the nations of Guatemala, El Salvador, Honduras, and Nicaragua. Costa Rica joined the CACM in 1963. The organization collapsed in 1969 due to a war between Honduras and El Salvador, and was reinstated in 1991. Because of its inability to settle trade disputes, the CACM has not been able to achieve all the goals of unification that were espoused in its founding. But, despite its shortcomings, the CACM has succeeded in removing duties on most products traded between its members, unifying external tariffs and increasing trade between its member nations.

Figure 6.8 compares the GDP of the different RTAs. The RTAs with the larger GDPs include nations such as the United States and Germany, while the smaller GDPs come from areas of underdeveloped countries as is the case with the AU.

From Figure 6.9 it can be determined that the EU is the only RTA with both a large number of members and a high amount of GDP. The AU has the highest number of members, but a relatively small GDP, when compared to the other RTAs. Also, as Table 6.9 and Figure 6.10 display in great detail, there are large population differences between the different RTAs. The AU easily has the largest population—in fact, almost double that of any other RTA—yet has one of the lowest GDPs. This further proves that this area

Table 6.9 Regional trade agreements and their scope

		(July 2007 est.)	PPP (2006 est.)	(2006 est.)	(2006 est.)	(2006 est.)	(2006 est.)	
Regional Trade Agreements	Nations	Population	GDP (millions)	Labor Force (millions)	Revenues (millions)	Expenditures (millions)	External Debt (millions)	Land Area (sq km)
NAFTA	3	443,230,979	$15,457,000	207	$2,789,000	$3,038,000	$1,867,000	20,227,980
DR-CAFTA	7	347,474,952	$13,392,910	170	$2,427,748	$2,680,121	$1,041,153	9,622,257
JFTA	2	307,193,140	$13,160,000	153	$2,413,880	$2,665,510	$1,011,300	9,253,894
FTA	2	321,574,123	$13,804,600	162	$2,676,000	$2,918,000	$1,589,100	16,779,853
EU	27	490,426,060	$13,349,081	236	$6,419,625	$6,721,840	$28,144,694	4,288,539
UNASUL	12	379,919,602	$3,378,837	175	$430,812	$397,384	$474,712	17,433,220
AFTA	10	526,368,563	$2,773,291	247	$207,076	$220,294	$353,063	3,703,563
AU	53	1,200,143,634	$2,426,322	290	$287,368	$285,027	$274,591	29,092,608
PARTA	18	33,222,512	$798,093	17	$314,022	$300,339	$634,944	8,425,404
CARICOM	20	15,405,363	$80,353	7	$14,835	$14,779	$17,511	439,045

Source: Compiled by the authors using data from the CIA's *World Factbook 2007*.

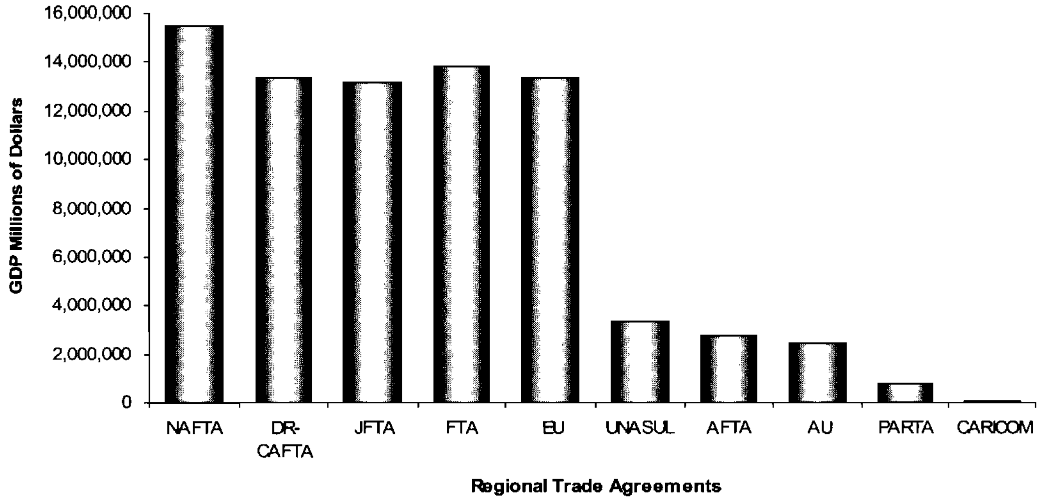

Figure 6.8 A comparison of regional trade agreements and their representative GDP

Source: Compiled by the authors using data from the CIA's *World Factbook 2007*.

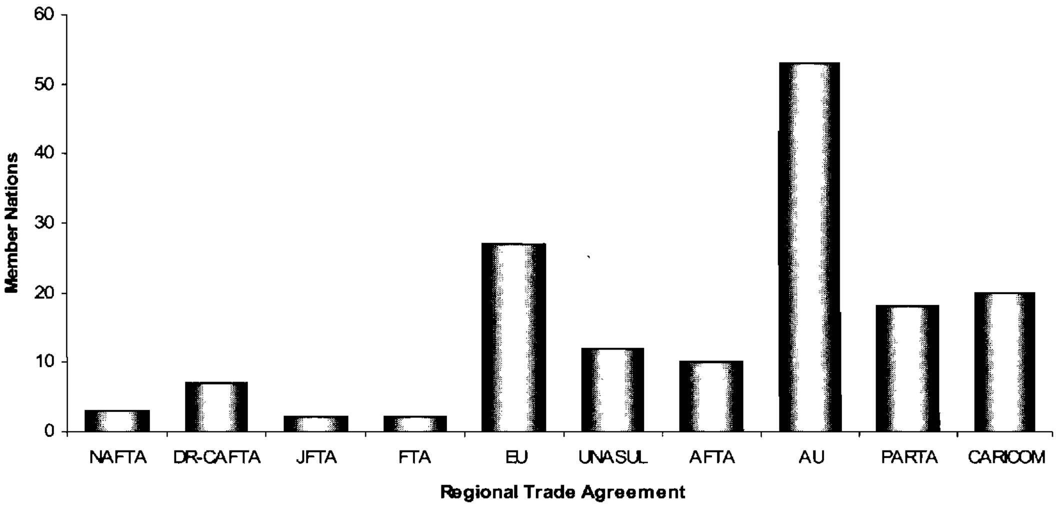

Figure 6.9 A comparison of regional trade agreements and their membership numbers

Source: Compiled by the authors using data from the CIA's *World Factbook 2007*.

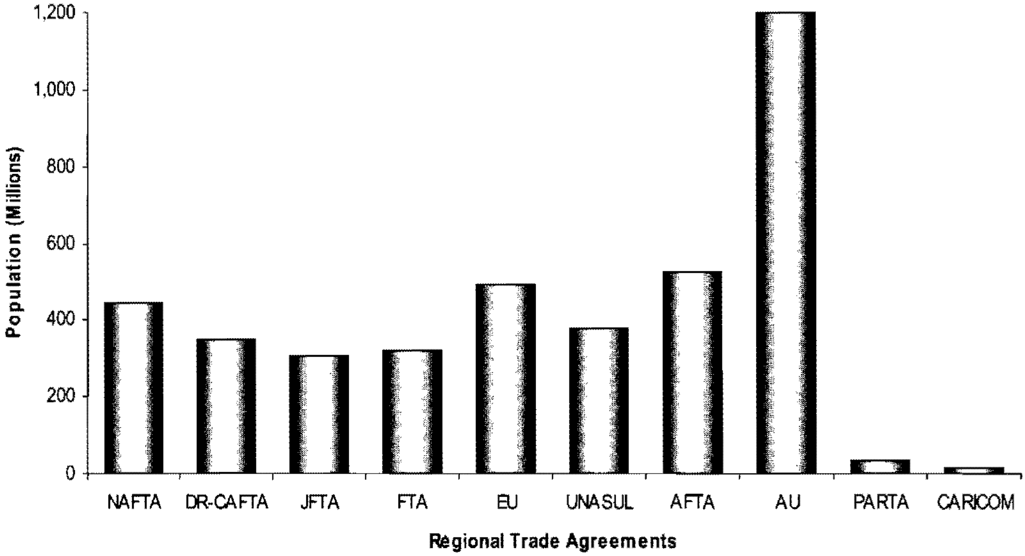

Figure 6.10 A comparison of regional trade agreements and their representative populations

Source: Compiled by the authors using data from the CIA's *World Factbook 2007*.

might be classified as underdeveloped. These figures prove that the AU has some huge untapped potential in terms of manpower.

A cursory inspection of the regional trade agreements discussed above reveals that there are many reasons for free trade agreements: to establish peace, to establish stronger economies in poor countries, to ease import and export restrictions and to jointly compete against other trade blocs. The regional trade agreements mentioned above are only a small example of the numerous arrangements between countries that promote trade. Despite often stated political objections to trade, it is clear that the nations of the world clearly believe that they will gain from trade and this is demonstrated by the international scope of the organizations discussed above.

REFERENCES

Ben-Yosef, E. (2005). *Evolution of the US Airline Industry*. The Netherlands: Springer.

Bradford, S. and Lawrence, R.Z. (2004). *Has Globalization Gone Far Enough? The Cost of Fragmented Markets*. Washington, DC: Peter G. Peterson Institute for International Economics.

Carbaugh, R. (2007). *International Economics* (11th edn). Mason, OH: South-Western.

Elwell, C. (2004). *The U.S. Trade Deficit: Causes, Consequences, and Cures* (Order Code RL31032). Washington, DC: The Library of Congress/Congressional Research Service.

EU Delegation (2004). *United States—Measures Affecting Trade in Large Civil Aircraft: Request for Consultations by the European Communities* (WS/DS317/1). Geneva: World Trade Organization, Dispute Settlement Body.

European Commission: Directorate General, Energy and Transport (2004). *International Aviation Agreements: Opening the Market for Efficient Air Travel*. Luxembourg: Office for Official Publications of the European Communities.

Hufbauer, G.C. and Elliott, K.A. (1994). *Measuring the Costs of Protection in the United States.* Washington, DC: Peter G. Peterson Institute for International Economics.

Husted, S. and Melvin, M. (2007). *International Economics* (7th edn). Boston, MA: Addison-Wesley.

Pugel, T. (2007). *International Economics* (13th edn). New York: McGraw-Hill.

US Air Force: Air Mobility Command (2006). *U.S. Air Force Fact Sheet: Civil Reserve Air Fleet.* Retrieved 13 June 2007 from US Air Force: Air Mobility Command Library at: http://www.amc.af.mil/library/factsheets/factsheet.asp?id=234.

US Census Bureau: Foreign Trade Division (2007). *Top Trading Partners Total Trade, Exports, Imports: Year-to-date December 2006.* Retrieved 13 June 2007 from US Census Bureau: Foreign Trade Statistics at: http://www.census.gov/foreign-trade/statistics/highlights/top/top0612.html.

US Central Intelligence Agency (2007). *The World Factbook 2007.* Dulles, VA: Potomac Books.

US Delegation (2004). *European Communities and Certain Member States—Measures Affecting Trade in Large Civil Aircraft: Request for Consultations by the United States* (WS/DS316/1). Geneva: World Trade Organization, Dispute Settlement Body.

US Department of Commerce: Bureau of Economic Analysis (2007). *US International Transactions.* Retrieved 13 June 2007 from Bureau of Economic Analysis: International Economic Accounts.

Van den Berg, H. (2004). *International Economics.* New York: McGraw Hill.

World Trade Organization (2006). *International Trade Statistics 2006.* Geneva: WTO Publications.

7

International Aviation: Open Skies and Global Alliances

> It is probably not love that makes the world go around, but rather those mutually supportive ALLIANCES through which partners recognize their dependence on each other for the achievement of shared and private goals.
>
> Fred A. Allen

One of the great benefits of aviation has been its ability to make the world smaller and promote globalization. From aviation's very roots, the international regulatory environment has been critical to the success of the aviation industry. This chapter will explore two major themes encompassed in international aviation: open skies and global alliances. Open skies deals with the legal framework surrounding the rights granted to airlines, with its roots tracing back to the original air transportation agreements between countries. (While the true ideal of open skies has yet to be achieved, it is the goal for international aviation.) Global alliances deal with the arrangements that airlines have made with one another to expand their scope on a global basis. The specific topics that are covered in the chapter are listed below:

- A brief history of international aviation agreements
- Bilateral air service agreements, including:
  - Freedoms of air transportation
  - The Bermuda Agreement
- Open skies agreements, including:
  - Characteristics of open skies
  - Benefits of open skies
- Open skies in Europe
- Open skies in Asia
- Global airline alliances, including:
  - History of global airline alliances
  - Benefits of global alliances
  - Disadvantages of global alliances
  - The future for global alliances

A BRIEF HISTORY OF INTERNATIONAL AVIATION AGREEMENTS

The first international agreement concerning air transportation occurred shortly after the end of the First World War in Paris. As a consequence of the tremendous leap in aviation that occurred during the war, delegates from 26 countries drew up the Convention Relating to the Regulation of Air Navigation (US Centennial of Flight [USCOF], 2006). The Convention voted to give each nation, "complete and exclusive sovereignty over the airspace above its territory" (USCOF, 2006). This was the first time that countries had been provided with an internationally recognized legal authority over their airspace, enabling them to allow or disallow aviation access into their country. At the end, neither Russia nor the United States signed the Paris Convention of 1919 (USCOF, 2006).

The United States signed its first international aviation agreement at the Havana Convention on Civil Aviation of 1928. This agreement guaranteed the innocent right of passage as well as the formulation of rules concerning such issues as aircraft navigation, landing facilities, and pilot standards. The Havana Convention also provided the right for each country to set the route to be flown over its territory. In total, the United States and 20 other Western hemisphere countries signed and ratified the Havana Convention of 1928 (USCOF, 2006).

The Convention for the Unification of Certain Rules Relating to International Carriage by Air was convened on 12 October 1929 in Warsaw, Poland. One of the major results of this convention was a formal definition of "international carriage." Article 1 of the Warsaw Convention states:

> … "international carriage" means any carriage in which, according to the contract made by the parties, the place of departure and the place of destination, whether or not there be a break in the carriage or a transshipment, are situated either within the territories of two High Contracting Parties, or within the territory of a single High Contracting Party, if there is an agreed stopping place within a territory subject to the sovereignty, suzerainty, mandate or authority of another Power, even though that Power is not a party to this Convention. A carriage without such an agreed stopping place between territories subject to the sovereignty, suzerainty, mandate or authority of the same High Contracting Party is not deemed to be international for the purposes of this Convention.
>
> (Warsaw Convention, 1929)

The Convention also established a general set of guidelines for the operation of the commercial air transportation industry for international flights. For example, Article 3 describes the requirements for a passenger ticket, while Article 4 outlines what needs to be included on a luggage tag (Warsaw Convention, 1929).

One of the more practical outcomes of the Warsaw Convention concerned air carrier liability. Article 17 states that the air carrier is liable for:

> …damage sustained in the event of the death or wounding of a passenger or any other bodily injury suffered by a passenger, if the accident which caused the damage so sustained took place on board the aircraft or in the course of any of the operations of embarking or disembarking.
>
> (Warsaw Convention, 1929)

This article states that the air carrier is liable for death or bodily injury suffered by a passenger on an air carrier's flight. However, Articles 20 and 21 provide escape clauses for the airlines if it is determined that they took all measures necessary to avoid the loss, or there was some contributory negligence on behalf the person (Warsaw Convention, 1929). Such issues are not as important today since aviation is extremely safe, but it was important at the time of this convention.

Eventually, the Warsaw Convention was completely overhauled by the Montreal Convention of 1999, which is the current convention governing international carriage liability. Article 21 of the Montreal Convention states that an air carrier has unlimited liability—that is, there is no maximum cap on the payment, and that in the event of death, the minimum that the airline must compensate is 100,000 Special Drawings Rights, SDR, (Montreal Convention, 1999).[1] This translates into roughly $150,000.

The next major international agreement concerning air transportation was the Chicago Convention of 1944 held near the end of the Second World War and hosted by US President Franklin D. Roosevelt. Roosevelt's goal for this convention was revolutionary in that he wanted an agreement that would allow any airliner from any country to fly to any other country with little or no restriction (Phillips, 2006). What Roosevelt was pushing for was a true open skies agreement, whereby there would be few, if any, restrictions on international flying. Unfortunately, few of the 54 delegations attending the Chicago Convention actually backed him on his goal for open skies (Phillips, 2006). Instead, Article 6 of the Chicago Convention created a system of bilateral air service agreements between countries for all scheduled international flying. The article states:

> No scheduled international air service may be operated over or into the territory of a contracting State, except with the special permission or other authorization of that State, and in accordance with the terms of such permission or authorization.
>
> (Chicago Convention, 1944)

A major outcome of the Chicago Convention was the creation of the International Civil Aviation Organization (ICAO) with the objective of, "develop[ing] the principles and techniques of international air navigation and to foster the planning and development of international air transport" (Article 44, Chicago Convention, 1944). The Chicago Convention superseded the previous Paris and Chicago Conventions and it still remains the major basis for all international aviation law.

BILATERAL AIR SERVICE AGREEMENTS

After the Chicago Convention, bilateral air service agreements between countries became the predominant method of regulating international air transportation. These agreements controlled market access, market entry, and, in many cases, market pricing (Doganis, 2001). In granting market access, countries allow various degrees of freedom of air transportation. There are up to eight degrees of freedoms that may be granted in bilateral air service agreements.

1 The SDR is an artificial currency unit based on several national currencies. DSR is used by the International Monetary Fund (IMF) for internal accounting purposes and by some countries as a peg for their own currency, and is used as an international reserve asset.

Freedoms of Air Transportation

The *first freedom* provides the right for an airline to fly over another country without landing (Doganis, 2001). An example of a first-freedom right would be an international nonstop flight from Los Angeles to London that overflies Canada on the way to England. In order for this to occur, Canada must grant the United States first-freedom rights. Today, with a few exceptions, almost all countries grant unilateral first-freedom rights. Moreover, countries may charge airlines for the permission to overfly their country, essentially placing economic barriers on the first freedom. For example, the overflight fee for the United States is $33.72 per 100 nautical miles over the continental United States (FAA, 2006). Russia is notorious for charging high overflight fees, especially on new polar flights from North America to Asia. This provides an economic hindrance for airlines flying such routes.

The *second freedom* of the air is the right to make a landing for technical reasons (such as refueling) in another country without picking up or setting down revenue passengers (Doganis, 2001). An example of a flight requiring the second freedom would be Cathay Pacific's flight from Hong Kong to Toronto, Canada, with a refueling stop in Anchorage, Alaska. In order for this flight to occur, the United States would have to grant Hong Kong second-freedom rights. The second-freedom right is usually granted since the airline provides revenue to the granting country in terms of landing fees and fuel purchase, but does not compete with the domestic airlines. With today's modern aircraft, the requirement for refueling stops has diminished greatly, but only a few years ago the second freedom was important to many operators. Today, it is cargo carriers that make the most use of second-freedom rights.

The *third* and *fourth freedoms* of the air are essentially two sides of the same coin. The third freedom grants the right to carry revenue traffic from your own country to another country, while the fourth freedom provides the right to carry revenue traffic from the other country back to your own country (Doganis, 2001). These rights are usually granted together in order to allow an airline to operate a return air service. Third and fourth freedoms may only be granted for certain city pairs in air service agreements, and this puts limitations on air travel. Any international flight that carries passengers between two countries requires third and fourth freedoms for that particular flight.

Fifth-freedom rights enable an airline to carry revenue traffic from their own country to another country, and then pick up and drop off traffic from the intermediate country to a third country (Doganis, 2001). For these rights to be useable, the third country must also agree to the right. A prime example of fifth-freedom rights in action is Cathay Pacific's flight from Hong Kong to Vancouver and then on to New York. On this flight Cathay Pacific is allowed to carry traffic from Hong Kong to Vancouver and also from Vancouver to New York. In order for Cathay Pacific to operate this flight, Canada and the United States must grant Hong Kong fifth freedom for the route. Fifth-freedom rights are rarely granted since the foreign airline is now competing with domestic airlines for the same traffic. However, there are other examples of fifth-freedom rights, such as Northwest's inter-Asia operation from Tokyo Narita and EVA Airways which operates the Taipei–Bangkok–London flight with full traffic rights from Bangkok to London. Fifth-freedom rights are highly desirable to airlines, as segments can be tagged on to an existing flight, and this increases its profitability.

The *sixth freedom* of the air allows an airline to carry traffic between two other countries by using its home base as a transit point (Doganis, 2001). A prime example of this is an airline that flies a passenger from Europe to North America and then transfers that passenger onto a flight to Mexico or Central America. Usually sixth freedoms are not granted explicitly, but are given implicitly when third and fourth freedoms are granted (Doganis, 2001). A slightly modified form of the sixth freedom, or "modified sixths," would allow an airline to transfer a passenger through its hub from two points in the same foreign country (Field, 2005). An example of a "modified sixths" would be a passenger flying on a Canadian airline from Boston to Seattle that uses Toronto, Canada, as its transfer point. Currently "modified sixths" are not allowed, but there is movement in North America to possibly allow this to happen (Field, 2005).

The *seventh freedom* allows an airline to carry revenue traffic between points in two countries on services which lie entirely outside its own home country. The liberalization of European airspace allowed seventh-freedom rights as airlines are now allowed to fly throughout Europe. For instance, Ryanair, an Irish airline, can fly from Germany to Portugal. The tremendous access provided by seventh-freedom rights enables increased competition; this has lowered air fares, and increased air travel demand. While Europe has allowed seventh-freedom rights, they are rarely granted by other countries.

Finally, the *eighth freedom* is probably the most controversial freedom. Also referred to as cabotage, the eighth freedom allows a foreign airline to fly between two domestic points in a country. For instance, cabotage rights would need to be granted if Qantas wanted to continue its Los Angeles flight on to New York with local traffic between Los Angeles and New York. Cabotage is controversial because it allows foreign competition on domestic routes. Few countries have granted cabotage rights, but they are actively sought during negotiations. True open skies between two countries would require cabotage rights from both countries. Europe's liberalization of air transport has allowed cabotage rights since the entire European Community is considered domestic from an air transportation perspective. Figure 7.1 graphically summarizes the eight freedoms of air transportation.

The Bermuda Agreement

The first bilateral air service agreement was signed between the United States and the United Kingdom on 11 February 1946 (shortly after the Chicago Convention) in Bermuda (DOT, 1978). While this agreement holds the distinction of being the first bilateral air service agreement to be signed, it is also one of the longest standing and most important bilateral agreements in aviation today. The agreement has been updated numerous times and overhauled in 1977, creating the Bermuda 2 agreement. The Bermuda 2 agreement is a good example of bilateral air transport agreements between countries; therefore, it is instructive to take a closer look at some of the more important provisions of the agreement.

One of the major themes of most bilateral agreements is regulatory approval on airfares. Both countries must approve the pricing of tickets for all carriers operating between the two countries. This "double approval" of tariffs is usually related to a cost plus profit formula, ensuring a profitable operation while keeping airfares artificially high. Other possibilities for tariff regulation are "dual disapproval," zone pricing, or free pricing (or no pricing regulation) (Doganis, 2001).

The tariff approval system in the Bermuda 2 agreement is based on Article 11 which states in part that "the designated airlines of one Contracting Party shall have a fair and

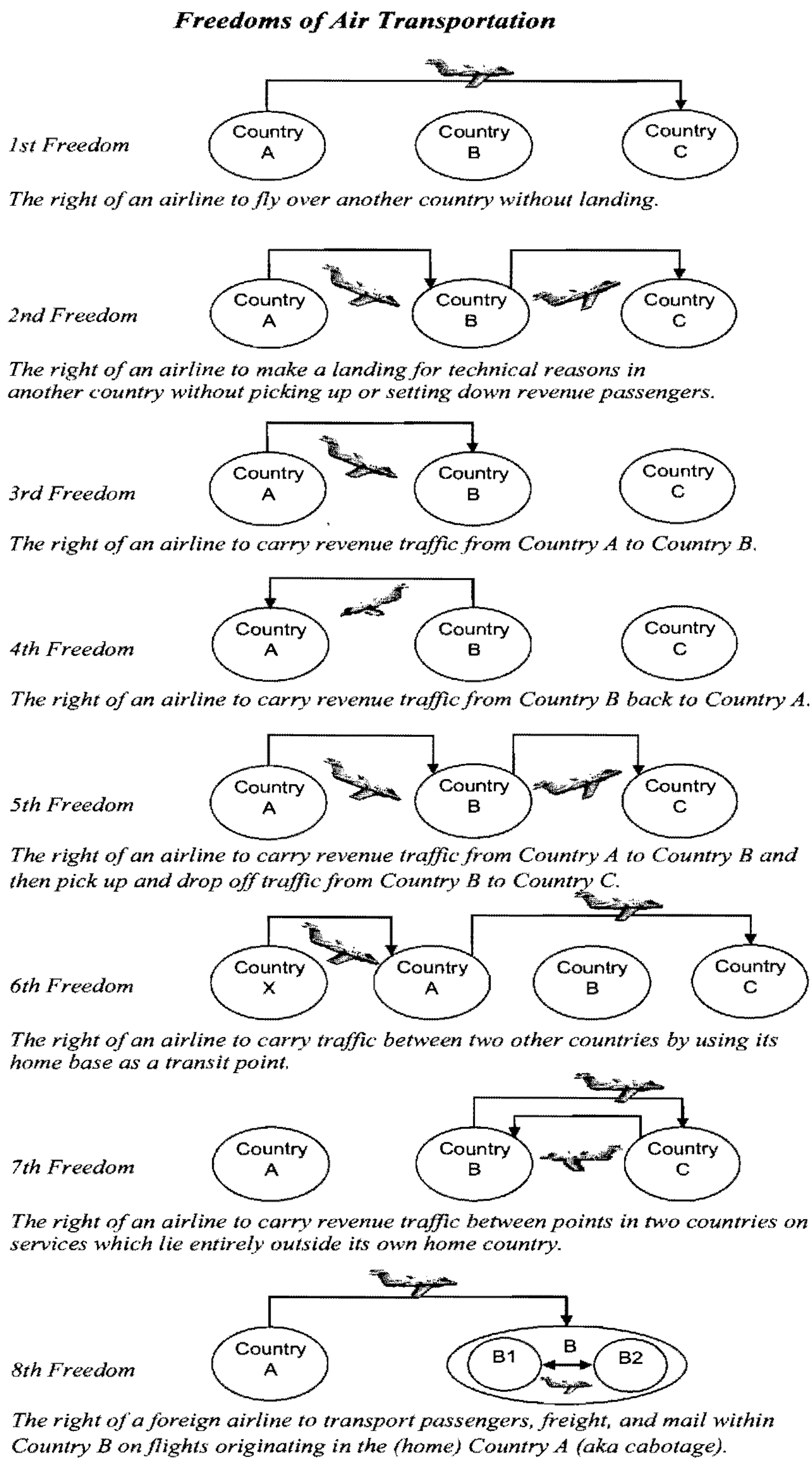

Figure 7.1 Freedom of air passage

equal opportunity to compete with the designated airlines of the other Contracting Party" (DOT, 1978). This stipulation governs all actions by the airlines, including tariffs, and therefore it acts to create some level of price fixing. The article contains stipulations against capacity dumping, which could severely impact the profits of the other operators. Today, with greater liberalization of competition laws, the tariff approval mechanism largely rubber-stamps airlines' requests for changes in fares.

The Bermuda 2 agreement also deals with other issues concerning security, airworthiness, dispute resolutions, and customs issues; however, from both an economic and an airline standpoint, the greatest impact of the Bermuda 2 agreement is the granting of air freedoms and route authorities. Although the rights granted by the Bermuda 2 agreement were considered quite liberal at the time, the agreement itself is now considered very bureaucratic as it places numerous restrictions on air transportation.

First, Article 2 of the Bermuda 2 agreement grants the rights bestowed on airlines of both the UK and the United States. Airlines are granted:

> a) the right to fly across its territory without landing; and
> b) the right to make stops in its territory for non-traffic purposes.
>
> (DOT, 1978)

Part (a) provides the first freedom to airlines from both countries, while part (b) grants second-freedom rights to all airlines. Article 2 goes on to grant fourth- and fifth-freedom rights as long as they are a part of the agreed-upon routes (DOT, 1978).

One of the major restrictions placed in the Bermuda 2 agreement is that only two airlines from each country are allowed to operate scheduled passenger services from London Heathrow to the United States (Competition Commission, 1999). Today only British Airways, Virgin Atlantic, American Airlines, and United Airlines are permitted to fly from Heathrow to the United States. This has effectively created a government-enforced cartel that considerably inhibits competition, especially considering the fact that Heathrow is London's most desired airport for passengers. American and United received these rights from Trans World and Pan Am respectively, and, in doing so, received considerable windfall profits. Heathrow is the only airport that has such restrictions, and all other carriers operating between the United States and London must do so from secondary UK airports, such as London Gatwick. US airlines, such as Delta and Continental, have been lobbying hard for access into London Heathrow, but the current bilateral agreement restricts them from doing so.

Under the Bermuda 2 agreement there are only certain cities that can be served by US and UK airlines from either London Heathrow or London Gatwick. On the other hand there are no route restrictions placed on routes not originating or departing from Heathrow or Gatwick. Airlines such as Continental have been successful at operating from other UK airports such as Manchester, Bristol, Birmingham, Belfast, Glasgow, and Edinburgh. The only US cities that US airlines can serve from Heathrow or Gatwick are: Anchorage, Atlanta, Boston, Charlotte, Chicago, Cincinnati, Cleveland, Dallas/Fort Worth, Detroit, Houston, Los Angeles, Miami, Minneapolis/St Paul, New York (both JFK and Newark), Philadelphia, San Francisco, Seattle, St. Louis, and Washington/Baltimore (Competition Commission, 1999). In addition to these cities, there are a few "switchable" cities that can be added to the list so long as an equal proportion of cities are dropped from the list (Competition Commission, 1999). These "switchable" cities include Fort Lauderdale, Honolulu, Kansas City, Las Vegas, New Orleans, Pittsburgh, and Portland

(Competition Commission, 1999). While this is an impressive list of cities, there are numerous other American cities that might benefit from a true open skies policy—that is, one that allowed flights from any US city to London. Moreover, because this policy amounts to a government-enforced restraint of trade, fares and profits from the favored cities will be higher than they otherwise would be with open competition.

On the reverse side, UK carriers have a slightly different list of US cities that they are permitted to serve from London. The cities are: Atlanta, Boston, Charlotte, Chicago, Dallas/Fort Worth, Denver, Detroit, Houston, Los Angeles, Miami, New York (both JFK and Newark), Orlando, Philadelphia, Phoenix, Pittsburgh, San Diego, San Francisco, Seattle, Tampa, and Washington/Baltimore (Competition Commission, 1999). The same comments about fares and profits that were made in the paragraph above apply here.

The archaic system of allowing only certain cities to be flown to from Heathrow not only impacts on airlines, but also on communities. For instance, American Airlines, one of the four carriers permitted to fly into Heathrow, cannot fly a Heathrow route from its largest hub, Dallas, because the city is not an approved Bermuda 2 city. Bilateral air service agreements, such as the Bermuda 2 agreement, place tremendous restrictions on international travel. While the Bermuda 2 agreement has some liberal aspects to it, other bilateral agreements may actually contain clauses that limit the number of seats that can be flown between the countries each day.

The general thrust of bilateral air service agreements has been to protect national interests and provide support for national airlines. While such protectionism helps carriers who receive the benefits, it is frustrating for airlines looking from the outside-in. Because of their protectionist nature, bilateral agreements curtail a market solution to international air travel and replace it with government regulation. Generally speaking, the artificial restrictions that are imposed by bilateral air transport agreements raise costs, create inefficiencies in the market, and allow rent-seeking behavior on the part of the favored airlines. Opening up the international skies would be similar to the deregulation movement that occurred domestically in the United States in 1978. Open skies would not only benefit consumers and the economy, but also increase the airlines' profits and reduce their costs.

OPEN SKIES AGREEMENTS

In 1992 the United States and the Netherlands signed the first "open skies" agreement (Doganis, 2001). This was followed by similar "open skies" agreement between the United States and Canada in 1995 (Field, 2005). As of 2006 the United States has signed 77 "open skies" agreements with countries ranging from Germany to Chad, and from Chile to Uzbekistan (DOS, 2006). The general trend in international aviation is to do away with complicated and restriction-laden bilateral agreements and move towards "open skies". In March 2007 the United States and the EU entered into an open skies agreement that will open all of the EU and United States. This agreement came into effect in March 2008 and is estimated to have an impact of over 12 billion dollars per year (Alford and Champley, 2007).

Characteristics of Open Skies

US open skies agreements generally contain eight key provisions (DOS, 2006). The first, and probably the most important, provision contained in all open skies agreements is the absence of restrictions on international route rights (DOS, 2006). This means that carriers from either country are free to fly between any two cities they wish, with whatever size aircraft they wish, as many times a day/week as they want. This lack of restrictions lowers barriers to entry for airlines, but does not entirely eliminate them as carriers may still require landing slots at foreign airports in order to initiate a new flight. However, open competition will allow airlines to bid for these rights and, ultimately, they will be assigned to their highest valued economic bidder.

The second major provision included in open skies agreements is that airline pricing should be determined by market forces (DOS, 2006). While, under a true open skies agreement, the governments would play no role in airline pricing, the US model does include a "double-disapproval" stipulation, whereby a fare can be disallowed if both countries agree (DOS, 2006). In practice, carriers are allowed to set whatever fares they want, but the presence of a "double-disapproval" stipulation could possibly prevent some low-cost carriers from entering certain international markets and offering deeply discounted fares, if disallowing such fares happened to be politically attractive to the two governments involved.

The third major provision contained in US open skies agreements is a clause ensuring fair and equal opportunity to compete (DOS, 2006). This clause covers a wide variety of issues, such as non-discriminatory airport slot allocations or user fees. This provision also covers issues such as availability of ground-handling and establishing sales offices. In essence, countries should allow airlines of both countries equal opportunity to be able to compete fairly.

The fourth major provision allows airlines to enter cooperative marketing agreements. (DOS, 2006). As will be pointed out later in the chapter, airline alliances are critical to the success of airlines. Prior to open skies agreements, restrictions could be placed on the air carriers' ability to enter alliance agreements with airlines of both countries. For instance, while both British Airways and American Airlines are both founding member of the oneworld alliance, they are not permitted to code-share on each other's flights under the current Bermuda 2 agreement (Button, 2002). Ideally, under open skies, airlines are permitted to enter whatever code-share agreements that they wish. Thanks to the open skies agreement between the United States and the Netherlands, Northwest Airlines and KLM were able to enter a strong alliance that included revenue-sharing between the airlines. In order for this cohesive agreement to occur, the open skies agreement had to be in place, and the extensive code-share agreement between the airlines was given antitrust immunity from the US Department of Justice (Doganis, 2001). While US open skies agreements allow full code-share agreements, they do not address issues pertaining to foreign ownership of airlines. This is a sticky point in current US–EU open skies negotiations.

Other provisions that are often contained in open skies agreements are mechanisms for dispute settlement, consultation pertaining to unfair practices, liberal legal charter agreements (whereby carriers can choose to operate under the charter regulations of either country), and agreements pertaining to the safe and secure operation of flights between the two countries (DOS, 2006). In open skies agreements the United States also seeks the provision that all-cargo flights be granted seventh-freedom rights (DOS, 2006), so that cargo flights can operate between the other country and a third country via flights that are not linked to their homeland. This stipulation enables airlines like FedEx and UPS to operate cargo hubs in foreign countries. Currently only about half of the open skies agreements signed by the United States contain this optional eighth provision (DOS, 2006). Fifth-freedom rights are rarely provided for in open skies agreements. However, one agreement that does allow this is the new US–Canada open skies agreement whereby Canadian carriers received unilateral fifth-freedom rights from the United States in exchange for seventh-freedom all-cargo rights for US carriers.

Benefits of Open Skies

The benefits achieved from open skies agreements are similar to those obtained from domestic deregulation. As a result of open skies agreements, airlines are able to fly more routes, which ultimately leads to increased competition, resulting in lower average airfares. Open skies also enables new city pairs (domestic to foreign) to be flown that were previously not possible. In general, consumers benefit from open skies as they receive more frequent service and pay lower prices as a result of increased competition. From the airline's perspective, however, open skies may also result in greater fluctuations in profitability.

Airlines also benefit from open skies agreements, but those benefits vary, depending on the airline's position in the market. For example, the airline that already has extensive rights to the foreign country is currently receiving some windfall profits from the protection it is receiving in the market. With open skies, that airline would no longer receive the protection and would face more competition. Carriers that are currently excluded from a market (or have limited service) gain more from open skies agreements than carriers that already have extensive route rights. This is the major reason why airlines such as Continental and Delta have been lobbying hard for an open skies agreement between the United States and the UK, while carriers such as American, which has extensive London access rights, have been relatively quiet in their lobbying.

As mentioned above, the United States signed one of the first open skies agreements with Canada. In a study conducted by the US Department of Transportation three years after the signing of the open skies agreement, it was found that trans-border traffic averaged an 11.1 per cent yearly growth rate compared to 1.4 per cent per year for the three years prior to the agreement (DOT, 1998). Moreover, the number of nonstop markets with over 50,000 annual passengers increased from 54 in 1994 to 77 in 1997 (DOT, 1998). It is estimated that 38 new city pairs were opened up between Canada and the United States as a result of the agreement. (DOT, 1998). While this tremendous growth rate in the market will not continue, the figures clearly show the large latent demand, from both business and tourism, which was being suppressed before the open skies agreement. The original open skies agreement between the two countries has been recently amended (in 2005) to provide both countries with increased freedoms, but some, including Air Canada President, Robert Milton, want to see the North American market resemble the European market in which a Canadian carrier would essentially be granted US cabotage rights, and vice versa (Field, 2005). However, the likelihood of such rights being granted is probably slim.

While the EU and United States have not signed and ratified a unified open skies agreement, the benefits from such an agreement would be immense. In a study conducted by the Brattle Group, it is estimated that an EU–US open skies agreement would increase transatlantic traffic by up to 11 million passengers per year—a 24 per cent increase—and boost intra-Europe travel by an additional 35.7 million passengers per year—a 14 per cent increase (Robyn, 2003). This increase in traffic would create an additional $5.2 billion increase in consumer benefits (through lower fares), while increasing economic output of related industries by $3.6 billion to $8.1 billion a year (Robyn, 2003). As these numbers indicate, an open skies agreement in the biggest international market would provide substantial economic benefits. This will be discussed in more detail in the next section.

OPEN SKIES IN EUROPE

Europe has had a successful experience of open skies with the creation (in 1997) of a single European aviation market (Kinnock, 1996). Under this single European aviation market, European carriers are free to fly routes throughout Europe. For instance, British Airways could fly from Paris to Frankfurt or Amsterdam to Rome. This granting of seventh-freedom rights also included the granting of eighth-freedom rights, or cabotage. This further enabled a British airline to be able to offer Frankfurt to Munich flights or Barcelona to Madrid flights. In fact, British Airways created a German subsidiary to fly domestic German routes. National ownership has become irrelevant for intra-European flights, and this is the primary reason why low-cost carriers such as easyJet and Ryanair have been able to expand rapidly. Although European liberalization has increased competition and helped reduce airfares throughout Europe, making aviation a viable competitor to train travel, it has also caused many airline bankruptcies; this, of course, is similar to what happened in the United States after deregulation (Kinnock, 1996).

Although the European Union has successful liberalized intra-Europe travel, global travel from Europe is still largely dominated by each country's respective flag carriers (de Palacio, 2001). German airlines cannot fly from London to the United States, while British Airlines cannot fly from Paris to Japan. This is a result of the current bilateral agreements between individual European countries and other countries around the world; these agreements limit international flights to airlines with full national ownership (de Palacio, 2001). This problem is made more complicated by the fact that individual European countries have signed bilateral agreements with foreign countries at a time when the European community is attempting to become a single market. Because it has had stronger leverage over each individual nation than over the collective whole, the United States currently has open skies agreements with most European nations. However, these nationality clauses have been deemed illegal in a 2002 European Court of Justice (ECJ) ruling, and this has placed pressure on the European Union to create multilateral aviation agreements with foreign countries (Baker, 2005).

As mentioned in the previous section, the largest such multilateral agreement would involve the European Community and the United States. Both sides want to reach an open skies agreement whereby all EU and US airlines can fly from any point in the European Union to and from any point in the United States, provided they can reach agreements at airports concerning landing rights, gates, and counter space (Phillips, 2006). Such a multilateral open skies agreement would permit Lufthansa to fly Paris to New York, while US carriers would have unlimited rights into Europe. This would effectively end the current nationality clauses that are deemed illegal in the current bilateral agreements.

Although a tentative open skies agreement between the EU and the United States has been reached, its ratification has been held up in courts (Phillips, 2006), but very recently has been ratified. A technically unrelated issue concerning foreign ownership of US airlines has been tied together with the open skies agreement, causing the process to stall (Phillips, 2006). Currently US airlines are allowed only 25 per cent foreign ownership with foreign owners being denied management control. The EU wants to increase the limit to 49.9 per cent foreign ownership and allow limited management direction of the US carrier (Baker and Field, 2004). Such a change in the ownership cap will require Congressional approval, while the EU Council of Ministers has stated that it will not approve the open skies deal unless Congress improves foreign ownership issues (Phillips, 2006).

This has left both sides at a stalemate, and has stalled an open skies agreement between the parties. This is unfortunate since an open skies agreement would benefit both airlines and consumers from all the participating countries. For the US airlines, one of the greater attractions of such an agreement is that it would end the Bermuda 2 agreement, and enable US airlines access to London Heathrow (assuming that there are slots available).[2] For their part, EU airlines would be provided with unfettered access to the world's largest aviation market, although they would probably not be granted cabotage rights to fly domestic US flights. For consumers on both sides of the Atlantic, the largest international market would grow considerably as the number of routes increased and ticket prices decreased. As history has shown, with any liberalization of aviation markets there have been airline bankruptcies and failures as airlines compete in the new markets. Since the current regulatory framework between both parties is inefficient and illegal, a multilateral open skies agreement between the United States and the EU is bound to happen but the question is: 'When?'

OPEN SKIES IN ASIA

While most of the world's aviation industries have adopted liberalized aviation markets, the Asia-Pacific region has not implemented open skies largely on account of perceived national interests identical to those that prevented the United States and Europe from liberalizing their airspace (Oum and Yamaguchi, 2006). One reason for this is the fact that the aviation industry in the Asia-Pacific region is still developing in comparison with the mature North American and European markets. Despite the fact that these markets have shown that the benefits of open skies far outweigh the benefits of protectionism, it may take some time before this is fully appreciated in this region. Two of the more successful airlines in the region are Emirates Airlines and Singapore Airlines. These two carriers play a large role in promoting their relatively small countries and have been successful, through the adoption of open skies agreements, in the creation of global aviation hubs.

Australia is another country that has implemented open skies by creating a single aviation market with New Zealand. Australia has also eliminated foreign ownership restrictions (Oum and Yamaguchi, 2006). This removal of foreign ownership restrictions has enabled Richard Branson to start up Virgin Blue in Australia. On the other hand, although Australia has considerably liberalized aviation, it has also denied Singapore Airlines' request to fly from Sydney to Los Angeles. This decision was taken undoubtedly to protect Qantas Airways from competition on this route

India, a long-time highly regulated aviation market, has slowly become more liberalized as the economy has grown; however, much of this liberalization effort has only occurred in the domestic market. Nevertheless, this liberalization has spurred tremendous growth in domestic air travel, and the industry has created several successful low-cost carriers. Only

2 In 2007 Delta Airlines gained access to Heathrow's runways after negotiations with Air France and KLM, its European partner airlines, to launch services in April 2008 when a new open skies regime will open up the airport.

recently has the Indian government permitted private Indian carriers to fly internationally, and many of the country's bilateral agreements have severe capacity restrictions placed on them.

In northern Asia, Japan and Korea still are highly regulated, with most international routes containing capacity restrictions. Liberalization of the Japanese market is even more difficult due to airport restrictions at congested airports such as Tokyo Narita. Although Japan does not have an open skies agreement in place with the United States, both Northwest and United operate significant hub operations out of Narita airport as a result of the fact that the United States was granted liberal fifth-freedom rights after the Second World War. In fact, very few countries in the Asia-Pacific region have open skies agreements with the United States.

Probably the most attractive country for foreign carriers to fly to in the Asia-Pacific region is China. However, for political reasons, China has also historically been one of the most restrictive countries in the region. One of the political sanctions that affect the aviation industry was the prohibition of direct flights between China and Taiwan. While this hindered both Chinese and Taiwanese carriers, the regulations benefited Macau which was used as a transiting point between China and Taiwan. While China has opened up its country to foreign business ownership, the liberalization of the aviation sector has progressed at a much slower pace.

Domestically, the Chinese aviation industry used to be all government-owned, but operated by multiple small carriers. Recently, the domestic industry has consolidated, forming three large Chinese carriers—Air China, China Southern Airlines, and China Eastern Airlines—which are based at Beijing, Guangzhou, and Shanghai respectively (Francis, 2004).This consolidation has enabled the Chinese carriers to become stronger internationally, and allowed slow, progressive reform in international aviation. In addition, China has allowed private ownership of airlines, including a foreign ownership cap of 49 per cent, which might possibly be raised in the future (Francis, 2004).

While China has liberalized some of their international aviation agreements, these agreements are still quite restrictive. For instance, China and Singapore reached an "open skies" agreement, but the "open skies" agreement forbids Singapore low-cost carriers from flying into Shanghai and Beijing (Francis, 2005). The agreement did, however, permit full open skies to all other Chinese cities. In this respect it is just a further example of certain "open skies" agreements not truly being open skies.

China is also experimenting with other liberalized air policies, such as creating an entirely open aviation policy for the Hainan region of China (Francis, 2004). Under this open skies policy, foreign carriers are permitted unlimited access to the Hainan region, and have full fifth-freedom rights, as well as limited cabotage rights to other Chinese cities other than Beijing, Guangzhou, and Shanghai (Francis, 2004). This policy constitutes an effort to open up aviation markets outside China's three dominant cities, and is probably also protectionist in nature since these three cities also happen to be the hubs of China's three largest airlines (Francis, 2004).

Probably the most ambitious economic liberalization project to occur in the Asia-Pacific region is the multilateral open skies agreement between the ASEAN (Association of Southeast Asian Nations) countries. The ten members of the ASEAN are Brunei, Singapore, Thailand, Cambodia, Indonesia, Laos, Malaysia, Myanmar, the Philippines, and Vietnam. Under the proposed agreement between these countries, all member nations would permit unlimited flights between their capital cities by December 2008 and full open skies to other cities by December 2010. These reforms would be a giant step forward for air transport liberalization in the region, since many of the ASEAN countries have historically had very protectionist viewpoints toward aviation. Such air transport liberalization would benefit not only individual countries, but also the economic region as a whole in both the short and long term (Forsyth, King, and Rodolfo, 2006). Although a few member countries have pushed for early adoption of this open skies agreement, the Asia-Pacific region still remains a heavily regulated industry.

GLOBAL AIRLINE ALLIANCES

In order to help overcome restrictive barriers to entry in international markets, many airlines have formed alliances with foreign carriers. The purpose of these alliances has generally been an attempt to introduce service in those countries and regions where there have been legal or financial restrictions. Although the initial global alliances started as simple code-share agreements between airlines, they have evolved into global alliances with multiple airlines that span the globe. The introduction of global alliances has also magnified the prominence of global or regional hubs. Today nearly 60 per cent of the world's total air traffic flies on some sort of a global alliance (Baker, 2006).

History of Global Airline Alliances

In 1986 the first international airline alliances was signed between Air Florida and British Island whereby Air Florida provided a passenger feed for British Island's London–Amsterdam route (Oum, Park, and Zhang, 2000). This basic form of an alliance between two carriers, commonly called a code share, enables an airline to place passengers onto the flight of another carrier. Code-share agreements can be signed that cover a few specific routes or flights, or they can cover almost all of the airline's flights. While this original code-share agreement was quite simple, code-share agreements have evolved to the point that they may even include some blocked seats for the code-sharing airline. For instance, in 1993 Air Canada and Korean Air entered into a code share agreement whereby each airline purchased 48 seats per departure on the other airline's flight.

In the above example, the inventory assigned to the airline is fixed, but in other cases it can be variable, depending on demand. After this initial code-share agreement was signed, multiple carriers signed code-share agreements. Some examples of these are: Japan Airline and Thai Airways in 1985; American and Qantas on Qantas's transpacific flights in 1986; and Air France and Sabena with a blocked space agreement on the Paris to Brussels route in 1992 (Oum, Park, and Zhang, 2000).

In 1992 Dutch carrier KLM and US carrier Northwest formed a major transatlantic airline alliance whereby a broad code-share agreement was put in place. In 1993 the alliance received antitrust immunity from the US Department of Transportation, thereby enabling the two airlines to closely coordinate their flights across the Atlantic. This led to a joint operation venture in which revenues were divided between the two carriers, regardless of the operating airline. Although joint ventures were not a new phenomenon to the airline industry (both Braniff and Singapore Airlines operated a quasi-joint venture with the Concorde aircraft), This agreement, because it covered so many flights not only across the Atlantic, but also in Europe and North America, was precedent-setting. Under the terms of the agreement, KLM purchased 25 per cent of Northwest's voting rights, and

49 per cent of Northwest's total equity share (Oum, Park, and Zhang, 2000). While equity investments bring the carriers closer together and enable the investing airline to help shape overall strategy, it may also limit the flexibility of the alliance. For instance, Gudmundsson and Lechner (2006) argue that one of the downfalls of the Qualiflyer alliance was that Swissair's equity investments made it difficult for the alliance partners to enter new agreements.[3] While Swissair's equity investment highlights the pitfalls involved, such agreements have also proven successful, such as British Airways' 25 per cent investment in Qantas in 1993, Air Canada's 27.5 per cent investment in Continental also in 1993, and Singapore Airlines' current 49 per cent investment in Richard Branson's Virgin Atlantic (Oum, Park, and Zhang, 2000).

The next development in the airline alliances model was the creation of global alliances. While Delta, Swissair, and Singapore had the initial roots of a global alliance in the late 1980s, the first truly global alliance was formed in 1997 between United, Lufthansa, SAS, Air Canada, and Thai Airways (Baker, 2001). This alliance was called the Star Alliance and was shortly followed by similar global alliances such as: the Qualiflyer (1998) that originally included Swissair, Sabena, Turkish Airlines, Air Liberté, and TAP Air Portugal; the oneworld alliance (September 1998) between American, British Airways, Qantas, and Cathay Pacific; and the Skyteam alliance (September 1999), originally between just Delta Air Lines, Air France, and Aeroméxico (Baker, 2001). Since then, the global alliances have added more airlines in an attempt to reach all the corners of the globe. Airlines have been more than willing to join these alliances on account of the undoubted benefits, and they also have not wanted to fall behind in the alliance game.

Today, three major global alliances exist (Star, Skyteam, and oneworld), and these alliances generated a combined 58.04 per cent of global airline revenue in 2005. The Star alliance comprises 21 airlines, and is the largest global alliance, with a 22.18 per cent global market revenue share. It also carries approximately 425 million passengers per year. Skyteam has been expanding recently, and obtained a 20.48 per cent global market share in 2005, with 372 million passengers per year. The Star alliance also initiated an associate program for smaller carriers in 2005. oneworld is the smallest alliance in terms of revenue with only 15.38 per cent global market share. However, this is expected to change with three new airlines scheduled to join the alliance. Of these airlines, Japan Airlines is the most important since it was the largest carrier in the world that had not yet joined a global alliance.

Table 7.1 gives a full list of carriers, with an entry date for each alliance, while Table 7.2 and Figures 7.2 through 7.4 provide various comparison statistics for the alliances; these include number of employees, countries served, and total aircraft fleet.

All three alliances have expanded to provide geographical coverage to all areas of the world. The two regions in which the alliances currently lack coverage are China and India. oneworld has opted to rely on Cathay Pacific for its China feed as the Hong Kong-based carrier took over Dragonair, an airline specializing in Hong Kong–China flights (Baker, 2006). With the considerable consolidation taking place in the Indian market it is unlikely that any of the alliances will enter into formal negotiations with the Indian Airlines in the near future (Baker, 2006).

Two major trends that have emerged in the global alliance game are the push for Eastern European members and the creation/acceptance of smaller regional airlines. Czech Airways was the first Eastern European carrier to join a formal global alliance, and

3 The Qualiflyer alliance was a European alliance led by Swissair and Sabena.

Table 7.1 Global alliance membership, 2006

oneworld		Skyteam		Star	
Aer Lingus[1]	Jun-00	Aeroflot	07-Apr	Air Canada	May-97
American Airlines	Sep-98	Aeromexico	Jun-00	Air New Zealand	Mar-99
British Airways	Sep-98	Air France/KLM	Jun-00	All Nippon Airways	Oct-99
Cathay Pacific	Sep-98	Alitalia	07-Jul	Asiana Airlines	07-Mar
Finnair	Sep-99	Continental Airlines	07-Sep	Austrian Airlines	Mar-00
Iberia	Sep-99	Czech Airlines	07-Mar	bmi	Jul-00
LAN	Jun-00	Delta Air Lines	Jun-00	LOT Polish Airlines	07-Oct
Qantas	Sep-98	Korean Air	Jun-00	Lufthansa	May-97
Royal Jordanian[2]	05-Jun	Northwest Airlines	07-Sep	SAS Scandinavian Airlines	May-97
Malev[2]	05-Jun	China Southern Airlines[3]	05-Jun	Singapore Airlines	Apr-00
Japan Airlines[2]	05-Jun	Air Europa[4]	07-Jun	South African Airways	07-Apr
		Copa[4]	07-Jun	Spanair	07-Apr
		Kenya Airways[4]	07-Jun	Swiss	07-Apr
		Tarom[4]	07-Jun	TAP Portugal	07-Mar
		Middle East Airlines[4]	07-Jan	Thai Airways	May-97
		PGA Portugalia[4]	07-Jun	United Airlines	May-97
				US Airways	07-May
				Varig[5]	Oct-97
				Air China[6]	05-Jun
				Shanghai Airlines[6]	05-Jun
				Blue[1,7]	07-Nov
				Adria Airways[7]	07-Dec
				Croatia Airlines[7]	07-Dec

Source: Compiled by the authors, using different sources.

Notes:
1. Aer Lingus withdrew from oneworld in April 2007.
2. oneworld member-elect airlines that will join oneworld in the near future.
3. Joined Skyteam in 2007.
4. Are associate members of the Skyteam alliance.
5. Varig withdrew from the Star alliance in January 2007.
6. Has formally applied to join the Star alliance and are currently limited members.
7. Are regional members of the Star alliance.

Table 7.2 Global airline alliances: facts and figures

Year\Revenues (% of global industry)	oneworld[1]	Skyteam[2]	Star[3]
2006	18.8	19.1	25.4
2005	15.4	20.5	22.2
2004	15.2	19.4	20.6
Passengers (millions)	319	364	405
Daily Departures	9,190	14,711	15,935
Number of Employees	266,426	279,133	351,761

Source: Compiled by the authors, using different sources.

Notes:
1. oneworld statistics include JAL, Malev, and Royal Jordanian.
2. Skyteam statistics exclude associate Skyteam members and Southern China Airlines.
3. Star alliance statistics exclude region members, Air China and Shanghai Airlines.

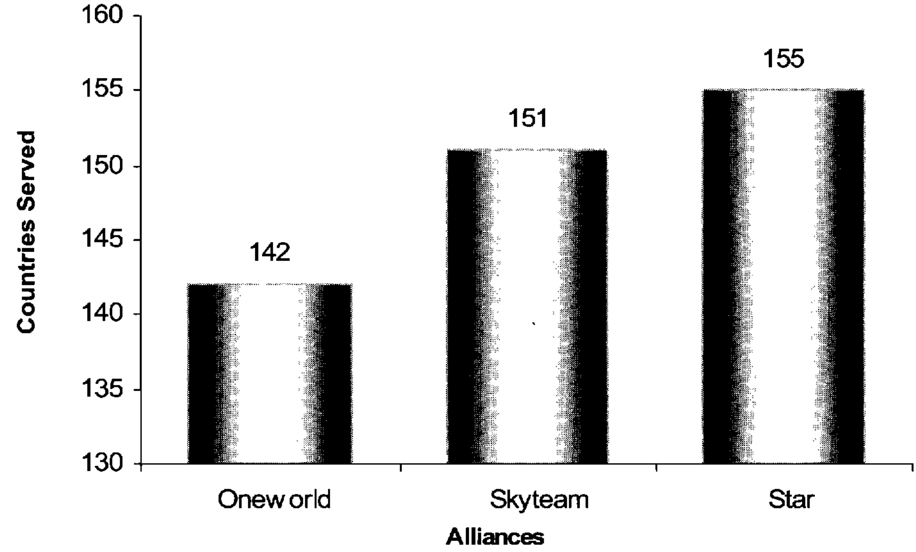

Figure 7.2 Global airline alliances: countries served as of June 2007

Source: Compiled by the authors, using different sources.

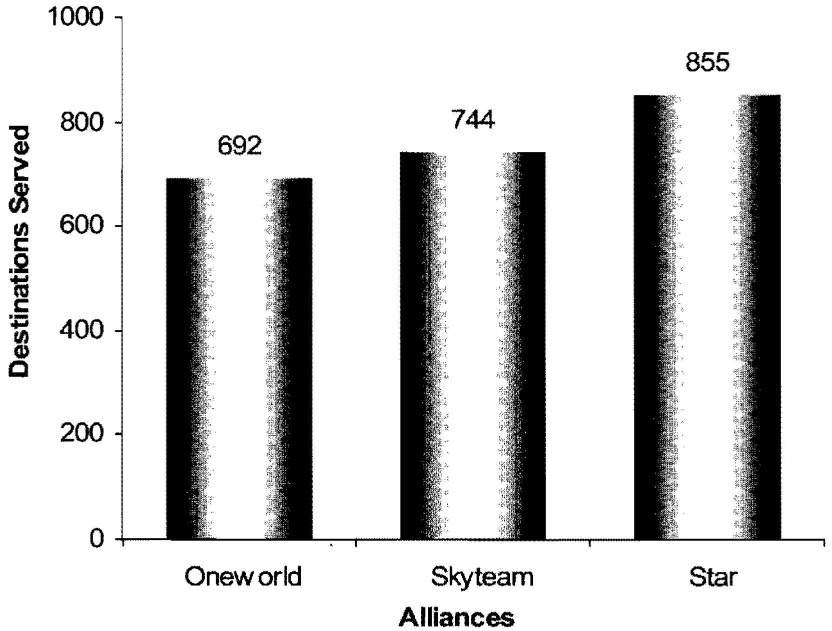

Figure 7.3 Global airline alliances: destinations served as of June 2007

Source: Compiled by the authors, using different sources.

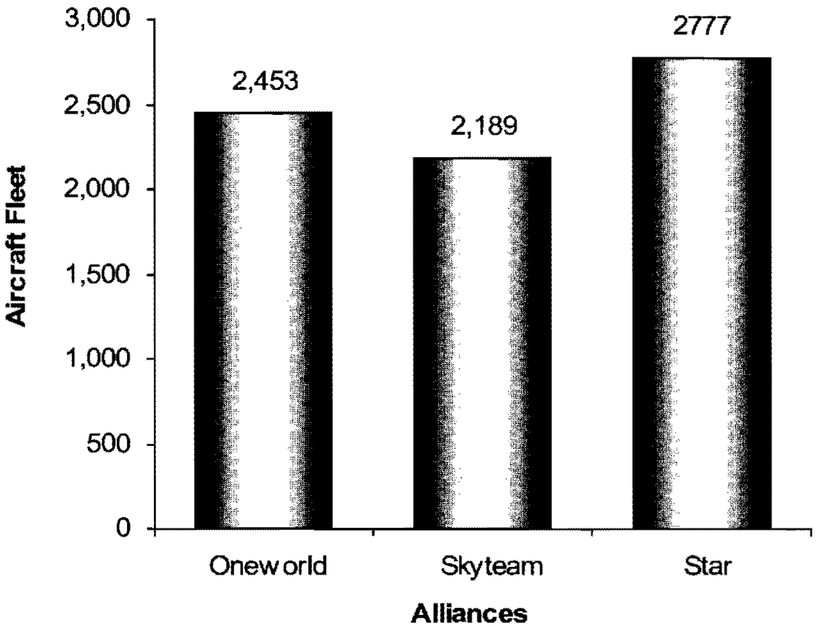

Figure 7.4 Global airline alliances: aircraft fleet as of June 2007 (aircraft fleet statistics exclude aircraft from related carriers)

Source: Compiled by the authors, using different sources.

since then numerous other Eastern Europe carriers have joined alliances as the region has prospered over the past few years. Aeroflot finally joined Skyteam in April 2006 after months of work in bringing the carrier into line with the other alliance members. Also, both Star and Skyteam have created regional/associate members for niche airlines that serve a particular need. These regional/associate members must have a sponsoring carrier for entry into the alliance. Although oneworld has no formal regional/associate program, the introduction of airlines such as Malev and Royal Jordanian show that it is also interested in specific regional carriers.

While most of the major airlines are part of a global alliance, a few carriers have decided not to join alliances. At one time, Japan Airlines was the largest airline not in a global alliance, but recently carriers such as Virgin Atlantic and Emirates have decided to go about business separately on the grounds that alliances would not provide them with any additional benefits—although both these carriers do have code-share agreements. Moreover, both airlines worry about declining service standards that might result from joining a global alliance. In the United States, Alaska Airlines has chosen not to join a global alliance, but has decided to have multiple broad code-share agreements with carriers from both oneworld and Skyteam. On the other hand, Aer Lingus has decided to back out of the oneworld alliance as it attempts to redefine itself as a low-cost carrier—of which there are currently none in a global alliance.

Benefits of Global Alliances

The major benefit of global alliances to the airlines and consumers is an expanded and optimized route network. Through global alliances, passengers can, in theory, easily travel from one destination to another, anywhere in the world, with one ticket from one airline. Before global alliances or code-share agreements, international travel was more complex because passengers might have to purchase multiple tickets on multiple airlines to fly to their desired destination. Through global alliances the airlines are now able to sell tickets to destinations that they previously could not serve for one reason or another.

Global alliances provide a traveler with greater flight options and help make the traveling experience more enjoyable. The various routings that are available from the three alliances reduce the international travel time for passengers. From an airline's perspective, while it might not find it profitable to fly directly to a certain city, it can still carry passengers to that city on connecting flights. In this way, alliances enable airlines to achieve benefits from economies of scope, and this is why the alliances are seeking a strong presence in every major market worldwide.

To be successful, the connections between carriers in an alliance must be seamless and easy. Information technology is critical to success in this area. Alliance members all operate different information technology platforms, but they still need to interact effectively. Each of the three alliances are addressing the information technology and ticketing structures differently. Whereas Star is creating a common platform that all carriers may choose to use, Skyteam is not pursuing a common information technology approach (McDonald, 2006). oneworld is the first airline alliance to have full e-ticketing between all its members. Skyteam undoubtedly will have difficulties achieving full e-ticketing connections since Russian law requires that Aeroflot provide paper tickets (McDonald, 2006). Nonetheless, full e-ticketing is probably critical to the future success of these alliances on account of its convenience for passengers and cheapness for the airlines.

Although the exact traffic and revenue benefits of joining an alliance are difficult to quantify, oneworld has estimated that the alliance generated almost $400 million in additional revenues for the eight alliance members. Oum, Park, and Zhang compared traffic increases on alliance routes to non-alliance routes. In the short-lived USAir/British Airways alliance, British Airways was able to increase traffic by 8.3 per cent on alliance routes over non-alliance routes (Oum, Park, and Zhang, 2000). Iatrou and Skourias compared the traffic difference on alliance routes between the pre- and the post-alliance periods for all three major alliances. On average, they found that traffic increased by 9.4 per cent as a result of airline alliances; however, the greatest increases in traffic were experienced by Skyteam and Star, while traffic for oneworld actually decreased (Iatrou and Skourias, 2005). They also found, not surprisingly, that traffic increased the greatest when an alliance received antitrust immunity.

Another major benefit of airline alliances is cost reduction in maintenance and operational activities as a result of bulk purchasing and sharing of resources. As mentioned above, joint information technology could greatly reduce costs among alliance members. One area that the Star alliance has pioneered is joint purchasing with the aim of achieving greater volume discounts. In an effort to reduce fuel costs for alliance members, the alliance launched Star Fuel Co. in December 2003 (Mecham, 2004). Using volume discounts, Star Fuel Co. was able to reduce fuel costs for Star alliance members at Los Angeles, San Francisco, and London Heathrow by $50 million in 2004 (Mecham, 2004). Star alliance members Air Canada, Austrian, Lufthansa, and SAS all explored a joint regional aircraft purchase in the same year, but the initiative did not come to fruition because Air Canada and Austrian made independent purchase decisions (Field and Pilling, 2004). Although the airlines concerned were able to develop common specifications for the aircraft, the failure of the joint purchasing agreement highlights the difficulties that may be encountered when separate alliance members have unique and different requirements and objectives. The Star alliance intends to pursue joint aircraft purchasing again in the near future, but its best strategy is probably to focus on the joint purchasing of commodities such as fuel and aircraft parts (Field and Pilling, 2004).

Airports are another major area in which alliances are looking to reduce costs. Airlines commonly use personnel from alliance members in order to help reduce costs, but alliances are seeking to extend this further by hosting operations all under one roof. Under this scenario not only would the carriers be able to share airport resources, but they would also provide passengers with easier connections. The airport at which all the alliances would like dedicated facilities is London Heathrow. With the opening of Terminal 5 in spring 2008, airlines will begin shuffling all over the airport in order to consolidate their operations (Thompson, 2004).

The creation of dedicated alliance terminals such as that described at Heathrow would not only help reduce costs, but also enable alliances to have shared business lounges, self-service check-ins, and possibly ground service. The Star alliance is actively pursuing similar "one-roof" initiatives in such airports as Paris Charles de Gaulle, Nagoya, Tokyo Narita, Miami, Bangkok, and Warsaw (Thompson, 2004). Additionally, in an effort to improve customer service, the Star alliance has initiated a program at three airports, where Star alliance customer service teams meet incoming flights and ensure that passengers and their baggage make their connecting flight (Thompson, 2004).[4] While the program may seem more like a customer service idea, the Star alliance claims that the program saves approximately $3.3 million per year by reducing passenger misconnects and lost baggage claims (Thompson, 2004).

4 The airports include: Chicago O'Hare, Frankfurt, and Los Angeles.

Having a strong global presence also enables the alliances to obtain global corporate travel agreements for corporations that require global travel (Field and Pilling, 2004). Skyteam estimates that there are about 75–100 organizations that require global corporate travel agreements (Field and Pilling, 2004). This additional source of revenue can be substantial for alliances, and the attractiveness of their products over competing alliances depends greatly on the alliances' global market coverage.

Another benefit that frequent travelers enjoy is the fact that global alliances enable travelers to accrue and redeem miles on a variety of airlines. No longer do frequent-flyer programs limit travel to just one airline; they may now include 20 or even 30 airlines. Furthermore, depending on the passengers' frequent-flyer status, they can receive reciprocal benefits such as upgrades, priority boarding/check-in, and/or lounge access. This is also one of the reasons why alliance members have to adhere to a certain level of quality, and it is also one of the principal reasons why no low-cost carrier has yet to join an alliance.[5]

Disadvantages of Global Alliances

The greatest cost involved in joining global alliances is that incurred in bringing information technology systems in line with other alliance members. These costs can be considerable, especially if the system needs a complete overhaul. The time needed for harmonization of IT systems is one of the reasons why alliance members usually announce their intention to join an alliance 6–12 months before they actually join. In the case of South African Airlines, it took nearly two years for the airline to finally join the Star alliance (Star Alliance, 2006).

There are additional costs associated with global alliances, such as increased overhead costs. The Star alliance is the most structured alliance with a full-time staff of around 70 people (Field and Pilling, 2004). The reason for this is not only the fact that the Star alliance is so large that it needs some oversight, but also because it is the most aggressive of the three alliances in creating a master brand (Field and Pilling, 2004). oneworld is slightly less structured, with a full-time staff of 23 people, but the majority of these people are employed in sales for the various oneworld ticket packages (Field and Pilling, 2004). Finally, Skyteam is even less structured with no true overhead group, instead relying on conversations between airlines to create alliance structure (Field and Pilling, 2004). Regardless of the alliance's structure, such coordination efforts raise overhead costs to some degree.

One area of potential concern for various alliance partners is that, as new airlines are added to the alliance, the importance of an existing airline may be reduced. In other words, carriers can be in competition with their own alliance partners. An interesting example of this is Delta's and Continental's battle on transatlantic routes, even though both airlines are Skyteam members and should technically be working together. Because of these issues, Gudmundsson and Lechner (2006) argue that airlines will drop out and switch alliances in order to achieve the greatest benefits. This has already begun with Aer Lingus and Mexicana dropping out of oneworld and the Star alliance respectively. While switching costs may prohibit a high degree of alliance movement by airlines, there will still be situations where airlines may feel that it is in their best interests to switch.

5 Note that both Westjet in Canada and Gol in Brazil have expressed an interest in joining a global alliance (Baker, 2006).

Future for Global Alliances

Table 7.3 and Figure 7.5 display the impact of global alliances on the top ten airports in the world, in terms of passengers. The market share for each airline is calculated based on 2006 ASMs published in the Official Airline Guide (OAG). In terms of concentration, the Star alliance commands the greatest market share. On the other hand, Atlanta airport is the most concentrated airport, where the Skyteam alliance, largely driven by Delta, commands an 87 per cent market share of the airport. This results in a highly concentrated Herfindahl Index value of 7,604. Almost all the airports are highly concentrated when

Table 7.3 Concentration of the top ten global airports by alliance in terms of ASMs

Alliances	ATL	ORD	LHR	HND	LAX	DFW	CDG	FRA	DEN	LAS
Star	4%	54%	17%	43%	30%	9%	10%	67%	62%	30%
oneworld	2%	31%	47%	0%	21%	79%	4%	5%	5%	3%
SkyTeam	87%	7%	2%	1%	20%	7%	62%	5%	8%	20%
Other	7%	8%	33%	56%	29%	5%	24%	24%	25%	46%
Herfindahl Index	7,604	3,995	3,621	4,967	2,572	6,361	4,571	5,048	4,607	3,439

Source: Compiled by the authors from 2006 OAG data.

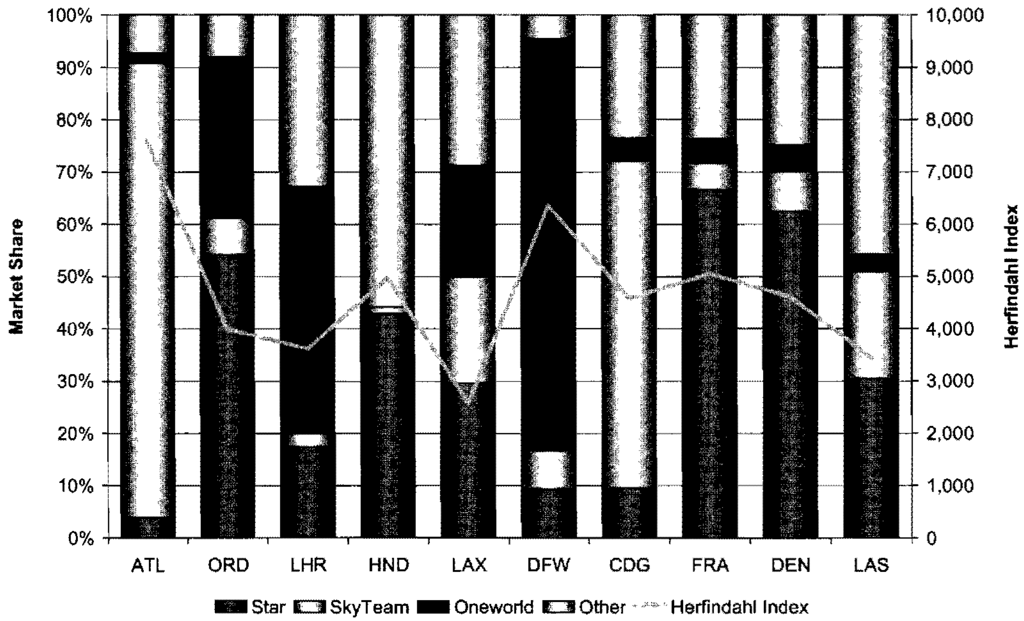

Figure 7.5 Concentration of the top ten global airports by alliance and market share

Source: Compiled by the authors from 2006 OAG data.

grouped by global alliances—much more so than when grouping the airlines individually. Of the top ten airports, Los Angeles is the least concentrated with all three alliances, and non-aligned carriers, having a fairly even market distribution. Even market distribution results in greater competition and lower airfares, while greater concentration generally leads to reduced service and higher air fares—a negative effect for consumers, but a positive effect for the airlines.

Since their inception, global alliances have been very successful as airlines have been able to expand their presence throughout the world. Whereas originally there were various alliance groupings, they have now consolidated into three major global alliances: oneworld, Skyteam, and Star. All three alliances have grown from their inception and will continue to do so, especially in India and China. Other than these two areas, all three alliances have sound geographic coverage in all of the world's major markets. And, with most of the world's major carriers already in a global alliance, there are few large airlines left to join established alliances. This means that any future alliance growth will probably be a result of the addition of smaller, regional airlines that can serve a specific niche that the alliance may be lacking. Moreover, given the uncertainties inherent in the aviation industry, the alliances are certain not to remain stable.

In the future, it will be interesting to see how aggressive each alliance is in promoting its brand. The Star alliance has heavily promoted its brand, with alliance members all painting their aircraft in the Star alliance color scheme. The extreme end of the spectrum for alliance branding is that every flight is sold as a Star alliance flight, but operated by an individual airline. This is not likely to happen, however, since there are legal issues associated with this approach, and the individual airlines would not like to lose their identity in their domestic markets.

International aviation is extremely important to the success of airlines as, for many, it is the last frontier for profitability. Awareness of international air service agreements and the push for open skies is critical to the evolution of international aviation, while global alliances enable airlines to exploit international opportunities.

SUMMARY

This chapter has presented an institutional discussion of aviation agreements and the concept of open skies. It has covered global airline alliances and their benefits and costs. A detailed presentation of the various aviation freedoms that are contained in international agreements has been given, and, finally, the future of open skies has been discussed.

REFERENCES

Alford, E. and Champley, R.(2007). *The Impact of the 2007 US -EU Open Skies Air Transport Agreement.* ITA Occasional Paper no. 07-001. International Trade Administration.

Baker, C. (2001). The Global Groupings. *Airline Business, July.* Retrieved on 12 October 2006 from Air Transport Intelligence at: http://www.rati.com/iata/oapd.

Baker, C. (2005). Back to the Table. *Airlines Business, September.* Retrieved on 28 September 2006 from Air Transport Intelligence at: http://www.rati.com/iata/oapd.

Baker, C. (2006). Stellar Orbit. *Airline Business, September.* Retrieved on 16 October 2006 from Air Transport Intelligence at: http://www.rati.com/iata/oapd.

Baker, C. and Field, D. (2004). Europe Rules out US Open Skies Offer. *Airlines Business, April.* Retrieved on 10 October 2006 from Proquest at: http://www.proquest.com.

Button, K.J. (2002). Toward Truly Open Skies. *Regulation, Fall.* Retrieved on 3 October 2006 from Proquest at http://www.proquest.com.

Chicago Convention (1994). *Convention on International Civil Aviation, Signed at Chicago, on 7 December 1944.* Retrieved on 3 October 2006 from: http://www.mcgill.ca/files/iasl/chicago1944a.pdf.

Competition Commission. (1999). *Bermuda 2.* Retrieved on 5 October 2006 from: http://www.competition-commission.org.uk/rep_pub/reports/1999/fulltext/430a4.2.pdf.

de Palacio, L. (2001). Open Skies: How to Get the Airlines Airborne Again. *Wall Street Journal (Europe),* 11 September. Retrieved on 10 October 2006.

Doganis, R. (2001). *The Airline Business in the 21st Century.* London: Routledge.

DOS (2006). *Open Skies Agreement Highlights.* Retrieved on 4 October 2006 from: http://www.state.gov/e/eb/rls/fs/2006/208.htm.

DOT (1978). *Air Services Agreement Between the Government of the United States of America and the Government of the United Kingdom of Great Britain and Northern Ireland.* Washington: DOT.

DOT (1998). *The Impact of the New US–Canada Aviation Agreement at its Third Anniversary.* Retrieved on 5 October 2006 from: http://ostpxweb.dot.gov/aviation/intav/canada2.pdf.

FAA (2006). *Overflight Fees.* Retrieved on 5 October 2006 from: http://www.faa.gov/airports_airtraffic/international_aviation/overflight_fees/index.cfm?print.

Field, D. (2005). True Open Skies? *Airline Business,* March. Retrieved on 28 September 2006 from Air Transport Intelligence at: http://www.rati.com/iata/oapd.

Field, D. and Pilling, M. (2004). Team Spirit. *Airline Business,* September. Retrieved on 12 October 2006 from Proquest at: http://www.proquest.com.

Forsyth, P., King, J. and Rodolfo, C. (2006). Open Skies in ASEAN. *Journal of Air Transport Management,* 12, pp. 143–52.

Francis, L. (2004). Liberal Values. *Flight International,* October. Retrieved on 12October 2006 from Air Transport Intelligence at: http://www.rati.com/iata/oapd.

Francis, L. (2005). Singapore–China "Open Skies" has Restriction on LCCs. *Air Transport Intelligence News,* 2 December. Retrieved on 12 October 2006 from Air Transport Intelligence at: http://www.rati.com/iata/oapd.

Gudmundsson, S.V. and Lechner, C. (2006). Multilateral Airline Alliances: Balancing Strategic Constraints and Opportunities. *Journal of Air Transport Management,* 12, pp. 153–58.

Iatrou, K. and Skourias, N. (2005). An Attempt to Measure the Traffic Impact of Airline Alliances. *Journal of Air Transportation,* 10(3). Retrieved on 12 October 2006 from Proquest at http://www.proquest.com.

ICAO (2004). *Database of the World's Air Services Agreements.*

Kinnock, N. (1996). The Liberalization of the European Aviation Industry. *European Business Journal,* 8(4). Retrieved on 10 October 2006 from Proquest at: http://www.proquest.com.

McDonald, M. (2006). When to Tie the Knot. *Air Transport World,* August. Retrieved on 12 October 2006 from Proquest at: http://www.proquest.com.

Mecham, M. (2004). Fueling Star. *Aviation Week & Space Technology,* 161(17). Retrieved on 18 October 2006 from Proquest at http.www.proquest.com.

Montreal Convention (1999). *Convention for the Unification of Certain Rule for International Carriage by Air, Montreal 28 May 1999.* Retrieved on 3 October 2006 from: http://www.jus.uio.no/lm/air.carriage.unification.convention.montreal.1999/.

oneworld (2006). http://www.oneworld.com.

Oum, T.H., Park, J.H. and Zhang, A. (2000). *Globalization and Strategic Alliances: The Case of the Airline Industry.* Oxford: Elsevier.

Oum, T.H. and Yamaguchi, K. (2006). Asia's Tangled Skies. *Far Eastern Economic Review,* January–February. Retrieved on 11 October 2006 from Proquest at: http://www.proquest.com.

Phillips, D. (2006). 'Open Skies' Reality Still Proves Elusive. *International Herald Tribune,* 4 June. Retrieved on 28 September 2006 from: http://www.iht.com/bin/print_ipub.php?file=/articles/2006/06/04/news/ravsky.php.

Robyn, D. (2003). The Benefits of Open Skies. *Airline Business,* March 2003. Retrieved on 28 November 2006 from Air Transport Intelligence at: http://www.rati.com/iata/oapd.

Star Alliance (2006). Star Alliance, Facts and Figures. Available at: http://www.staralliance.com/en/press/facts_figures/index.html.

Straus, B. (2006). Delta Buys UA's JFK–London Route, Minus Heathrow. *ATW Daily News,* 31 July 2006. Retrieved on 6 October 2006 from: http://www.atwonline.com/news/story.html?storyID=5892&print=Y.

Thompson, J. (2004). Come Together. *Airline Business,* November. Retrieved on 18 October 2006 from Air Transport Intelligence at: http://www.rati.com/iata/oapd.

US Centennial of Flight Commission (USCOF) (2006). *International Civil Aviation.* Retrieved on 3 October 2006 from: http://www.centennialofflight.gov/essay/Government_Role/Intl_Civil/POL19.htm.

Warsaw Convention (1929). *Convention for the Unification of Certain Rules Relating to International Carriage by Air, Signed at Warsaw on 12 October 1929.* Retrieved on 3 October 2006 from: http://www.jus.uio.no/lm/air.carriage.warsaw.convention.1929/doc.html.

World Airline Report (2005). *Air Transport World,* July.

8

Market Structure and Monopolistic Markets

My grandfather once told me that there are two kinds of people: those who work and those who take the credit. He told me to try to be in the first group; there was less competition there.

Indira Gandhi

Chapters 8 and 9 deal with market structure. Market structure refers to the competitive environment that surrounds the firm. It generally can be described in terms such as barriers to entry, numbers of buyers and sellers competing in the market, the individual seller's control over price, extent of product substitutability, and the degree of mutual interdependence between firms. Market structure determines the type of competition that is found in the industry. The topics covered in this chapter are as follows:

- Perfect competition, including:
  - Conditions of perfect competition
  - Perfect competition in the short run
  - Perfect competition in the long run
- Monopoly, including:
  - Legal and government barriers
  - Capital requirements
  - Technology
  - Natural barriers
  - Labor unions
  - Airports
- Price/output decision for monopolies
- Monopoly pricing and consumer well-being
- Market structure in the aviation industry
  - Aircraft manufacturing
  - Jet engine manufacturing.

Table 8.1 depicts the market continuum, and displays the four principal market structures. Perfect competition occurs when there are many buyers and sellers who have very little or no control over price. At the other extreme of the continuum are monopolies.

Table 8.1 Market continuum

	Perfect Competition	Monopolistic Competition	Oligopoly	Monopoly
Number of Sellers	Large	Many	Few	One
Type of Product	Homogenous	Unique	Homogenous or differentiated	Unique
Control over Price	None	Very Little	Good	Very Good
Entry Condition	Very Easy	Easy	Difficult	Impossible
Example	Agriculture	Retail	Airlines	Public Utilities

Here, the market contains only one seller who has almost complete control over the output or price. These two market structures are the focus of this chapter. Monopolistic competition and oligopolies, otherwise known as hybrid market structures, are the focus of Chapter 9.

PERFECT COMPETITION

A perfectly competitive industry is one in which there are a large number of small buyers and sellers who can enter and exit the industry with no restrictions. This creates a situation in which the individual firm has little or no power over the price of their good. The price is dictated by the market, and this makes the individual firms price-takers. On this basis, each firm has a simple decision: to sell at the market-bearing rate or not to sell at all.

Conditions of Perfect Competition

In order for perfect competition to exist, four conditions must be met. Since very few markets satisfy all four conditions exactly, examples of perfectly competitive markets are limited. Nonetheless, as many industries approximate these conditions, the idea of a perfectly competitive market is a useful construct when comparing market structures. The four conditions are:

- homogenous identical product
- many small buyers and sellers
- perfect dissemination of information
- very low barriers to entry.

The first condition for perfect competition is that the product must be relatively homogenous. This condition flows naturally from the idea that buyers of the product should feel that they are receiving essentially the same product (regardless of which seller they purchase from). The more heterogeneous the product becomes, the more sellers can take advantage of their market position. Therefore, we should expect—and, indeed, we see—sellers attempting to differentiate their product in any type of market. However, there are many types of markets, notably those for agricultural products, where it is difficult for sellers to differentiate their product. These markets are generally thought to be good examples of competitive markets.

The second condition of perfect competition is that many small buyers and sellers exist in the industry. Multiple buyers and sellers enable the market to dictate the price of the product through competition between buyers and sellers. As mentioned above, a good example of perfect competition is agriculture. In agriculture there are multiple individual producers and numerous buyers so that no one buyer or seller can control the market and artificially raise (or lower) the price of the commodity. A counterexample of an industry where sellers have strong influence over price is the commercial aircraft manufacturing industry. Because Boeing and Airbus are the only two major sellers of large commercial aircraft, they have a strong influence on price; it follows, therefore, that the commercial aircraft manufacturing industry is clearly not a perfectly competitive market.

The third condition of perfectly competitive markets is the generally available dissemination of information. In order for markets to react efficiently, information needs to be available to all buyers and sellers. If information is not available, then distortions in price may occur within the market. An example of a perfectly competitive market with full dissemination of information is the stock market. Due to the demands of investors and requirements set by the Securities Exchange Commission (SEC), companies disseminate information relating to the company's financial well-being. Using this information, investors buy or sell, and this adjusts the stock price, making the market efficient. Similar scenarios occur for other perfectly competitive markets, such as agriculture, where commodity markets provide readily available information on prices.

The final condition for perfect competition is that the barriers to enter the market must be low. In order for the market to act efficiently and adjust prices accordingly, firms must be able to easily enter and exit the industry. With low barriers to entry, the market can bear many small sellers. Although barriers to entry can vary by market, common barriers are financial requirements, economies of scale, market power, customer loyalty, technology, and government regulations. For example, the barriers to entering the airline industry are thought to be quite large. Airlines generally require huge capital investments, economies of scale, economies of scope, and strong customer loyalty in order to be profitable. These massive barriers to entry are the primary reason why almost all start-up airlines since deregulation have failed. These issues are discussed in more depth in Chapter 12 on low-cost airlines.

Perfect Competition in the Short Run

In the short run, in a perfectly competitive market, firms are considered as price-takers. Individual firms are unable to affect the price for their product, and must accept the market price. Since the market determines the price for the product, the market price is set at the point where market demand equals market supply. This is displayed graphically in Figure 8.1, where the equilibrium point is the market clearing price.

The equilibrium point between market supply and demand determines the market clearing price, but individual firms perceive this price as fixed for all quantities simply

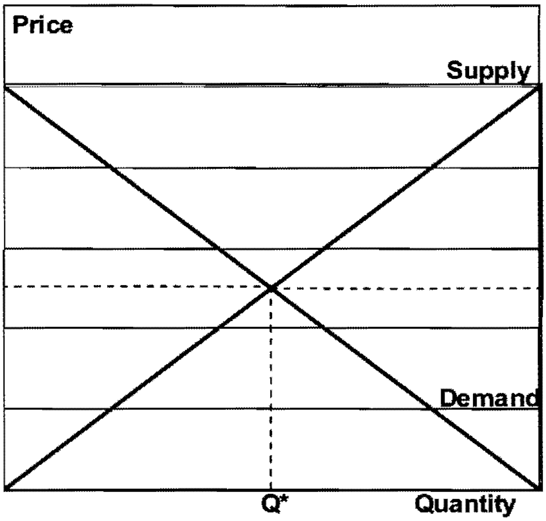

Figure 8.1 Market equilibrium: industry

because individual firms do not produce a sufficient quantity of the homogeneous product to affect the market price. Because of this, price is displayed as a horizontal line in Figure 8.2, which equates to the equilibrium point determined in Figure 8.1. Furthermore, since the price is constant for all quantities, price also equals average revenue and marginal revenue. Using a hypothetical example, Table 8.2 displays how price equals average revenue and marginal revenue for price-takers.[1] The horizontal price line in Figure 8.2 can be considered the demand for a firm in a perfectly competitive industry.

Since price is held constant in the short run, a firm's output decision is to produce where marginal revenue equals marginal cost. Given that the price line is also the marginal revenue curve, a firm's optimal output (Q^* in Figure 8.2) is where the marginal revenue line intersects the marginal cost curve, which is point B in Figure 8.2. At this optimal level of output, the firm is producing Q^* goods at price P. When the firm's average revenue is compared to the firm's average cost, the difference between points B and C in Figure 8.2

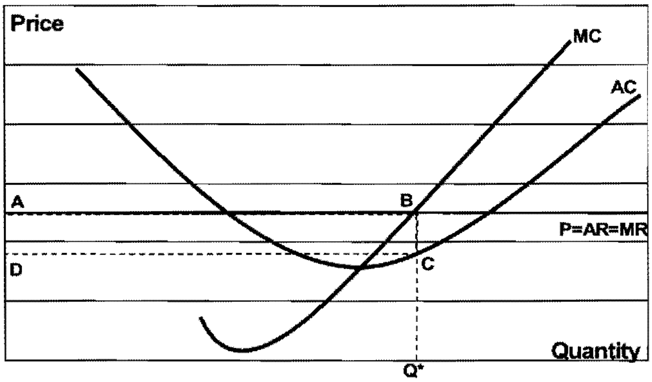

Figure 8.2 Market equilibrium: firm

1 The formulas and explanations of total, average, and marginal revenue were all contained in Chapter 4.

Table 8.2 Total revenue, average revenue and marginal revenue

Quantity	Price	Total Revenue	Average Revenue	Marginal Revenue
0	$100	$0	-	-
10	$100	$1,000	$100	$100
20	$100	$2,000	$100	$100
30	$100	$3,000	$100	$100
40	$100	$4,000	$100	$100
50	$100	$5,000	$100	$100

represent the contribution margin of the firm. Therefore, the shaded box ABCD represents the total profit for the firm in a perfectly competitive market.

However, the average cost curve may vary from firm to firm. For example, the firm in Figure 8.3 has a much higher cost structure, and, instead of making a profit, the company is losing money, which is represented by the shaded region ABCD. Since the firm is a price-taker, it has two options. The first is to lower its cost structure to the point at which it can actually make a profit or the second is simply to shut down. Since a firm can only continue to produce if its revenues exceed its variable costs (that is, it must meet its wage and supply bills), then the firm must shut down if it cannot cover its average variable costs. In the short run, average variable costs do not include sunk costs since these must be paid regardless of output level.

Perfect Competition in the Long Run

The situations presented above highlight the price/output decision for firms competing in a perfectly competitive market in the short run. However, markets are not static, but continually evolving. Since the barriers to enter and exit in perfectly competitive markets are low, new firms will enter the market when the industry is profitable and firms will exit

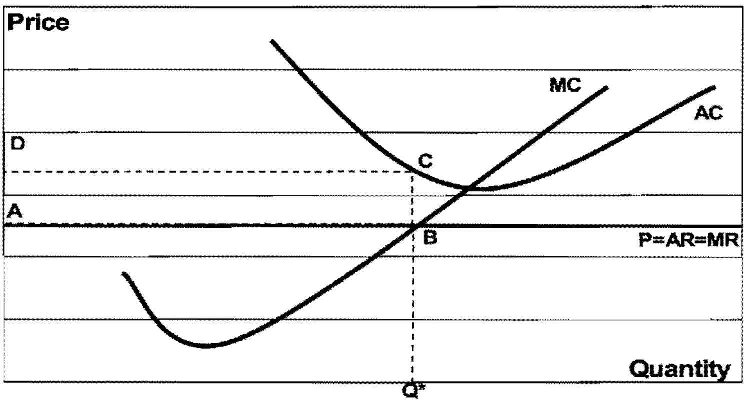

Figure 8.3 Short-run loss in perfect competition

the market when the industry is sustaining losses. So, it is not the high number of firms that keep profits low but, rather, the entry of new firms.

To see the effects that market entry and exit have on firms in a perfectly competitive market, consider the situation where the firm is making an economic profit, or where the intersection between marginal revenue and marginal cost is greater than average total cost.[2] Since firms are making a short-run profit, new firms easily enter the market (due to low barriers to entry) in hopes of also obtaining a profit. New entry ultimately increases supply in the market, causing a shift in the supply curve, which is represented in Figure 8.4. The shift in the supply curve also creates a new market clearing price, which all firms in the industry must accept.

Over time as more and more firms enter the market, the supply curve will continue shifting to the right. Each movement of the supply curve increases market supply, and

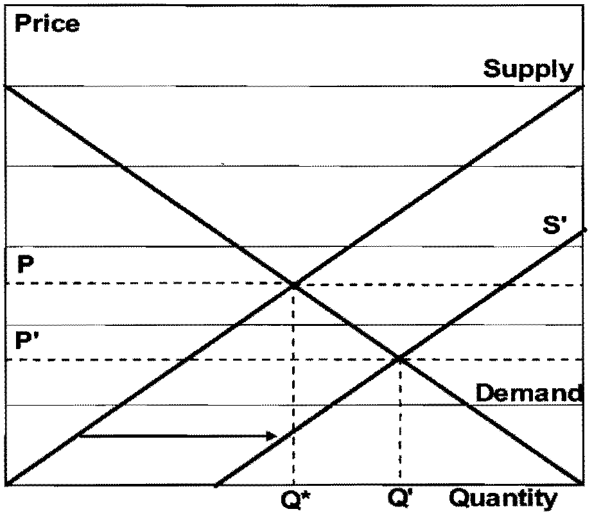

Figure 8.4 Industry level

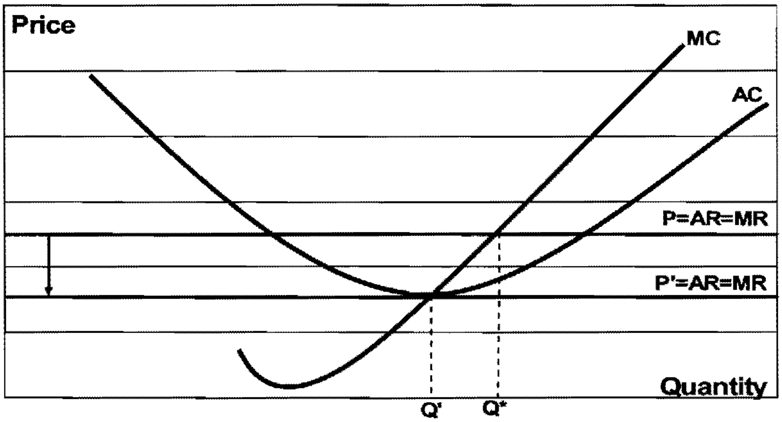

Figure 8.5 Firm level

2 Recall that, by definition, the average total cost curve contains a normal rate of return (profit) on investment. Anything above this is called an above average rate of return or economic profit.

also decreases both the price and the quantity demanded for an individual firm. In order to maximize profits in the face of decreasing demand, the firm reduces output. These movements and market reactions occur until marginal revenue equals marginal cost and average cost for individual firms in the market. At this point, the economic profit for the individual firm equals zero. Figure 8.5 displays the change in market output and individual firm output from a short-run scenario to a long-run scenario.

On the basis of these market adjustments, firms in perfectly competitive markets have zero economic profits in the long run. Since this is a long-run phenomenon, the market will continuously adjust in response to firms freely entering and exiting the market.

Market adjustments do not occur in just one direction. If some firms are losing money in the industry, then these firms will exit the market, causing a leftward shift in the market supply curve. This reduces the total quantity supplied by the market, but also increases the quantity demand (and the price) for an individual firm. Over time, the individual firm will return, to the point where marginal revenue equals marginal cost and average cost and economic profit (or loss) equals zero. Figures 8.6 and 8.7 display the market reaction when firms that are losing money exit a competitive industry.

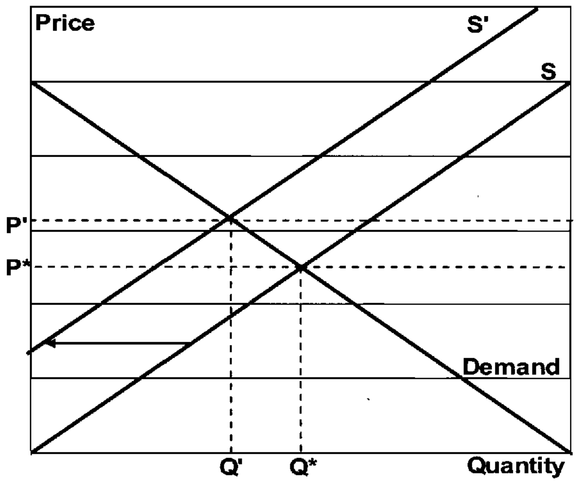

Figure 8.6 **Industry adjustment**

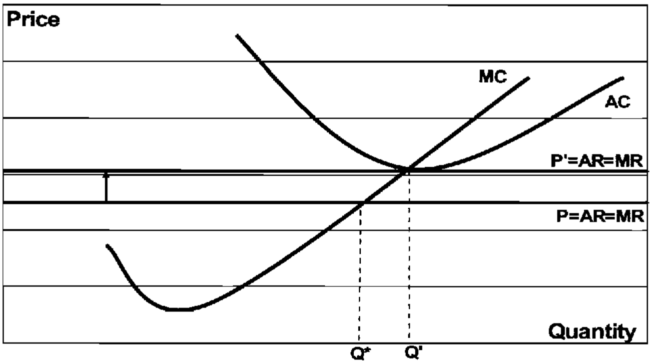

Figure 8.7 **Firm adjustment**

In order to change this situation, firms want to be able to differentiate their products so that they will gain some market power. Because of this, there are very few markets where all the conditions of perfect competition hold, but there are numerous markets where the perfect competition model described above is a good approximation of reality.

MONOPOLY

At the opposite end of the market continuum are monopolies. By definition, a monopoly is a market where there is only one seller. However, monopolies are all judged with respect to a certain market or geographical area. For example, the only paper manufacturer in a small town may be considered to have a monopoly over paper manufacturing with respect to the town, yet when the geographical scope is widened to include other towns, the county, or the state there are invariably other paper manufacturers competing in the market. Thus, the market structure for paper manufacturing would be considered a monopoly with respect to just the small town, whereas the market would be an oligopoly on a state level. This means that, depending on the level of analysis, a monopoly can exist in many different ways. Since monopolies can exist in industries, especially when the term is narrowly defined, it is important to understand their economic principles.

Monopolies are largely created by the existence of barriers to entry in a given industry. As mentioned, a barrier to entry is any obstacle that makes it unprofitable or impossible for new firms to enter an industry. However, there is some disagreement as to which barriers are significant. Economists agree that government can and does erect barriers that seriously harm the economy. The common government prohibition of cabotage, for example, obviously eliminates foreign competition and allows domestic air carriers to keep prices somewhat higher; incumbent carriers benefit at the expense of consumers, and the industry is prevented from being as large and efficient as it would otherwise be. The issue, discussed more fully in the Chapter 9, is whether a barrier to entry outside of government can seriously harm the economy. Some commonly mentioned barriers in the aviation industry are listed below and explained in the following paragraphs:

- legal and government barriers
- capital requirements
- technology
- natural barriers
- labor unions
- structure of airports.

Legal/Government Barriers

A major barrier to entry, particularly in international markets, is legal or government restrictions. Government regulation can help prevent market access, creating situations where an artificial monopoly may be created. One of the most common legal barriers to entry is patents and copyrights. When a patent or copyright is held, the holder of the patent has the exclusive legal right to sell the product. At first glance, it might seem that these are inherently anti-competitive. When a company obtains a patent on say, an aircraft design, this prevents anyone else from manufacturing that same aircraft, increasing the likelihood that the firm will enjoy some degree of monopoly power and be able to keep prices higher. In the absence of the patent, another firm might well start producing the same aircraft and drive prices down substantially. However, this simplistic, static analysis is misleading. From a dynamic, long-run prospective, this monopoly power serves to motivate far more research and development of new aircraft, thereby creating more competition and a wider variety of products.

For example, suppose there had been no patent laws back when Boeing was planning the development of the 747. Boeing would have known that once the 747 was developed and had met necessary regulatory approval, a competitor could have bought one and, by simply copying the design, could have eventually created copies while incurring only a fraction of the development costs that Boeing itself had incurred. Knowing this, Boeing might well have abandoned the risky project, and the 747 wouldn't exist. Thus, in this case, the patent prevents there being too much competition in the short run, so that more new products can be developed over time: more monopoly power in the short run increases output and competition in the long run. The pharmaceutical industry also has extensive experience with monopolies; new drugs receive legal protection, enabling the company to have a monopoly over the drug. As in the example above, however, an argument can be made that patent and copyright monopolies are justified as an incentive for research and development—in other words, without patents, pharmaceutical companies would have much less incentive to spend significant resources on the development of new drugs.

The aviation industry also has extensive experience with government restrictions creating barriers to entry. During the US regulation of air transportation, the Civil Aeronautics Board (CAB) dictated which airlines were to fly specific routes. In many cases, the CAB created monopolies on many individual routes. Since deregulation, domestic markets no longer have extensive government barriers to entry; however, international markets are still heavily regulated, and this creates a situation where government restrictions still provide substantial barriers to entry. For example, a bilateral air service agreement may restrict access into a particular market, creating a situation where a monopoly is created on a specific route. The topic of aviation monopolies will be covered in more detail later in the chapter.

Capital Requirements

Another possible barrier to entry in any industry is the capital required to enter the market. The capital necessary to commence production may be sufficiently large so that the potential profits do not justify the investment, the risk is too large, or the capital cannot be obtained. Firms only enter a market when they feel that they can obtain a reasonable rate of return. Because of these factors, the larger the capital required to enter a new market, the smaller the number of firms. For example, the commercial aircraft manufacturing industry requires very large capital requirements for new entrants. A new aircraft manufacturing company would call for sizeable capital for production facilities, research and development, and general overhead expenses. The capital requirements are the main reason why extremely few firms enter aircraft manufacturing. Conversely, the capital requirements for a restaurant or small retail store are considerably less than aircraft manufacturing; restaurants are therefore far more numerous. An aviation application of capital requirements as a barrier to entry is the project costs for the Airbus 380 project which. had an estimated $10.7 billion price tag for development costs—a sizeable sum of money that prevents other firms from entering the very large commercial aircraft market (Kjelgaard, 2002).

Airlines also require tremendous capital and physical assets to enter the commercial aviation market. While start-up airlines and start-up general aviation manufacturers often complain about the problems they have raising capital, it should be emphasized that this does not necessarily reflect any inefficiency in capital markets. Profits rarely come easy in any industry and seem to be particularly elusive for most airlines. Sometimes "no" is the efficient answer to entrepreneurs long on enthusiasm but short on viable business plans. On the other hand, capital is clearly accessible to those who do have a persuasive business plan. For instance, JetBlue obtained $130 million in start-up investment, making the airline the most heavily financed start-up in US airline history (Kjelgaard, 2000).

Related to capital requirements are the prospects of profitability in the industry. If the industry has narrow profit margins in addition to large capital requirements, then it is even less likely that new firms will enter that industry. Of course, if profits are not high there is no social need for new entry.

Technology

Depending on the industry, technology can be a substantial barrier to entry. Without a certain required level of technology, firms may be unable to compete effectively in a market. This is especially true in high-technology markets, where new technology drives sales. For example, in the microprocessor industry an entering firm requires a substantial level of technology to provide a product that might compete with Intel and AMD. In addition, the required technology is not a one-time occurrence, but must be continually upgraded in order to keep up with the industry as a whole.

While technology gains do not create sustainable monopolies, they can create monopolies for a period of time. For over 30 years Boeing held a monopoly in the very large commercial aircraft industry with its 747 aircraft; although, as discussed in the following chapter, Boeing's profits seem to be about normal. This monopoly was finally broken when Airbus launched its double-decker A380 aircraft. Although Boeing's 747 monopoly was a result of several factors, the required level of technology played a significant role in preventing other companies (until Airbus) from creating a very large commercial aircraft. Another aviation example is Aerospatiale's Concorde, which remains the only supersonic passenger aircraft to undergo commercial production.[3] Before its demise, the Concorde had a monopoly, though not a very profitable one, over supersonic commercial aircraft through a technological advantage that other companies could either not replicate or replicate efficiently. In this case, the technology for supersonic commercial aircraft represented a significant barrier to entry.[4]

3 Tupolev did create the TU-144, a similar supersonic commercial aircraft, but it did not go into widespread production.

4 On the other hand, many economists would argue there is no problem in any of this. Concorde did not succeed financially—the time savings were apparently not great enough to motivate enough travelers to pay for the higher costs; regular jet service across the Atlantic turned out to be a very viable substitute. Perhaps Boeing and others could have readily mastered the technology, but had the good business sense to choose to stay out of this market.

Natural Barriers

Natural monopolies can occur in the market due to economies of scale. In industries that have very high fixed costs, significant economies of scale can be achieved as production increases. For example, there usually is an extremely high fixed cost in constructing a hydroelectric power plant, but once the plant is constructed, the cost of generating extra electrical power is very low. This means that, if there is a competing hydroelectric power plant, each of the plants can lower their prices down to the marginal cost of producing electricity. As neither plant can recover their fixed costs under this sort of a pricing arrangement, one or both of the plants must go out of business in the longer term. In the late nineteenth century, railroads were subject to this sort of situation (sometimes called "ruinous competition"). Extremely high fixed costs were entailed in acquiring the land and constructing the railroad, but the costs of adding the extra cars and engines to carry more freight were relatively low. Therefore, competing railroads could lower their prices to just cover their variable costs (so they could still operate), but could not cover the fixed costs that they had incurred in building the line. The result was a predictable bankruptcy for one of the railroads. In this situation the bankrupt railroad was usually acquired by the competing line, and a monopoly was the final outcome. Any prospective new entrant would be faced with the prospect of requiring extensive capital costs and faced with the prospects of losing money so the monopoly was likely to last for some time.

Labor Unions

A final barrier to entry, particularly in the aviation industry, can be labor unions. Labor unions essentially band workers together to bargain as a monopolist of labor supply, and can thereby raise members' wages above the competitive level.[5] This monopoly power stems mainly from supportive government regulation, sometimes supplemented by direct government subsidy—for example, the US government prohibits employers from requiring new employees to contractually agree to not join a union, and generally limits the efforts firms can make to avoid or expel unions. Labor unions can have significant power in bargaining relationships, and this can increase the barriers to entry or completely restrict entry into certain markets. Depending on the contract negotiations, conditions may be imposed that make it difficult or unprofitable for an airline to enter a specific market.

Labor unions can also increase the barriers to entry in airline markets through scope agreements—contracts with labor groups that dictate various requirements, such as the size of aircraft that regional airlines can fly. For example, a scope agreement that requires all aircraft with more than 51 seats to be flown by mainline pilots, as opposed to cheaper regional affiliates, increases the barriers to enter the 70-seat regional jet market. Scope agreements can also exert monopoly power in terms of the number of aircraft that an affiliate can fly. For example, US Airways launched discount carrier MetroJet in 1998; however, the pilot contract limited MetroJet's operation to 25 per cent of US Airways,

5 Note that when unions succeed, at least temporarily, in forcing wages above the competitive level, this more expensive labor cost will tend to ultimately reduce employment in the union sector. In turn, workers who can't get jobs in the union sector will enter non-union markets and depress wages there. Thus, unions push wages down in some sectors even as they raise them in others; the net impact on wages tends to be about zero. Thus, it is a myth that unions tend to be a significant cause of inflation.

thereby restricting the number of aircraft that MetroJet could use to roughly 100 aircraft (Daly, 1998). This agreement restricted MetroJet's operations creating an artificial barrier to entry.

More broadly, unions may restrict entry into new markets by indirectly restricting capital inflows. Though unions arguably produce some benefits, they are viewed negatively by investors who see them as tending to depress rates of return. Imagine, for example, the likely surge in stock prices and borrowing prospects for legacy carriers if laws were changed so that airline strikes and other union work actions became illegal. The fact that such a law does not exist often makes it much harder for unionized firms to raise capital.

Airports

Another aspect of the aviation industry where market structure plays a significant role in shaping the industry is airports. Once again, the major issues concerning airports are the significant barriers to entry. The lack of available gates and slot controls at some airports serves to protect incumbent carriers and make it difficult for new entry. And at hub airports, it can allow an airline to exert monopoly power.

For an airline looking to serve a new airport, one of the first requirements is the availability of gates and ticket counters. Without the availability of both, the airline will not be able to service the airport. Although these requirements are universal, the ease of obtaining them varies considerably worldwide.

In the United States, the majority of gates and ticket counters are closely held by the individual airlines either through lease agreements with the airport or under full ownership. Leases with the airport provide the airline with exclusive control over a set number of gates, enabling the airline to utilize the facilities as they see fit. In a practical sense, the airline acts as the owner of the gates so that the airline can update its facilities as it sees fit. In return for this control over airport assets, airlines must make lease payments that are prescribed in a lease agreement with the airport authority. Depending on the terms of the lease agreement, airlines may be stuck with the facilities unless they can find another user for them. In some instances, airlines have found it profitable to sublet these facilities to competing airlines; in other instances, airlines have successfully terminated the lease agreements under Chapter 11 bankruptcy,[6] although this is a rather drastic remedy.

At hub airports airlines may own facilities. For example, Northwest Airlines invested heavily in their WorldGateway hub at Detroit International Airport, while Continental Airlines owns Terminal E at Houston Intercontinental Airport. Under these scenarios airlines may invest a large sum of capital in the airport in order to construct or revamp an entire terminal.

Regardless of the technical structure, US airlines have significant control over airport facilities, and this control represents a significant barrier to entry as any new airline will have to acquire airport facilities in order to commence service. At airports where there are idle facilities, this usually does not pose a significant problem to the new entrant, but at airports with scarce resources, the reverse is true. While the airport authority will

6 In the United States, Chapter 11 bankruptcy proceedings allow a firm to continue operating (as opposed to an immediate shutdown) under court supervision. This supervision continues until the firm is able to work its way out of the adverse financial situation. If this does not happen within a specified time, then the firm goes out of business.

attempt to work with all the airlines to accommodate the new entrant airline, this may not always work if all the resources are already committed. The other option for the new entrant would be to enter into an agreement with an incumbent airline to either use or sub-lease their facilities. This option is usually costly, and there may be some instances where, in order to suppress competition, incumbent airlines may be unwilling to allow a new entrant to the airport. This is particularly true at hub airports where the hub carrier may go to extensive means to obtain gates in order to block new entrants, especially low-cost competitors.

Moreover, such a structure may not be economically efficient since airlines may sign gate leases to block competition, even if they themselves do not require the gate. Even if leases contain minimum usage requirements, airlines can adjust their schedule or add extra flights to meet the minimum requirements. This creates a situation in which the assets may not be used in the most efficient manner.

Another solution to airport facility usage is common use facilities, implemented by the majority of airports outside of the United States and even by a few airports within the country. These are facilities that all airlines can use and are assigned by the airport daily, depending on demand, on a fee per use basis. In this way, use of facilities is maximized, and, in theory, because of this, the cost per use should be become lower. However, airlines dislike common use because it offers less control over the facilities and their operation. Common use facilities generally do not constitute significant barriers to entry, since, if facilities are available, an airline can use them. Even though facilities may be full at peak periods, an airline is unlikely to be unable to enter an airport using common use facilities at some time of the day.

However, whereas gates and other airport facilities rarely provide a significant barrier to entry outside of the United States, in Europe and other parts of the world, barriers to entry remain in the form of airport slots—namely, the rights to land or take off from an airport at given times. More formally, Article 1 of the European Council Regulation No. 95/93, the key regulation concerning airport slot allocation in Europe, defines a slot as:

> ...the entitlement established under this Regulation, of an air carrier to use the full range of airport infrastructure necessary to operate an air service at a coordinated airport on a specific date and time for the purpose of landing and take-off as allocated by a coordinator in accordance with this Regulation.
>
> (Commission of the European Communities, 2001)

Slots are allocated through a variety of mechanisms, but there are two primary rules to slot allocation according to EEC No. 95/93. The first rule, the "grandfather right," entitles an airline to the same slot in the future (if they are currently using it). Although this rule was enacted to provide stability, in practice it provides the airline with quasi-ownership of the slot. Slot usage is determined by the "use-it or lose-it" rule which states that the airline must use the slot for at least 80 per cent of the time during the scheduled period. If the airline fails to meet this requirement, the grandfather right does not apply and the slot is lost. This provides an incentive for the airline to continue using the slot, even if it is not economically efficient. The remaining slots are then dispersed to applicants, with only 50 per cent of new slots allocated to new entrants (Matthews and Menaz, 2003). As a result of the "grandfather right," the "use-it or lose-it" rule, and the new entrant slot limit, very few slots become available for new entrants. For example, in the summer of 2000, 97 per cent of London Heathrow's slots were grandfathered, leaving only 3 per cent of the total slots available, with only 1.5 per cent available to new entrants (DotEcon, 2001). Moreover, in all likelihood, the slots were available at inconvenient times. These slot controls represent a significant barrier to entry, and the grandfather rule enables an airline to accumulate a significant number of slots over time.

According to EC 95/93, the legal method of obtaining a slot in Europe is through the formal slot allocation process. However, although the directive permits the swapping of slots, a 1999 UK court ruling opened the door to the "grey market," where slot swapping is permitted with monetary compensation (Matthews and Menaz, 2003). This ruling created a situation in the UK where slots could be traded, leased, and sold to other airlines, although this practice remained illegal in the rest of Europe. The "grey market" is opposed throughout the rest of Europe under the rationale that a private firm should not benefit financially from a public good (Mackay, 2006). However, this rationale certainly favors incumbent airlines over new entrants into the market. While the "grey market" does reduce the barriers to entry, it also shifts the entry requirements from a legal/structural requirement to a financial requirement. From an economic point of view, financial barriers are more desirable than legal barriers since, under financial barriers, the slot is allocated to the firm that is willing to pay the highest price and therefore values the slot the most. With the recent open skies agreement signed between the European Union and the United States, the "grey market" will probably become more active, especially for slots at London Heathrow.

The United States has limited experience with slot controls, largely because it has been successful at constructing additional runways (Mackay, 2006). However, in 1969 the FAA implemented a slot system at four airports: New York LaGuardia, New York Kennedy, Chicago O'Hare, and Washington National. In 1985 the FAA permitted a full secondary market for slots, similar to the UK's "grey market," where slots could be sold, traded, or leased (Department of Justice, 2005). However, secondary slot trading was not as successful as anticipated due to two major issues:

- market power
- uncertainty of duration and value.

Airlines decided to retain slots, rather than sell/lease them, in order to prevent new carriers from entering the market. Market power also caused airlines to increase their slot ownership in order to become more dominant. This was especially true at Chicago O'Hare where United and American both increased their slot ownership, thereby making the airport less competitive. Finally, slots were only deemed temporary by the FAA, creating uncertainty over the lifetime and true value of the slot. (DOJ, 2005).

A lack of slot trading made the system relatively ineffective, and the slot mechanisms were eventually abolished at all four airports. As a result of the slot removal, airlines immediately increased service to the airports, especially at LaGuardia and O'Hare where tremendous delays were experienced (Mackay, 2006). This caused the FAA to mediate with the airlines to reduce the number of flights to acceptable levels. The immediate increase in service to these airports illustrates that a slot system represents a barrier to entry, even with a buy/sell market for slots.

Slots in Europe and airport facilities (gates, ticket counters) in the United States pose significant barriers to entry at the airport level. Furthermore, Dresner, Windle, and Yao (2002) have determined that high gate utilization during peak periods is the major airport barrier to entry in the United States. All these factors have allowed incumbent carriers to increase their market power at airports, particularly at hub airports. Indeed, the near monopoly power that some airlines have acquired at hub airports has been a significant barrier to entry at these airports.

Many studies have investigated the hypothesis that airlines have exerted their strong market power at the hub airport in the form of hub premiums which involve charging originating or terminating passengers a higher fare than other passengers traveling throughout the carrier's system (Gordon and Jenkins, 1999). However, opinions on whether hub premiums actually exist are mixed. Gordon and Jenkins (1999) used proprietary Northwest Airlines data to show that there was actually a hub discount at Minneapolis St Paul airport. However, some studies have found that hub premiums do not exist, and others have shown that they do (Borenstein, 1989). Lijesen, Rietveld, and Nijkamp (2004) claimed that a few carriers in Europe, specifically Lufthansa, Swiss, and Air France charged hub premiums. However, even if hub airlines do charge higher prices, this may simply reflect higher quality not adjusted for in such studies. For example, the dominant carrier may have more amenable airport facilities, may benefit from goodwill in the community where they are perceived as the "home-town company," may be perceived as having established greater credibility on safety, or have higher valued frequent-flyer awards. The fact that most businesses and communities seem to like having a hub airline nearby supports this view.

Since the evidence for a hub premium is not conclusive, and since hub airlines are often in bankruptcy, it stands to reason that, if any actual premium exists, it is likely to be small. Furthermore, even if there is some slight monopoly power at that airport, the positive network effects from hub-and-spoke systems also produce benefits that, as Economides (2004) shows, may be substantially greater than the damage of higher prices at the hub. In other words, the efficiency gained from economies of scale, scope and density in the hub network may help reduce prices and increase product availability in the overall network.

While specific studies appear to have reached differing conclusions on hub premiums, the actual case of Pittsburgh International Airport provides a good real-world example of what may happen when a hub airport is open to competition. In fall 2004, US Airways announced that it was going to significantly scale back its Pittsburgh hub. The total number of flights at Pittsburgh was dramatically reduced, and a greater number of gates became available. As the barriers to entry were lowered, low-cost carriers Southwest, JetBlue, and Independence Air all entered the market. Whereas the total number of passengers using the airport decreased, originating passengers increased by 12 per cent (McCartney, 2005). Moreover, airfares dropped significantly, especially on a few dominated routes. For example, the average airfare between Pittsburgh and Philadelphia (a route between two US Airways hubs) fell from $680 to $180 (McCartney, 2005).

PRICE/OUTPUT DECISION FOR MONOPOLIES

The first thing to note about a demand curve for a monopolist is that the market demand equals the firm's demand. This makes sense since, with only one firm competing in the market, the firm faces the entire market. The profit-maximizing output is the point where marginal revenue equals marginal cost. The reason for this is the fact that profit increases as long as the incremental revenue achieved from producing one additional unit is greater than the incremental production cost incurred to produce the additional unit. Figure 8.8 displays the profit maximization point for a monopolist.

In Figure 8.8 the monopolist would produce at Q^*, which is the point where marginal revenue equals marginal cost. At a production level of Q^*, the firm would have profits equal to the shaded area between the demand curve and the average cost curve. Since new market entry is not possible (by definition) in a monopoly, these profits can be enjoyed in the long run, *ceteris paribus*. However, simply having a monopoly does not guarantee long-run profits, as a combination of shifts in the demand curve and changes in average cost could put the firm in a loss-making situation. This scenario is presented in Figure 8.9

It should be noted that the profit-maximizing (or loss-minimizing) quantity for the firm represents the total supply to the industry. Without competition, the monopolist dominates the market so that the monopolistic firm effectively defines the supply side of the market, and this determines the price for the market.

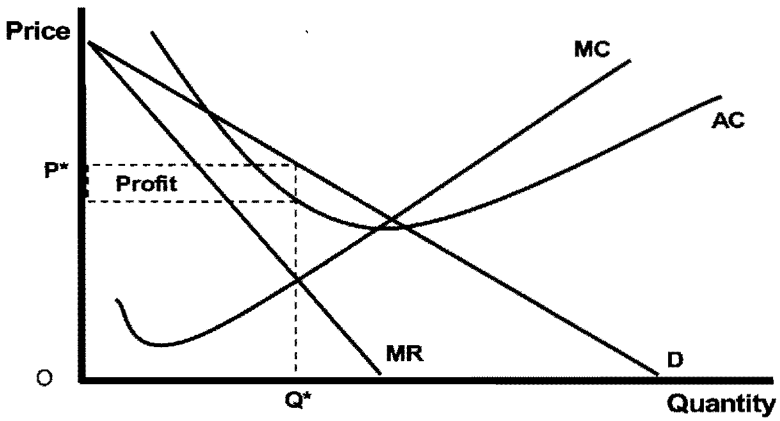

Figure 8.8 Profit maximization for a monopolist

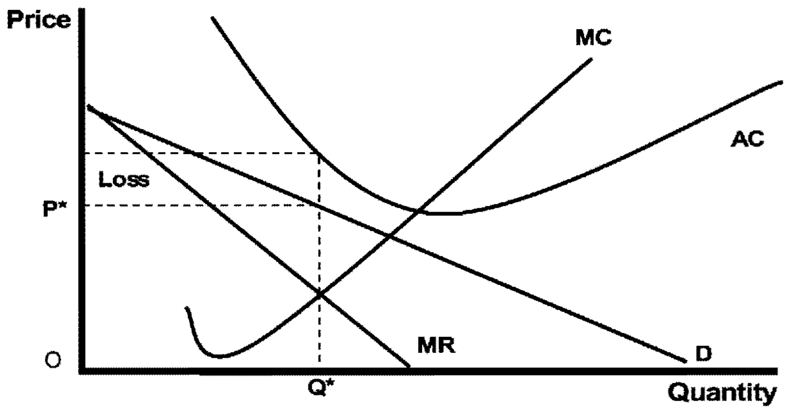

Figure 8.9 Loss scenario for a monopolist

In order to see the relationship between pricing under monopolies and other market structures, consider Figure 8.10.

As mentioned previously, monopolists will set their price where marginal revenue equals marginal cost. Therefore the corresponding price would be P_m in Figure 8.10. Once a monopoly is no longer available and competition in the market increases, the firm will move to a price/output decision where marginal cost equals demand. This point corresponds to P_c in Figure 8.10, which lies below P_m.[7] In addition, the quantity increases from Q_m to Q_c. Because of the increased competition, the price is lowered, which increases the quantity demanded. Finally, P_r corresponds to a price where government regulation requires cost-based pricing. Under cost-based pricing, the optimal output for the firm would be where average cost equals the demand curve. This point represents a further decrease in price and a further increase in the quantity demanded. Figure 8.10 graphically portrays why prices drop when new entrants enter a formerly held protective market.

The appendix to this chapter presents a formal mathematical derivation of profit maximization under monopoly. Here we simply state the formula derived in the appendix:

$$P = \frac{MC}{1 + \frac{1}{E_p}}$$

This formula provides the optimal price for a monopolist to maximize total profit. However, since price elasticity is rarely constant for all levels of output, the profit-maximizing profit will also vary by price elasticity. To illustrate this point, consider Table 8.3 which finds the profit-maximizing profit for different price elasticities, assuming that marginal cost is held constant at $1,000.

As Table 8.3 shows, the profit-maximizing price decreases as consumers become more price-elastic. Eventually, consumers may become so price-elastic that the monopolist's profit-maximizing price would be equal to the marginal cost. On this basis, cost plays a

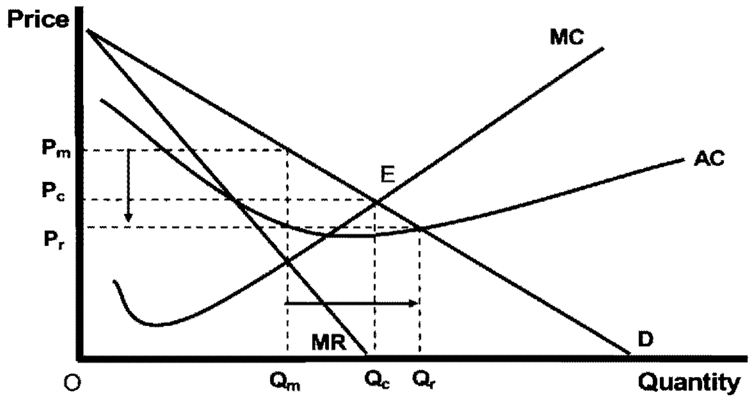

Figure 8.10 Equilibrium price under monopoly and perfect competition

7 All this assumes that the production costs are the same for the tiny competitor and the huge monopolist, and that no economies of scale exist. Alternatively, if there were significant economies of scale, it is very possible that the lower costs of the monopolist would translate into a price below the competitive level.

Table 8.3 Profit-maximizing price for various price elasticities

Marginal Cost	Price Elasticity	P
$1,000	-1.5	$3,000
$1,000	-2	$2,000
$1,000	-3	$1,500
$1,000	-4	$1,333
$1,000	∞	$1,000

significant role in shaping the price/output decision for a monopolist, and that decision varies because the elasticity of demand changes. This creates a situation where monopolists must still be reactive to the market and understand their consumers.

MONOPOLY PRICING AND CONSUMER WELL-BEING

To many, the pricing decision for a monopolist might seem simple: charge the highest price that the market will bear. However, while such a policy may provide short-run extraordinary profits, in the long run this practice will tend to reduce profitability. This type of pricing policy will cause demand for the product to fall as consumers find ways to adapt to substitute goods or otherwise change their habits to buy less of the monopolized product. Though there may be no perfect substitute for airline travel over long distances, it is easy to see how a monopoly airline could shift travelers over to corporate jets, chartered airlines, auto travel, train travel, or utilization of communication techniques such as video conferencing to minimize travel. Aggressive monopoly pricing will also cause other firms to invest more in overcoming barriers and entering the market. Monopoly pricing (as described above) basically pits one firm against the rest of the world, and the world tends to ultimately win in that sort of contest. Therefore, even a true monopolist protected by an iron-clad barrier, such as an effective patent, tends to moderate price somewhat in an attempt to discourage the consumer adjustments and market innovations that will eventually slash monopoly profits.

Economists point out another factor, overlooked by most people, that greatly mitigates the net harm of monopoly. If a monopolist restricts output and raises price in one market, this will simultaneously increase output and reduce price in other markets. If, for instance, a firm were to somehow monopolize the grapefruit market and raise price this would result, of course, in fewer grapefruit being sold and therefore fewer resources employed in producing grapefruit. Then, as more land, fertilizer, orchard workers, and so forth would now be available for orange production, there would be a surge in production there. The harm done by a monopoly in the grapefruit market would be substantially, though not completely, offset by cheaper and more abundant oranges. Likewise, any monopoly pricing in some air travel markets would shift aircraft and employees to other markets and reduce prices for consumers there. There would still be some net harm since this reallocation of resources is triggered by monopoly manipulation rather than consumer preference, but the point is that the net cost of any monopoly is far lower than most people realize.

MARKET STRUCTURE IN THE AVIATION INDUSTRY

Aircraft Manufacturing

Modern aircraft manufacturing is a heavily capital-intensive industry requiring immense expenditures in research, development, and manufacturing. As technology has increased and economies of scale benefits have become more important, the cost of designing and marketing an aircraft have become substantial, strengthening the barriers to entry into the industry. This trend has also caused a drastic reduction in the number of firms competing in the commercial aircraft industry, so that there are now very few firms competing in the industry, creating a market structure that resembles an oligopoly.

> The United States, at the beginning of 1960, had 12 commercial aircraft manufacturers. By 1980, following the exit of several firms from commercial aircraft production, only three remained: McDonnell Douglas, Boeing, and Lockheed. By 1981, Lockheed's L-1011 Tristar project had been a dismal failure, costing the company $2.5 billion over 13 years and forcing it to exit the market as well. Lockheed's exit left Boeing and McDonnell Douglas as the only two U.S commercial aircraft manufacturers.
>
> (Harrison, 2003)

Lockheed's L-1011 Tristar project is a prime example of the risk involved with aircraft manufacturing. The Tristar project experienced delays that reduced its competitiveness against the similar McDonnell Douglas DC10; these delays ultimately doomed the project. Although Lockheed is still a defense contractor, the financial failure of the L-1011 caused the company to exit the commercial aircraft manufacturing market, leaving only two (at that time) large US commercial aircraft manufacturers.

The remaining two companies, Boeing and McDonnell Douglas, pursued separate strategic business paths. Boeing developed new aircraft types, while McDonnell Douglas focused on redesigning existing aircraft types. The McDonnell Douglas strategy was not successful in the commercial aircraft market, and their US market share fell below 20 per cent in 1993 (Harrison, 2003). In 1996 Boeing announced a $13 billion merger with McDonnell Douglas. This left only one US commercial aircraft manufacturing firm and created a duopoly in the global market for large commercial aircraft (Harrison, 2003). A similar trend occurred in Europe where BAE and Fokker exited the commercial aircraft manufacturing business. This global trend is largely a result of the enormous cost of doing business in the market.

Today there are four major aircraft manufacturers, with two competing fairly evenly in each market. Boeing and Airbus compete in the large commercial aircraft market, comprising aircraft with over 100 seats, while Bombardier and Embraer compete in the regional jet market. The firms compete aggressively with each other in almost all segments of their individual markets. Until recently, the biggest exception to this rule was the very large aircraft market, in which Boeing had a monopoly with its 747. However, with the introduction of the Airbus A380, Boeing will no longer continue to enjoy a monopoly in that market.

The primary reason for Boeing's long-held monopoly of the very large aircraft market was the very high barriers to entry in this market. Not only were the financial requirements immense, but the technological and manufacturing requirements were also huge. In 1965, when Boeing decided to develop the 747, the projected launch costs were $1.5 billion (Esty, 2001). The project was widely viewed as a "daring, bet-the-company gamble on an untested product" (Esty, 2001). In addition to the large capital requirements, Boeing needed new technological breakthroughs, and it also had to construct an entirely new manufacturing complex in Everett, Washington. And, there were times when the project appeared to be a failure:

> Boeing's problems with the 747 sounds like a litany of the damned ... (and almost) threatened the company's survival...Boeing not only had to pay penalty fees for late deliveries, but, far worse, didn't receive the large last installments until the deliveries were made. Deprived of an adequate...cash flow, Boeing found itself seriously short of funds yet obliged to finance a huge inventory of partly build 747s
>
> (Newhouse, 1982)

The tremendous project risk, coupled with the other extensive technical and capital requirements, make it difficult to enter the very large aircraft market, especially since a new entrant will not have a monopoly and must compete with Boeing. However, in 2000, Airbus announced its intention to develop a very large double-decker aircraft to compete with the Boeing 747 in the very large aircraft market. The project was estimated to cost Airbus $13 billion, yet with subsequent delays and problems that figure has climbed to $14–15 billion (Esty, 2001). Furthermore, the original break-even forecast of 270 aircraft has climbed to 420 (BBC News, 2006). Although it is a risky project, the new competition in the very large aircraft market will undoubtedly cause greater competition between Boeing and Airbus. Examples of competitive actions already utilized in other commercial aircraft market segments include discount prices, attractive financing, and the purchase of older aircraft.

Very few firms, or countries, have the resources to design and manufacture commercial aircraft, and this is a primary reason why manufacturing aircraft has become something of a global enterprise For example, Boeing's 787 Dreamliner has development costs between $8–10 billion and utilizes partners from multiple countries, especially Japan (Kotha, Nolan, and Condit, 2005). Another example is Boeing's 777, which was launched in 1990 with an estimated development cost of $4–5 billion (Gollish, Clausen, Koggersvol, Christey, and Bruner, 1997).

In 2007 China announced its intention to start manufacturing large commercial aircraft by 2020, providing a possible new entrant to the market (Associated Press, 2007). Although the barriers to entry are immense, China may be able to overcome them by having a robust economy with a positive trade balance, a strong domestic economy, and an abundance of labor. China is expected to buy 2,230 new aircraft before 2025, so the demand is clearly there; however, the required technological skills will still be a major barrier for the Chinese (Associated Press, 2007). In an attempt to overcome this barrier, China reached an agreement with Airbus to open an A320 final assembly line (Associated Press, 2007) In the hope that the technological skills gained from this venture will translate into success for its large aircraft program. So while the barriers to enter the market are extremely high, they are not impenetrable.

Whereas the broad market for commercial aircraft is largely a duopoly, for some airlines the market is more like a monopoly. Thus, while most airlines are able to play one manufacturer against the other in negotiations, some are more constrained by operational limitations. For example, consider Southwest Airlines, an all-Boeing 737 operator. Part of Southwest's success has been due to having a common fleet; this has increased crew flexibility, reduced crew training cost, and reduced spare part inventories, to name just a few benefits. Under these circumstances, switching to an Airbus aircraft would entail substantial costs. This may put Boeing in a position where it can exert a degree of monopoly power over Southwest Airlines, although this is greatly tempered by the resale market for aircraft—Southwest could buy new aircraft indirectly through a third-party airline that was better positioned to strike bargains with Boeing, then the third party could immediately resell to Southwest.

Furthermore, Boeing must also consider the possibly severe damage to their reputation if it ever were somehow able to exploit such situations. If it is perceived to have betrayed one of its best and most loyal customers, then other airlines take care to avoid repeating Southwest's mistake. It is likely that any possible gain from exploiting Southwest would be far less than the loss from declining sales as the world learned that it was a grave mistake to become too dependent on Boeing. Indeed, since most companies are likely to be hesitant about relying too much on a single supplier, it may be vital for Boeing to be able to point to a very satisfied, successful "dependent" in order to encourage other airlines to follow Southwest's example.[8]

Jet Engine Manufacturing

An industry with a market structure similar to the commercial aircraft manufacturing industry is the commercial aircraft jet engine manufacturing industry. Not only are the products of the industries highly related, but they exhibit similar barriers to entry: high capital requirements, economies of scale, and advanced technological knowledge and skills. Unlike commercial aircraft manufacturing where there are four major firms competing in two distinct market segments, engine manufacturing is an oligopoly of three, or the "Big Three": General Electric (GE), Rolls-Royce, and Pratt & Whitney (P&W).

In addition to the "Big Three," two consortiums were formed in order to add more players to the engine market. CFM International (CFMI) was a partnership between General Electric and Snecma, the French state-owned engine manufacturer. International Aero Engines (IAE) was formed in 1983 between Pratt & Whitney, Rolls-Royce, Daimler-Benz, Fiat, and Japan Aero Engines. These consortiums not only pooled technological knowledge, reduced risk, and lowered production and development costs for individual manufacturers, but also created an interesting situation where manufacturers could be both partners and competitors concurrently. In addition to the two consortiums, P&W formed an alliance with GE for the development of the GP7200 platform, an engine designed for very large commercial aircraft such as the A380 (Bowen and Purrington, 2006).

Figure 8.11 displays the worldwide market share for commercial jet engines in terms of deliveries for the past 45 years. Up until the early 1980s, Pratt & Whitney was the dominant jet engine manufacturer. However, at that point the company a principal McDonnell-Douglas supplier, made a strategic decision that shaped the market for many years—namely, to focus on supplying engines for Boeing's new 757 instead of the 737, in the belief that the 757 was going to be the aircraft of the future (Bowen and Purrington,

8 Also, if Boeing bargained too aggressively, then Southwest could ultimately switch over. This scenario may have already occurred when easyJet placed a large order for Airbus aircraft, having previously been a consistent Boeing customer.

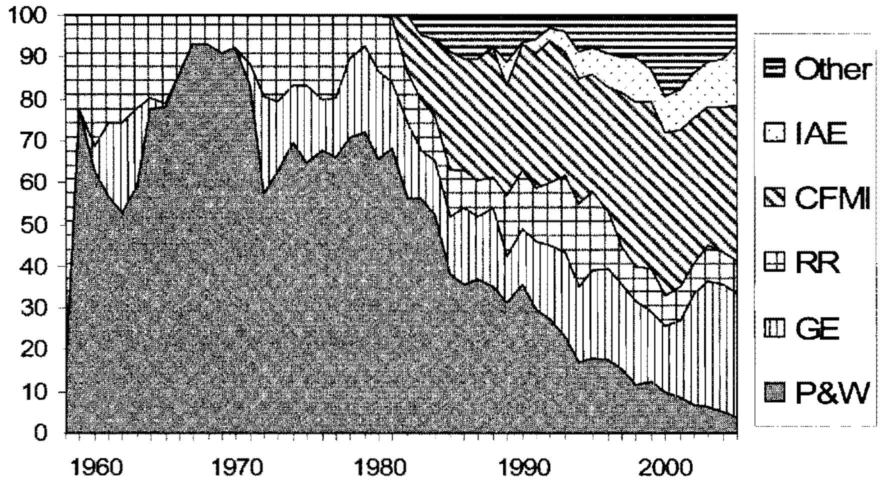

Figure 8.11 Commercial jet engine manufacturing market share by engine deliveries

Source: Compiled by the authors using the Airline Monitor (2006).

2006). This created a situation where CFMI became the sole jet supplier for 737s, while P&W and Rolls-Royce split orders for the 757s. While the 757s sold well, the 737s became the most successful commercial aircraft in history. Pratt & Whitney lost even more market share with the demise of McDonnell-Douglas, delivering only 3 per cent of the global commercial jet engines in 2005—well below the 90 per cent market share they enjoyed in the late 1960s.

The main benefactor of Pratt & Whitney's declining market share has been General Electric. GE, a long-time military engine manufacturer, entered the commercial market with the backing of its large parent company (which includes aircraft lessor GECAS). GE strengthened its position in commercial engine manufacturing with its CFMI consortium with Snecma. When GE's and CFMI's market share are combined, General Electric had over a 65 per cent presence in the commercial jet engine manufacturing industry in 2005. The next closest competitor is the P&W/Rolls consortium, IAE, with 15 per cent.

Figure 8.12 provides another comparison between the major commercial engine manufacturers in terms of in-service engines. Figure 8.12 is lagged, by roughly an engine's life, from Figure 8.11. This lag is the principal reason why Pratt & Whitney still remains a market leader, as many of their older products, such as the JT8D remain in service on aircraft such as the Boeing 727 and MD-80 (Bowen and Purrington, 2006). Therefore, P&W can still turn substantial revenues through maintenance agreements and selling spare-part inventories. However, as the older aircraft are retired, P&W's market share has declined, while GE and CFMI have gained.

The technological knowledge and skills required in the commercial aircraft engine manufacturing industry are substantial, and are major barriers to entry. Moreover, this knowledge requirement is a primary motivator for the number of alliances and agreements generated in the industry, as resources and risk are spread across multiple firms.

While the market broadly has three major firms (plus the two major consortiums), just as in the aircraft manufacturing industry there may be certain situations where the market is, in reality, a duopoly or monopoly. For example, on the new Boeing 787 only

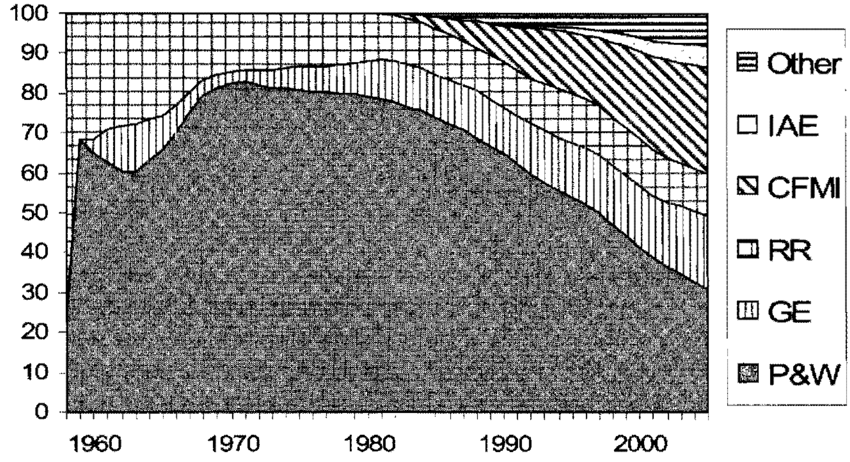

Figure 8.12 Commercial jet engine manufacturing market share by in-service engines

Source: Compiled by the authors using The Airline Monitor (2006).

two engines are being offered to customers: General Electric and Rolls-Royce. Therefore, the competitive actions of the firms will largely be based on duopoly competition theory. Furthermore, all Boeing 737NGs are powered by CFMI engines, which means that CFMI has a monopoly on the supply of engines for 737 aircraft. This creates an interesting dynamic as an engine's success is highly tied to an aircraft's success. Moreover, if CFMI wants to enforce monopoly power by attempting to raise the engine's price, this would hurt the competitiveness of the 737. However, with CFMI also being one of two suppliers for the rival Airbus 32X family, CFMI holds an extremely strong position in the narrow-body jet engine manufacturing industry.

Although engine manufacturers can hold some market power, airlines routinely play one engine manufacturer against the other in order to get the best deal. This means that engine manufacturers may have to provide deep discounts on engine purchases, but are able to recover costs through maintenance agreements or "power by the hour" contracts. Therefore, whenever possible, engine manufacturers capitalize on economies of scope benefits. This is especially true of General Electric which leases aircraft with GE engines through GE Commercial Aircraft Services (GECAS). Economies of scope represent another barrier to entry into the industry because they give incumbent firms a competitive advantage.

The capital requirements for engine manufacturing are quite large and, for the existing firms, have come mainly from earlier military contracts. The "Big Three" all have roots tracing back to military applications. Rolls-Royce and Pratt & Whitney were heavily involved with engine manufacturing during the Second World War, while General Electric made the jump from military applications in the late 1960s (Smith, 1997). Without military applications and grants, a new entrant would find the required technology extremely expensive to acquire. If a new aircraft project looked profitable, then undoubtedly one of the "Big Three" would become involved, leaving new entrants the less desirable aircraft designs.

The two industries discussed above—namely, aircraft manufacturing and jet engine manufacturing—are industries with firms that are capable of exerting substantial market power. The airline industry is somewhat unique in that its major suppliers all have substantial market power and, depending on circumstances, can exert near-monopoly power. On the other hand, the airlines compete in a fierce oligopoly where the customers have substantial power and where switching costs are low. The combination of these forces squeezes the airlines from both sides, creating a difficult market structure for the industry.

SUMMARY

This chapter has introduced the concept of market structures, and has covered not only the more theoretical economic models of a competitive market, but also those of a more monopolistic market. It has shown that although very few markets are perfect examples of competition or monopoly, the models are useful as benchmarks against which one can measure real-world market structures. The last part of the chapter comprised a more institutionalized discussion of the market structure of aircraft manufacturing and the manufacture of jet engines, both of which have certain monopolistic aspects.

APPENDIX: THE RELATIONSHIP BETWEEN MONOPOLY PRICING AND ELASTICITY OF DEMAND

Since monopolists are the only firm in the industry, their objective is to maximize total profit. Since profit is total revenues minus total costs, the optimal price for a monopolist can be determined mathematically. Total profit (π) can be stated as:

$$\pi = TR - TC$$
$$\pi = P(Q) * Q - C(Q)$$

Since price is not constant for all levels of demand, the formula above represents price as a function of quantity [P(Q)]. The price is then multiplied by Q to take into consideration the number of units sold by the firm. Cost is also represented as a function of quantity, although it is not multiplied by Q since the function would be total cost. In order to determine an optimum price, the first order derivative with respect to quantity leads to:

$$\frac{d\pi}{dQ} = P(Q) + \frac{dP(Q)}{dQ}Q - \frac{dC(Q)}{d(Q)}$$

Setting the above equation equal to zero yields:

$$P(Q) + \frac{dP(Q)}{dQ}Q - \frac{dC(Q)}{d(Q)} = 0$$
$$P(Q) + \frac{dP(Q)}{dQ}Q = \frac{dC(Q)}{d(Q)}$$

Since price and cost per additional unit is the definition of marginal revenue and marginal cost, the equation could be restated in those terms. However, in order to explain things further, the cost function will only be rewritten in terms of marginal cost. The rewritten formula is:

$$P(Q) + \frac{dP(Q)}{dQ} Q = MC$$

To help solve the equation, all terms are divided by P(Q):

$$\frac{P(Q)}{P(Q)} + \frac{dP(Q)}{dQ} \frac{Q}{P(Q)} = \frac{MC}{P(Q)}$$

The term $\frac{dP(Q)}{dQ} \frac{Q}{P(Q)}$ is the mathematical definition of the inverse of price elasticity; therefore the formula can be simplified into:

$$1 + \frac{1}{E_p} = \frac{MC}{P}$$

Solving for the optimal price in which a monopolist should price at provides:

$$P = \frac{MC}{1 + \frac{1}{E_p}}$$

If, for example, an airline is faced with a price elasticity of demand equaling -2 and the marginal cost of $100 per seat, the airline can mark up by 50 per cent and charge each additional seat up to $200.

REFERENCES

Associated Press (2007). China to Develop Large Commercial Aircraft by 2020. *International Herald Tribune*, 12 March. Retrieved on 28 March 2007 from: http://www.iht.com/articles/ap/2007/03/12/business/AS-FIN-China-Homegrown-Jet.php.

BBC News (2006). Airbus Hikes A380 Break-even Mark. Retrieved on 28 March 2007 from: http://news.bbc.co.uk/go/pr/fr/-/1/hi/business/6067540.stm.

Borenstein, S. (1989). Hubs and High Fares: Dominance and Market Power in the US Airline Industry. *Rand Journal of Economics*, 20, pp. 344–65.

Bowen, K. and Purrington, C. (2006). Pratt & Whitney: Engineering Standard Work. Harvard Case, 9-604-084, 27 March.

Commission of the European Communities. (2001). *Proposal for a Regulation of the European Parliament and of the Council Amending Council Regulation (EEC) No. 95/93 of 18 January 1993 on Common Rule for the Allocation of Slots at Community Airports.* Retrieved on 4 April 2007 from: http://www.ectaa.org/ECTAA%20English/Areas_dealt_with/en_501PC0335.pdf.

Daly, K. (1998). Winds Rise in the East. *Airline Business*, September. Retrieved on 12 March 2007 from Air Transport Intelligence at: http://www.rati.com/iata/oapd.

Department of Justice (DOJ) (2005). Congestion and Delay Reduction at Chicago O'Hare International Airport. *Comments of the United States Department of Justice.* Docket No. FAA-2005-20704. Retrieved on 5 April 2007 from: http://www.usdoj.gov/atr/public/comments/209455.pdf.

DotEcon (2001). Auctioning Airport Slots. *A Report for the HM Treasury and the Department of the Environment, Transport and the Regions.* Retrieved on 4 April 2007 from: http://www.dotecon.com/publications/slotauctr.pdf.

Dresner, M., Windle, R., & Yao, Y. (2002). Airport Barriers to Entry in the US. *Journal of Transport Economics and Policy*, 36, pp. 389–405.

Economides, N. (2004). Competition Policy in Network Industries: An Introduction. Stern.nyu.edu/networks/.

Esty, B. (2001). Airbus A3XX: Developing the World's Largest Commercial Jet (A). Harvard Case, 9-201-028, 24 August.

Gollish, D., Clausen, H., Koggersvol, N., Christey, P. and Bruner, R. (1997). The Boeing 777. Darden Case, UVA-F-1017, November.

Gordon, R. and Jenkins, D. (1999). *Hub and Network Pricing in the Northwest Airlines Domestic System.* The George Washington University.

Harrison, M. (2003). US versus EU Competition Policy: The Boeing-McDonnell Douglas Merger. *American Consortium on European Union Studies on Transatlantic Relations Cases*, 2.

Kjelgaard, C. (2000). JetBlue Holds $50m-plus Cash Despite Cost Over-run. *Air Transport Intelligence News*, 2 July. Retrieved on 9 March 2007 from Air Transport Intelligence at: http://www.rati.com/iata/oapd.

Kjelgaard, C. (2002). Airbus Expects A380 Break-even at under 250 Aircraft. *Air Transport Intelligence News*, 11 April. Retrieved on 9 March 2007 from Air Transport Intelligence at: http://www.rati.com/iata/oapd.

Kotha, O., Nolan, D. and Condit, M. (2005). Boeing 787: The Dreamliner. Harvard Case, 9-305-101, 21 June.

Lijesen, M., Rietveld, P., and Nijkamp, P. (2004). Do European Carriers Charge Hub Premiums? *Networks and Spatial Economics*, 4, pp. 347–60.

McCartney, S. (2005). The Middle Seat: Why Travelers Benefit When an Airline Hub Closes. *Wall Street Journal*, 1 November. Retrieved on 2 April 2007 from Proquest at: http://www.proquest.com.

Mackay, L. (2006). Overview of Mechanisms to Deal with Airport Congestion. Unpublished work. Embry-Riddle Aeronautical University.

Matthews, B. and Menaz, B. (2003). Airport Capacity: The Problem of Slot Allocation. Retrieved on 4 April 2007 from: http://www.garsonline.de/Downloads/Slot%20Market/031107-matthews.pdf.

Newhouse, J. (1982). *The Sporty Game.* New York: Alfred A. Knopf.

Smith, D. (1997). Strategic Alliances in the Aerospace Industry: A Case of Europe Emerging or Converging? *European Business Review*, 97(4), pp. 171–78. Retrieved from Emerald on 27 March 2007.

9

Hybrid Market Structure and the Aviation Industry

> United has little to fear from numerous small competitors. We should be able to compete effectively by advertising our size, dependability, and experience, and by matching or beating their promotional tactics .... In a free environment, we would be able to flex our marketing muscles a bit and should not fear the treat of being nibbled to death by little operators.
>
> Richard Ferris, CEO, United Airlines, 1976

While the previous chapter introduced the two extremes of the market structure continuum (see Table 9.1), this chapter will analyze the two middle hybrid market structures: monopolistic competition and oligopolies. Unlike perfect competition, which rarely exists in reality, both oligopolies and monopolistic competition are prevalent in modern industry, with the airline industry heavily affected by the characteristics of oligopolies. This chapter will analyze both hybrid market structures, and show how the market structure impacts on companies operating in this environment. The topics covered in the chapter are as follows:

- Monopolistic competition, including:
  - Price–output decision
- Oligopolies, including:
  - Differing views of oligopoly
  - High cost of capital
  - High exit barriers
- Examples of oligopoly
- Contestability theory
- Kinked demand curve theory
- Cournot theory
- Profitability issues
- Competition and antitrust issues, including:
  - Predatory pricing
  - Cartels and collusion
- Industry consolidation, including:
  - Four-firm concentration ratio
  - Herfindahl-Hirschman Index

- Beyond market concentration considerations
- Antitrust, market evolution, and cooperation.

MONOPOLISTIC COMPETITION

Monopolistic competition is, perhaps, the most common market structure of the four types displayed in the market continuum in Table 9.1. Monopolistic markets contain many sellers, but not quite to the degree of a perfectly competitive market. Consequently, each firm produces a small fraction of industry output; in other words, there exists little market consolidation. As with perfect competition, freedom of entry is reasonably easy. There are, of course, still entry costs, the key difference being that it is much easier to obtain capital in a monopolistically competitive market than in an oligopoly market. Other barriers to entry may include customer loyalty or regulatory restrictions. A monopolistically competitive industry has the following characteristics:

- a large number of sellers
- low barriers to entry
- product differentiation
- full dissemination of information.

Examples of monopolistic competition are:

- law firms
- accountancy firms
- bookstores
- convenience stores
- radio stations
- the restaurant industry
- clothing industry.

Table 9.1 Market continuum

	Perfect Competition	Monopolistic Competition	Oligopoly	Monopoly
Number of Sellers	Large	Many	Few	One
Type of Product	Homogenous	Unique	Homogenous or differentiated	Unique
Control over Price	None	Very Little	Good	Very Good
Entry Condition	Very Easy	Easy	Difficult	Impossible
Example	Agriculture	Retail	Airlines	Public Utilities

The key difference between perfect competition and monopolistic competition is that, in perfect competition, companies mainly sell homogeneous products while monopolistically competitive firms sell heterogeneous products with many close substitutes. Firms in a monopolistic competitive environment sell products that are more or less similar, and they have some power to set their own prices. From the firm's point of view this is a more desirable situation, since the more they can differentiate their product, the more control they have over the price. However, since all the firms have the same incentive, there is still a good deal of price competition in monopolistically competitive markets.

Monopolistic competition is extremely common in today's business environment. Restaurants are an example of monopolistic competition since they provide more or less similar services, are fairly common in most markets, and exist in a market with relative ease of entry. Bookstores, grocery stores, and pharmacies are all further examples of monopolistic competitive environments.

Price–Output Decision

Demand for a firm's product in perfectly competitive markets is essentially horizontal, since the firm is a price-taker and has virtually no power in setting prices. In this situation, demand is perfectly elastic. As product differentiation increases, the elasticity of demand decreases, which shifts the firm away from a perfectly competitive market and toward a more monopolistically competitive or oligopolistic market. As mentioned above, this increase in product differentiation creates greater potential for a firm to have more control over the price that it charges. This situation is depicted in Figure 9.1. In this figure, the demand curve for a firm shifts from a horizontal line to a downward-sloping line. As the firm progresses through the market structure continuum, the slope of the demand curve will become steeper until it eventually reaches a point where it encompasses the entire

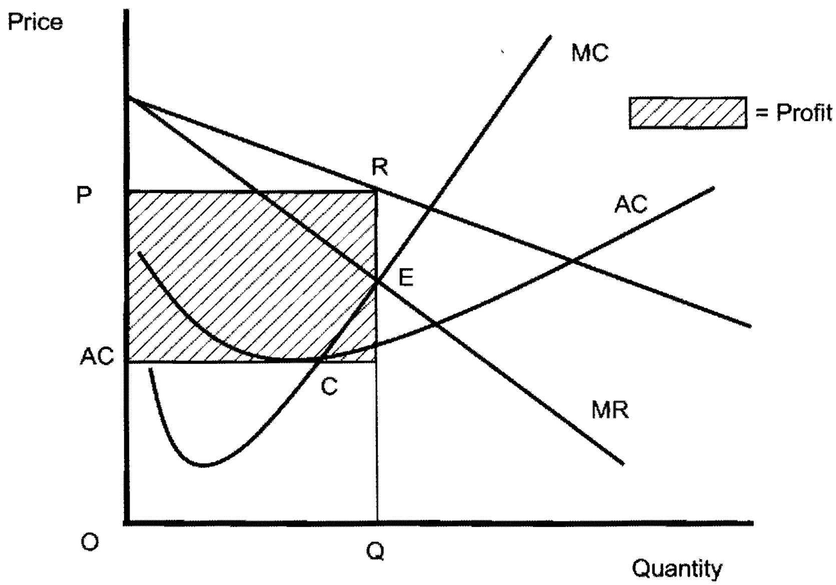

Figure 9.1 Short-run equilibrium: monopolistic competitive market

market, and it is at this point that the firm becomes a monopoly. Just as in a monopoly, in a monopolistically competitive environment, firms produce the quantity where marginal cost equals marginal revenue and charge the highest price that sells that quantity:

$$MC = MR$$
$$P > MR$$

The price they charge would be on the demand curve. As should be clear from the figure, the more elastic the demand curve facing the individual firm, the less control it has over its price. Figure 9.2 below illustrates typical demand curves across the various market structures.[1]

Since it is relatively easy to enter a monopolistically competitive market, we would expect that any supranormal profits that might be earned in the short run would be competed away in the longer term through the entry of new firms into the industry.

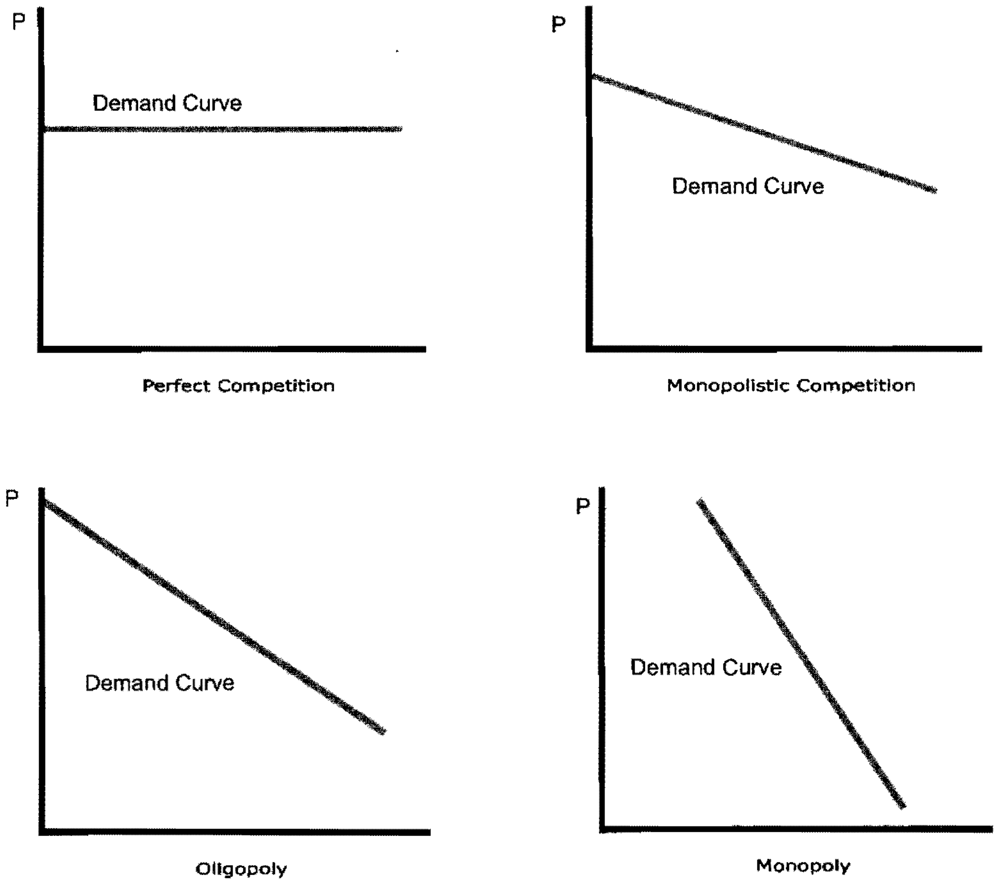

Figure 9.2 Progression of the demand curve: perfect competition to monopoly

1 Oligopoly is an especially complex case and the nature of demand can vary significantly given the particulars of each industry. Airlines, for instance, often seem to have no more control over price than the typical monopolistic competitor.

Figure 9.3 shows the longer-term equilibrium in a monopolistic competitive industry where a typical firm earns zero economic profit.

OLIGOPOLIES

The next step along the market continuum from monopolistic competition is oligopoly. Oligopoly is most relevant to aviation and will be the focus of the remainder of this chapter. Unlike firms in perfect competition or monopolistic competition, the oligopolist's actions will substantially affect the market. In turn, this may alter the actions of other competitors. This creates a complex interdependence amongst the firms; each firm's actions will be conditioned on how they believe the competition will react. For example, an airline might be more likely to reduce fares if it thought that competitors would leave their fares unchanged, but would prefer to leave fares constant if it believed that competitors would instantly match price cuts. Thus, each airline's pricing is based, in part, on what it believes competitors will do. Clearly, this is a complex problem, and one that often presents no clear, optimal strategy.

It follows, then, that almost any short-run outcome is theoretically possible in oligopoly. Firms might, for instance, sometimes practice "tacit collusion," where they keep prices relatively high and "go along to get along" by avoiding any aggressive competitive act that would lead to price wars. On the other hand, oligopoly can, as seems to often be the case for airlines, produce aggressive "cut-throat" competition where the typical firm is routinely operating in the red.

Whereas normal long-run profits are a given in monopolistic competition, it is theoretically possible for long-run profits to be above normal in oligopoly if there is a sufficiently insurmountable barrier to entry. Currently, though, many oligopolies, including such former paragons as General Motors, seem to struggle just to earn normal long-run profits. Indeed, most legacy airlines' long-run profits have been well below normal. Nevertheless, there are usually barriers to entry of some sort in oligopoly markets, so the number of competitors

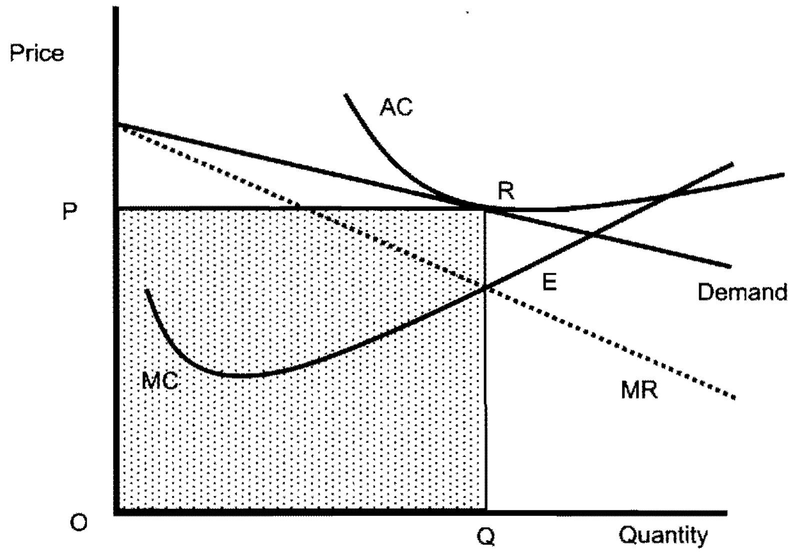

Figure 9.3 Longer-run equilibrium: monopolistic competitive market

is relatively small. These barriers include high start-up/fixed costs, the existence of sizeable economies of scale, control over scarce resources, or exclusive patent/legal rights.

Oligopoly is often viewed as inherently undesirable—better than monopoly, but not nearly as good as perfect competition. While most economists would probably agree that there is a certain degree of truth in this perspective there are some complications. A few large firms can often enjoy economies of scale and thereby produce at far lower costs and sell profitably at far lower prices than could an industry composed of smaller, more numerous firms. This is easy to see in the case of aircraft production. It is far cheaper for Boeing or Airbus to develop and produce 1,000 aircraft than it would be to have 100 small manufacturers develop and produce ten comparable aircraft each. Boeing or Airbus can spread research and development costs over more units, thus pricing them lower, and can benefit from the experience gained, becoming increasingly efficient with each additional aircraft produced. There is no doubt that airlines and air travelers are better served by having two, rather than 200, manufacturers of large aircraft. Economies of scale, scope and density are also important in the airline industry; so, again, we are probably better off with an airline oligopoly than any feasible alternative.

This does not mean, of course, that any movement to fewer, larger firms is automatically more efficient and good for consumers. Such industry consolidation might be beneficial, but it might also artificially suppress competition. So, how can we decide, say, when two competing airlines should be allowed to merge? Let us sketch two different views.

Differing Views of Oligopoly

There is a common presumption that any substantial increase in market concentration is generally undesirable and that oligopolistic competitors should therefore generally not be allowed to merge. The main exception is the situation where denying a merger cannot prevent increased market concentration because the weaker firm will simply liquidate if the merger is not approved. This was essentially the rationale in the case where the financially desperate McDonnell-Douglas was allowed to merge with Boeing. Underlying this staunch anti-merger view is the perception that efficiency gains from economies of scale are likely to be less significant than the increased danger of oligopoly abuse in increased pricing power and artificially high profits.

An important foundation for this view is the belief that new entry is often exceedingly difficult because many barriers to entry tend to be naturally powerful. The staunchest proponents of this view would even argue, for example, that advertising and brand-name recognition are potentially enough in themselves to seriously impede new entry and thereby allow oligopolies to enjoy high profits even in the long run (see, for example, Galbraith, 1979). An alternative view—call it the market process view—is that very few, if any, entry barriers, other than legal barriers erected by government, are significant. According to this theory, it is best to let the market evolve in whatever way firms choose. Efficiency gains are likely, and will generally be impossible for government regulators to estimate and predict, so it is best for them to merely stand aside and let the market process work. High profits, as in the perfect competition model, will always be short-lived as new firms will enter, and, existing firms will increase capacity, thus driving down prices and profits. Government can improve efficiency only by getting its own house in order—

eliminating international trade barriers and other government polices that seriously limit competition and harm the economy.

High Cost of Capital

Let us compare and contrast these differing views by considering the difficulties posed by the very high capital requirements for a new entrant into, for example, aircraft manufacturing. Any entrepreneur intending to start a company that would compete with Boeing and Airbus in the production of large aircraft would face quite a challenge. Most likely, he or she would need billions of dollars worth of specialized capital equipment to begin production. Since raising billions of dollars of capital isn't easy, this would, according to the traditional view, constitute a serious barrier to entry.

However, market process proponents would point out that modern capital markets have many trillions of dollars worth of assets; billions can be readily raised if one offers a persuasive business plan to investors. Many of the Internet start-ups of the 1990s demonstrated just how easy it can be to quickly raise billions of dollars if investors are excited about your prospects. At this point, it seems that two producers, Airbus and Boeing, are enough. If, somehow the industry grows sufficiently, or Airbus and Boeing somehow otherwise manage to enjoy high prices and profits, then Lockheed Martin, or some other firm, will enter their market. New entry is not kept at bay by high capital costs, according to this view, but by the fact that the industry seems to be sufficiently competitive with just two firms.

High Exit Barriers

Another potentially serious entry barrier is high exit barriers. Investors must consider worst-case scenarios—what happens to their investment in a company that performs so poorly that it must be liquidated, with its assets being sold off to the highest bidder? If our hypothetical aircraft manufacturer liquidates, it will face the problem of very limited resale markets. Indeed, much of its equipment might have only two possible buyers, Airbus and Boeing, and end up being sold as scrap metal if neither of those two is interested. Thus, any major investment in such illiquid assets may face unusually high risk. Theoretically, this risk might inhibit new entry and thereby allow Boeing and Airbus to enjoy higher than normal returns.

However, market process proponents reply, it is a standard principle of finance that riskier investments must offer higher expected profits. In other words, if Boeing ever does earn unusually high long-run profits, it can be reasonably argued that such profits reflect merely the greater risk—that is, the risk-adjusted rate of return would still be a normal rate of return. After all, their investors also face the risk of illiquid resale markets should Boeing ever fail. More broadly, since economies of scale are so important in this industry, and since the physical capital may be so specialized, it makes sense in terms of social welfare to be cautious before capital is plunged into aircraft production. In other words, because it is so hard to exit that industry, it is perfectly appropriate, in this view, to hesitate before plunging in. From society's perspective, returns should, arguably, be quite high, before new entry occurs.

EXAMPLES OF OLIGOPOLY

Oligopoly is very common in modern economies. When you go to buy a soft drink, you find that almost all of the soft drinks are produced by a relatively small number of producers: Coca-Cola or Pepsi Cola, and a few other firms produce a vast majority of soft drinks sold in the United States. Common examples of oligopolistic industries are:

- airlines
- airports
- automobile Industry
- breakfast cereal
- cigarettes
- long-distance telephone companies
- film and camera
- soft drinks
- supermarkets
- television cable companies.

Oligopolistic markets are of particular interest since most major aviation-related industries are oligopolies. The commercial aircraft manufacturing industry is largely a duopoly with Airbus and Boeing competing in the 100-seat plus aircraft category and Bombardier and Embraer competing in the regional aircraft market. Consolidation in aircraft manufacturing has occurred with the acquisition of McDonnell-Douglas by Boeing and the exit of Lockheed from the commercial aircraft industry. Since commercial aircraft manufacturing is extremely capital-intensive, it is unlikely that another manufacturer (except possibly a Chinese manufacturer) will enter the market, in the next 10–15 years.

In both aircraft manufacturing duopolies, the manufacturers offer similar products (that is,. 737 versus A320; CRJ versus ERJ) closely matched in price. The only major difference has been Airbus's insistence on a super-jumbo aircraft (A380) as opposed to Boeing's decision to focus on a smaller super-long-range aircraft (B-787). The B-787 aircraft is very fuel-efficient with a cruising speed of Mach 0.85, and, being smaller than the A380, it accesses regional airports with no problems. The aircraft, which carries about 280 passengers, also has a range that can extend to 8,500 nautical miles. Since both manufacturers price their aircraft almost identically, competition occurs in the area of additional services, such as financing agreements or agreed buy-back of older aircraft.

Oligopolistic market characteristics also apply to the aircraft manufacturer's suppliers. Rolls-Royce, GE, and Pratt & Whitney are the major engine manufacturers and service providers. Engine manufacturers also compete in service categories by creating the most fuel-efficient engines and by providing the most attractive "power-by-the-hour" contracts. Under a power-by-the-hour arrangement, the engine manufacturers provide fixed-cost maintenance based on the number of hours flown each year. The airlines provide a fixed level of funding, and expect to receive a given level of support by the engine manufacturers. The contractor expects to be provided a fixed level of funding upfront, and anticipates a long-term support arrangement. Similar conditions prevail for avionics, aircraft interiors, and in-flight entertainment systems.

Table 9.2 Large aircraft manufacturers' market share

Airline Manufacturers	Market % 2005	Market % 2006
Boeing	55	60
Airbus	45	40

The airline industry is clearly an oligopoly market as it only has a few firms participating in the typical city-pair market. While oligopoly market theory suggests that firms should compete on service, since price cuts can be so readily matched, the US domestic airline industry has totally reversed this trend as the airlines have cut costs and service amenities. This development partly stems from the fact that many non-price competition aspects can be easily copied by competing airlines, such as frequent-flyer programs, but the key point is probably that consumers are often driven mainly by price concerns. Airlines might prefer otherwise, but customer preferences seem to consistently force vigorous price competition. However, even an airline like Southwest Airlines, which has traditionally had a price leadership strategy, is also known for its friendly service, and provides a frequent flight schedule on many of its city-pairs. Southwest's awareness of service quality is one of the reasons for its success. Other airlines such as Emirates and Singapore, which pride themselves on their service quality, have also very successfully adopted this strategy.

CONTESTABILITY THEORY

Thinking back to the perfect competition model presented in Chapter 8, note that it is not the large number of competitors that reliably drives profits to normal level but, rather, the expansion of output and entry of new firms. Profits can temporarily be quite high regardless of how many multitudes of competitors there are in wheat farming or anything else. Likewise, having only a few competitors in a market in no way guarantees any prospect for high profits. If new competitors can readily enter a market, even a single firm may be driven to behave in a basically competitive manner, earning only normal profits on average, in order to discourage new entry.

Pure *contestability theory* takes this idea a step further, and posits that, in the absence of significant entry barriers, the number of firms in an industry is completely irrelevant. The key element of contestability theory is:

> That a market is vulnerable to competitive forces even when it is currently occupied by an oligopoly or a monopoly. That is, if any incumbent is inefficient or charges excessive prices or exploits consumers in any other way, successful entry must be possible and profitable.
>
> (Bailey, 1981, p. 179)

Before deregulation, many economists speculated that pure contestability theory might well apply to the airline industry. Since aircraft are inherently mobile, it was thought that they could, under certain conditions, readily be reallocated to whatever routes were commanding higher prices, thus driving those prices down. Since each airline knew this, they would refrain from significantly raising prices—in other words, potential competition would have the same effect as actual competition. However, several studies examined the airline industry and found a positive relationship between airfares and market concentration levels: the fewer the airlines in a given market the higher the fares on average, which suggests that airline markets are not perfectly contestable (Strassmann, 1990; Whinston and Collins, 1992; Oum, Zhang, and Zhang, 1993).

One possible explanation is that economists underestimated the cost of entering a new market. Suppose, for example, that an airline had service to airport A and to airport B but no nonstop service connecting A and B. It might seem that an aircraft could be reassigned to a new A–B route almost instantly—that the aircraft could fly out of one market and into another in pursuit of the greatest profit. But, in reality, a new route must be planned and announced to consumers well in advance, normally at least three months before, and probably some special spending on advertising will be necessary. These are not massive costs, but they may create enough friction to the entry process to prevent pure contestability results.[2]

However, there is considerable doubt that these slight entry costs can fully explain observed price variances. Network effects in the context of intense competition may offer a better explanation. An airline network is more than the sum of its separate routes, particularly for the legacy carriers that aspire to offer seamless travel to "almost anywhere." Suppose, for instance, that such a carrier found it necessary to operate a nonstop route to Las Vegas not only because it is such a popular vacation destination, but also because, among other things, many key customers preferred to redeem their frequent-flyer awards for a Las Vegas trip. In this case, the value of the Vegas route might far exceed the actual revenue garnered from paying customers on that particular flight. Thus, the airline would sensibly keep that route rather than reallocating the aircraft to another route—say, nonstop to Minneapolis—that would produce more direct revenue, but less total value and revenue for the network as a whole. Thus, the price of a flight to Minneapolis could remain higher than the price to Las Vegas even if the market were purely contestable. Roughly the same thing happens in grocery stores when a particular item, a "loss leader," is sold at an especially low price in order to bring customers into the store, who will then, hopefully, buy other items with higher mark-ups while they are there.

In this case we need to look at profits for the entire airline rather than the prices of particular city-pairs to judge the contestability of the industry. Since, as we shall see in more detail later, the long-run profits and rates of return for the airline industry are exceedingly low, it may well be that the industry is basically contestable in this broader sense.

Of course, it is possible to drown in a deep spot within a pond where *average* depth is only knee-deep. Likewise, in this network set-up there can still be considerable pain for those particular cities and city-pairs that face relatively high prices even though the system-wide average fare is an extraordinary consumer bargain, actually below the cost associated with a normal profit. Fortunately for consumers in those 'high end" markets, low-cost carriers with simpler point–to-point networks are entering these markets with increasing frequency. There have been numerous examples where carriers which abuse their market power have generated competition from other carriers. Low-cost carriers AirTran and Frontier created their own hubs in the Atlanta and Denver markets that were formerly dominated by oligopolistic legacy carriers. Virgin Atlantic evolved to provide British Airways with

2 One of the key elements of contestability theory is that entry and exit from markets must be free and easy (Bailey, 1981). The complete absence of barriers of entry would satisfy pure contestability theory, but in the airline industry there can be sizeable barriers of entry into a given airport. This is certainly the case at airports such as London Heathrow which is slot controlled, restricting competition. However, there are competing airports, and this makes contestability theory apply to a certain extent, even in this market. Ultimately, in oligopoly markets where there are low barriers to entry, any market power held by carriers will be much less than in markets where there are high barriers to entry.

legitimate competition on long-haul flights. Yet, although these examples illustrate an application of contestability theory on a widespread scale, contestability theory is probably most relevant in small markets where there may only be one or two airlines serving the market. In these situations airfares may remain somewhat high, but not exorbitantly so, since this would encourage competition in that market.

Figure 9.4 displays a hypothetical demand curve for a firm in an oligopoly market. Under the contestability theory, there are two components to the firm's demand curve. The first component occurs where the firm can increase prices and not lose a substantial amount of demand. This is shown as the solid portion of the right line in Figure 9.4.[3] Here, demand is relatively inelastic since there is not much competition in the market. This occurs up to a point where new firms are encouraged to enter the market because of higher profits. If the firms actually do enter the market, then this will, increase competition, and ultimately cause demand to become more elastic. This is shown as the solid portion of the left line in Figure 9.4.[4] The intersection of these two demand curves is marked by the dotted vertical line, and this indicates the price that will encourage new firms to enter the market. Ideally, a firm wishes to charge a price just below this kinked point, so that there is no incentive for new entrants into this market. The problem is that this entry point is not precisely known, varies from market to market, and may not be high enough to support normal profit for more than one firm.

KINKED DEMAND CURVE THEORY

The kinked demand curve is another theory that predicts price stability in oligopolistic markets. The theory states that in an oligopolistic market there exists a band where price stability exists. This band is the kinked portion of the demand curve.

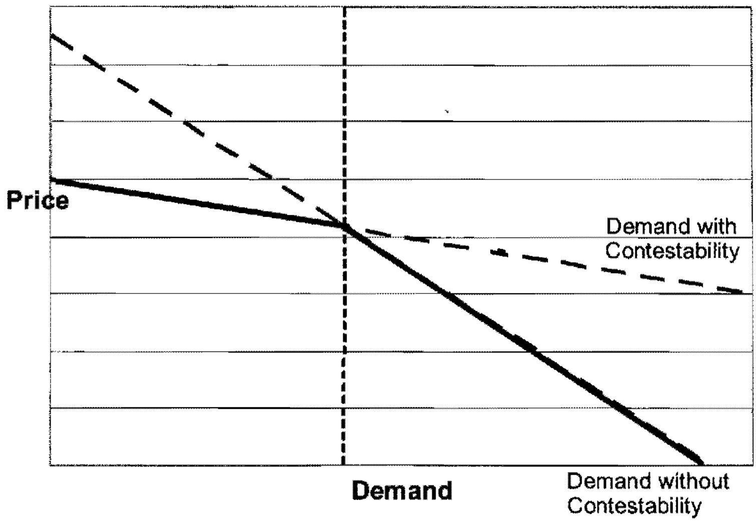

Figure 9.4 Price competition and market reaction

3 In October 2006 United Airlines raised fares in several markets, but when it became clear that airlines such as JetBlue and Northwest would not raise their fares, it rescinded the fare increases.

4 Other airlines may not follow a price increase by one airline; therefore demand will remain relatively elastic. Furthermore, an increase in price would not lead to an increase in the total revenue of the airline.

Consider a duopoly with two firms who have slightly different demand curves for their product. Figure 9.5 displays the two demand curves for firms, D(1) and D(2). These two demand curves intersect at point A on Figure 9.5; this point lies roughly above the $800 price level and the 320 demand level.

Both companies want to operate at a point where marginal revenue equals marginal cost. From this point, the optimum price can be found from the corresponding point along the demand curve. In a duopoly market, the firm whose prices are the lowest will provide the market price which the other firm will match in order to remain competitive.

The market demand curve can be constructed on the basis of this information. Up until point A in Figure 9.5, firm 1's demand curve is less than firm 2's; therefore, based on simple supply and demand, firm 1 would charge a lower price than firm 2. This situation is reversed after point A where firm 2's demand curve is significantly less than firm 1's. Therefore, the market demand curve will be the point $1000:0 to point A and from point A to $0:750.

In order to construct the market marginal revenue curve, firm 1's marginal revenue curve should be used up to a demand level of 320 units, and firm 2's marginal revenue curve should be used thereafter. This creates a situation where the market marginal revenue curve contains a vertical portion, line B-C. Since the intersection of the marginal revenue and marginal cost curve will yield the optimum price for the industry, when the marginal cost curve intersects between points B and C, the market clearing price would be roughly $820, or the price at point A. Therefore, if the marginal cost curve's intersection shifts anywhere in between points B and C, the market clearing price will remain the same. This creates a situation of long-run price stability in the market since there is a relatively wide range over which marginal cost can change (B to C on Figure 9.5) without changing the profit-maximizing price (A on Figure 9.5).

At times, the airline industry seems to provide some indication that this might be occurring. Only major shifts in the demand curves will create significant fluctuations in the market marginal revenue curves, thus changing the market clearing price. Such a major shift in demand occurred shortly after the terrorist attacks of 11 September 2001. In this case, demand shifted down, disrupting and altering previously stable equilibriums.

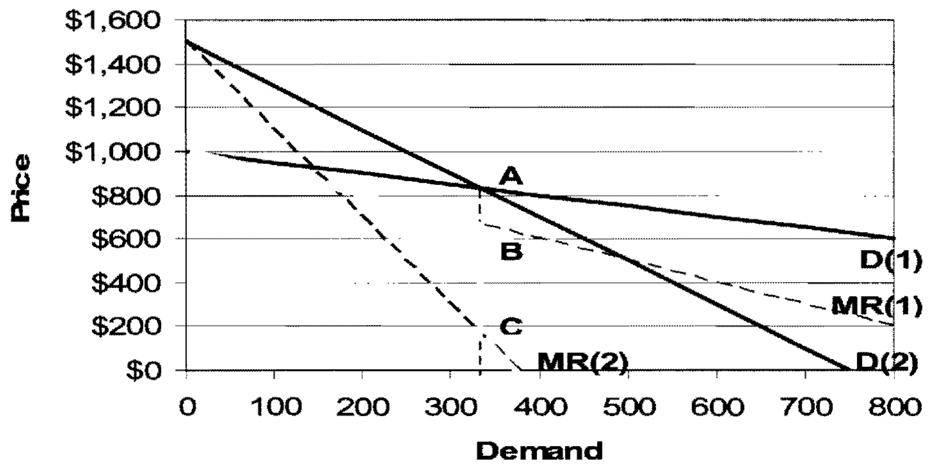

Figure 9.5 Kinked demand curve

COURNOT THEORY

Cournot theory helps explain competition and market equilibrium, based on firms competing through output decisions. The theory assumes that products are homogenous, market entry is difficult, firms have market power, cost structures are similar, and each firm assumes that the other will not respond to changes. To clarify this last point, the model assumes that, for instance, Boeing believes that Airbus will not respond to any changes in price and output initiated by Boeing. Likewise, Airbus assumes Boeing will also be completely unresponsive. Though this assumption is probably not realistic, the model still offers some insights and, given that Airbus and Boeing cannot be sure how the other will respond, the theory may sometimes approximate reality.

Consider a duopoly market where the firm's marginal costs are zero. The demand curve for the entire market is described in Figure 9.6, with the total output in the industry being Q. The first firm's marginal revenue curve is also described in Figure 9.5, and is exactly half the demand curve, since the demand curve is linear. The optimal output decision for the first firm is where marginal revenue equals marginal cost. Since marginal costs are assumed to be zero, the firm would want to produce at the point where the marginal revenue curve crosses the x-axis. This point is exactly half of the total demand, or point Q/2. The corresponding price for this level of output from the demand curve is P_1.

Since firm 1 takes half of the market for itself, the second firm's maximum demand would be Q/2. Therefore, the demand curve for the second firm is shifted to the left and intersects the x-axis at the point Q/2. This is displayed in Figure 9.7. Using this new demand curve, the second firm's optimal output level is where its marginal revenue curve intersects the x-axis. This occurs exactly at Q/4 and the corresponding price point is P_2. Based on this, the first firm would take half the market and the second firm would take a quarter of the market, meaning that the firms total market share is 3/4Q.

However, these firms have not reached equilibrium since the second firm's price is dramatically below the first firm's price. With this disparity in prices, most consumers would opt for firm 2's product over firm 1's, since the products are homogenous and the price is significantly lower. Based on this, the firms will readjust their output in an effort to

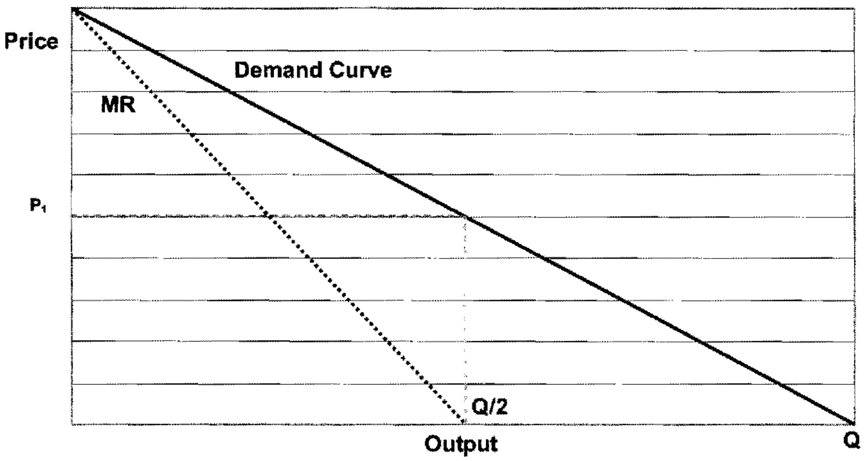

Figure 9.6 Initial output decision for the first firm

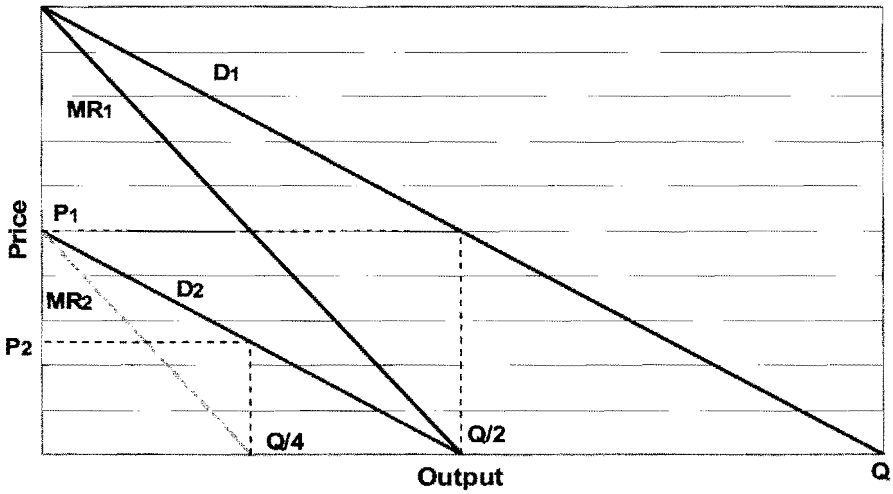

Figure 9.7 First-round Cournot theory

obtain equilibrium. These readjustments do not occur all at once, but over various rounds. Table 9.3 provides the various rounds and corresponding market output for the firms as they progress through their readjustments.

The Cournot solution is where both firms are in equilibrium, with the same price level and the same level of output. In a duopoly the Cournot solution would have both firms obtaining 1/3Q; therefore the total market share obtained by both firms would be 2/3Q. Note that the total market share declines as firms readjust to equilibrium, and that firm 1's market share declined while firm 2's market share increased. Cournot theory predicts that firms will continue to readjust the level of output until they have achieved market equilibrium at the same price level.

The characteristics of the aircraft manufacturing industry seem to resemble those of the Cournot theory. Although Boeing's and Airbus's products are not identical, they are fairly close in technical performance and requirements. The barriers to entry in the aircraft manufacturing industry are high, and each firm has tremendous market power. Through government aid, both firms have similar cost structures, but not identical ones. Both Airbus and Boeing have roughly equal market share, and the readjusting progression has been evident in past years as Boeing's market share has slowly declined and Airbus's market share has increased (although this situation has turned around in 2006 and 2007 as Boeing has regained market share). Differences between the two companies' output can be attributed to cost structure differences and/or product differentiation.

Table 9.3 Cournot theory progression

	Round 1	Round 2	Round 3	Round 4
Firm 1	1/2 Q	3/8 Q	11/32 Q	1/3 Q
Firm 2	1/4 Q	5/16 Q	21/64 Q	1/3 Q
Total Market	3/4 Q	11/16 Q	43/64 Q	2/3 Q

On the other hand, the Cournot solution is rarely achieved in the airline industry because airlines are able to differentiate their product (mainly through route structure and frequency of flights), each airline has a different cost structure, and each individual carrier has little market power. Consequently, the market share is rarely evenly distributed across multiple carriers. In fact, the development of the hub system has led to situations where one airline tends to dominate the market share at a hub airport.

PROFITABILITY ISSUES

Normal long-run profit levels might be explained by the industry being contestable in the manner discussed above, but contestability in itself should not produce such below-normal returns. The explanation may lie in the industry's oligopolistic nature, combined with very high fixed and very low marginal costs. Recall that, since an airline's schedule is usually fixed approximately three months in advance, most costs, even labor and fuel, are essentially fixed for that period. The marginal cost of placing a passenger in an otherwise empty seat on an aircraft that will be flying in any case (with or without that passenger) is extremely low—consisting mainly of the cost of processing their ticket or, in some cases, of paying a travel agent's commission. Thus, each individual airline is in a position where even a very low price is better than nothing for an otherwise empty seat. As will be discussed in Chapter 11 on the revenue management, each airline strives to make this low fare available only to those passengers who would not have flown at a higher price on *their own airline*. However, each airline is naturally happy to draw passengers away from a competitor. Suppose, for example, that Joe would have paid $200 to fly airline A but is lured into flying airline B for $150. Likewise, Jane would have flown airline B for $200 but is lured over to airline A for $150. Each individual airline, acting independently, is acting rationally, but the collective result is that each receives $50 less, perhaps incurring a loss rather than making a profit.

If each airline could refrain from "stealing" the other's customers, they might both enjoy a normal profit rather than being in bankruptcy. This is a classic *prisoner's dilemma.*[5] Since each airline can not "trust" the other to refrain from this "cut-throat pricing," they both do it, even though both would prefer to cooperate and stop such aggressive price competition. The reality is, if only one airline stops offering the $150 deal, that airline will lose both Jane and Joe to the competition. Of course, any attempt at cooperating to rein in this cut-throat competition is complicated by the fact that US government antitrust policy typically prohibits arranging such cooperation via a formal contract and makes it difficult, and probably prohibitively risky, to make an illegal arrangement in secret. Another complication is that other airlines will tend to enter the market with aggressive price cuts even if the two airlines do somehow manage to cooperate. US airlines also argue that the bankruptcy laws exacerbate the problem by providing a subsidy that keeps failed airlines from actually leaving the market, so that, even in the long run, it is difficult for the industry to decrease capacity sufficiently to keep prices high enough to support a normal profit.[6]

---

5 The term follows from the idea that two criminals might both go free if each lies to protect the other. But, with the police questioning them separately, each knows that he will face a very stiff sentence if he lies to protect his partner, but his partner then tells the truth. Unless each can somehow be pretty certain that the other will lie, they have an incentive to implicate each other in exchange for a lighter sentence.

6 For example, many airlines have effectively "dumped" the cost of their pension programs on to taxpayers

At first glance, the misery experienced by airline investors from pricing below costs appears to be a joyous gain for air travelers—as if most carriers were perpetually selling at below-cost, "going out of business" bargain rates. And that may, in fact, be the case. Conventionally, most economists have viewed excessively low long-run returns as a problem that will eventually take care of itself as needed. In other words, if many investors and lenders don't think it is worth investing in the airlines, they can stop financing them until capacity decreases enough to raise prices, normalize profits, and warrant future investment. Moreover, the strong performance by airlines like Ryanair and Southwest suggests that the business can be profitable if it's "done right." So, most economists are probably content to let the industry evolve as it will, even with rampant bankruptcies, until or unless there is clear evidence of a problem for consumers.

However, some economists, proponents of the "empty core" theory, suggest that there may already be a serious problem.[7] It is possible for the aforementioned cut-throat competition to be so severe that it actually does harm consumers by preventing airlines from offering some higher-priced products that consumers would prefer to have. For example, suppose there is no available nonstop service on a given route and that two competing airlines offer service, via a stop and transfer at their respective hubs, for a price of $200. Suppose that consumer demand is such that *one* airline could offer nonstop service on this route for $280 and, if the other competing airline maintained its same $200 service, both airlines would be financially viable in the market. In other words, suppose the market is capable of supporting both airlines, but only one with a nonstop flight. However, if competition leads either to the duplication of the nonstop service and/or a significant price cut, then the non-stop service becomes financially impossible to maintain. So, it is possible to offer consumers a chance to pay a premium and obtain the more desirable nonstop service only if there is no strong competitive response.

Suppose, though, that each time one airline adds a nonstop service, the other either duplicates that service or substantially slashes prices on its stop-and-transfer service, so that the nonstop service becomes a financial loser and is abandoned.[8] Moreover, since the two airlines compete in numerous markets, they each learn of this tendency and therefore choose never to start such nonstop service in similar markets. This problem might explain why airline customers complain so much about declining quality, while at the same time making choices that drive airlines to cut quality.[9] There is no practical way for an airline to contract with customers to get them to keep flying the new nonstop route after competitors respond, even though customers might be willing to do so if they understood that booking a bargain today would eventually result in poorer service/higher prices in the future.[10]

via government assumption of these pension obligations, though with some cuts for wealthier pensioners. Also, although this may be changing of late, the much higher cost structure of the legacy carriers also leads to increased capacity from low-cost carriers more than offsetting the capacity cuts by the legacies. Unusually powerful unions at the legacy carriers are also often cited as contributing to the legacy airlines' ongoing struggles.

7 See Raghavan and Raghavan (2005) for a discussion of an "empty core problem."

8 This situation explains, incidentally, the puzzling fact that a flight from A to B to C is sometimes cheaper than a flight from A to B.

9 Of course, it is also very possible that consumers really do want lower prices primarily and just enjoy complaining about quality because they unrealistically want extremely low prices *and* high quality!

10 To illustrate, suppose many customers' first choice is $150 with stop-and-transfer, their second choice is $280 nonstop, and their third choice is $200 with stop-and-transfer. Suppose airline D starts the nonstop $280 service, but then airline U offers the $150 service in response. This renders the nonstop service non-viable and it is abandoned, and the market returns to the $200 service by both airlines, leaving these customers with their least preferred option.

In such a case, consumers would be better served if the two airlines could freely negotiate a solution: perhaps one airline could be induced to not respond by receiving a small side payment, or the two airlines might make a trade where they take turns adding nonstop service in various more marginal markets. We now have the counterintuitive situation where consumers might actually be better served by an alliance that seems to closely resemble a cartel! However, this is in fact a less unique situation than it might at first appear. It has, for example, long been common practice, especially in manufacturing, for firms that are normally competitors to occasionally team up on particular projects. Similarly, an airline might conduct maintenance or baggage-handling for a competitor. One reason for this is that economies of scale for certain products may be such that only a single producer can be efficient enough to viably deliver the product to market. This relates to the fact we mentioned earlier that reducing the number of competitors may sometimes increase both efficiency and consumer well-being. If the airlines were allowed to act in concert, it is easy to see how, for instance, in a given market two large aircraft with 95 per cent load factors might have much lower costs per passenger, and therefore lower prices, than three smaller aircraft with 75 per cent load factors, operated by three separate airlines. Just as aircraft prices might be lower with only the two current producers than if we had several more, it may be that, at least in some cases, reducing competition on some routes could benefit consumers.

It is probably fair to say that most traditional economists would view allowing competing airlines to cooperate and form alliances as too radical a step, seeing the risk of price collusion conspiracy as outweighing potential gains. However, many other economists would probably be willing to consider it, especially on some limited experimental basis. Some would argue that, given the financial plight of legacy carriers, the risk of them colluding to make "too much profit" is not significant. Also, strong proponents of market process would argue that new entry, and perhaps potential new entry (contestability), could effectively restrain any harmful anti-competitive impulses. In other words, price conspiracies are unlikely and, even if the airlines in a given market attempted collusion, a new entrant would undercut them—that is, the market is contestable, at least in a basic, practical sense.

Ironically, in this situation staunch antitrust regulation may ultimately decrease the number of competitors. The fact that current regulation may prevent airlines from cooperating to improve efficiency in particular markets may result in more airline failures, eventual liquidation, and/or desperate mergers that antitrust enthusiasts have no choice but to accept. It is speculatively possible, for instance, that Lockheed or McDonnell-Douglas might have remained as competitors to Boeing and Airbus had they been allowed to cooperate on some projects.

When regulatory preferences conflict with the economic reality of substantial economies of scale, scope or density, economic reality will ultimately win. If more cooperative efforts between airlines are needed, they will eventually emerge, if not through approved alliances then through bankruptcy, liquidation, and mergers that reduce the number of independent airlines. Just as regulators could not stop the aircraft manufacturing market from evolving into a duopoly, with each firm able to enjoy considerable economies of scale, it may be that a similar process is unfolding for airlines. Of course, there is tremendous uncertainty about all of this. However, it does seem likely, given the financial disarray of many legacy carriers, that major changes of some sort will occur in the industry.

COMPETITION AND ANTITRUST ISSUES

Antitrust regulation has become very controversial. Virtually all economists see substantial problems with at least some aspects of antitrust regulation, and some even maintain that the costs of these regulations clearly exceed the benefits and that they should be completely abolished (Crandall and Winston, 2003; Armentano, 1986).

One difficulty is that the most problematic anti-competitive behavior is completely exempt from antitrust oversight. Economists would generally be thrilled if it were possible, for instance, to challenge international trade barriers under antitrust law, but all government policies are exempt from these laws. Since, as shown above, there is also some question about the real power of any barrier to entry outside of government, it follows that there is some debate as to whether there is enough of a private monopoly problem to justify a government regulatory program. Even if there are some imperfections in market competition, as most economists would probably agree, it is not easy for regulators to make things better for consumers by imposing fines and escalating legal costs on firms; after all, such costs tend ultimately to be passed on to consumers in the form of higher prices. There is also likely to be substantial bias on the part of regulators. If, for instance, regulators publicly admit that there are few problems in the marketplace, then they will probably experience significant budget cuts and may soon find themselves out of a job. If, on the other hand, they claim that there are not only monopoly problems, but also potential problems lurking everywhere, they tend to receive budget increases and greater chances of job security, promotion, and higher pay.

The problem is exacerbated by the fact that, as we shall see, truly anti-competitive behavior is usually very difficult to distinguish from healthy, competitive business practices. Indeed, the purpose of any business is to attract consumers and thereby *harm their competitors*. It is theoretically possible for a firm to harm its competitors too much, or in a certain way, such that it is ultimately harmful to the overall economy, but it is difficult—some would say impossible—for regulators to draw a clear line to divide appropriate competitive acts from inappropriate competitive acts. Let us consider some examples.

Predatory Pricing

Predatory pricing theoretically occurs if a firm: (1) cuts price intentionally low for the express purpose of driving competitors out of the market; then (2) raises prices to a monopoly level once the competition has gone. Note that aggressive price cuts, even if they drive competitors out of business, are not in themselves predatory. Airlines often find themselves losing money, fighting over a market that isn't big enough to sustain all existing firms. In this case, they may fight it out until some firms leave the market and prices rise high enough to support normal profits. Thus, aggressive price cuts and less efficient firms going out of business are the routine results of healthy competition. Only if prices go up to monopoly levels can the action be said to be predatory.

Also, there is some question as to whether predatory pricing is likely to occur at all. It is certainly a high-risk strategy that would likely fail, even in the absence of any regulation, in many cases. Assuming a firm survives stage 1—that its competitors go bankrupt before it does—it is likely to face new entrants once price is jacked up at stage 2. Indeed, new entrants would know that the predator could not easily afford another round of predatory cuts after already enduring losses in driving out the first group of competitors. If the predator did successfully repeat his predation, then a third round of new entrants would be likely to be drawn by the knowledge that the predator would struggle to survive a third duel to the death, and so on. In other words, a successful predator must be so fierce that he completely frightens off the rest of the world. This is theoretically not impossible, and so could constitute a barrier to entry, but it is not likely to be common.

However, normal price competition is often mistaken for predatory behavior. If, for example, a competitor enters a market with a close substitute offered at a lower price, then clearly the incumbent firm must either match that price cut, at least approximately, or leave the market altogether. Airline A can't charge a price much above Airline B if B offers essentially the same product. If A was previously charging a much higher price before it then matches the much lower price of B, then A will also probably increase output. In effect, low-cost B forces A to abandon its higher-price/lower-volume strategy and embrace a low-price/high-volume strategy; the only alternative for A is to abandon the market completely. If B does eventually pull out of the market, then A may find it optimal to return to the high-price/low-volume strategy. The standard business procedure of matching a competitor's price cuts when necessary is indistinguishable from predatory pricing. Does it make any sense to forbid legacy carriers from matching the lower prices of low-cost entrants?

US courts considering predatory pricing charges in recent decades have focused not on the price cuts, but on the feasibility of a "predator" raising prices to monopoly levels. In, for example, *Frontier Airlines* v. *American Airlines*, the judge summarily dismissed the case because, he maintained, the government really had no case at all—that is, no credible evidence of predation. American Airlines merely matched the prices of Frontier but, in the court's view, had no hope of gaining any monopoly power even if it destroyed Frontier, since there were numerous other competitors in the market. The fact that American Airlines, like all legacy airlines, struggled to earn even normal profits over the long run is supportive of the court's decision. Though some economists may disagree with this approach, the courts' deep skepticism, combined with the fact that the legacy carriers have been struggling just to survive in recent years, seems to have dampened regulators' enthusiasm for bringing predatory pricing charges.

Cartels and Collusion

Firms that fix prices and output in a formal agreement are called a cartel. Firms that fix prices and output in covert informal agreements are said to be in collusion. Cartels and collusion may enable firms to exert monopoly-like power in their pricing policies (Hirschey, 2003). While cartels and collusion are generally illegal in the United States, they are allowed in many foreign markets. In the United States, under the Sherman Antitrust Act of 1890, the Clayton Antitrust Act of 1914, and the Federal Trade Commission Act of 1914, such collusive agreements are illegal. Nonetheless there are several examples where they can be found, particularly in the sports professions. The National Football League, Inc. (NFL), Major League Baseball, Inc. (MLB), the National Basketball Association (NBA), and the National Collegiate Athletic Association (NCAA) are often cited as examples of cartels.[11] Around the world, there have been famous cartels in oil and diamonds. Probably

11 However, it might also be argued that these sports leagues are not really cartels at all, since their product, entertainment, faces many substitutes. Cooperation among sports franchises might be of the same sort that exists among different franchises of a given restaurant chain.

the most famous and most important cartel in the world economy is OPEC (Organization of the Petroleum Exporting Countries). While OPEC cannot directly set the price of a barrel of oil, its control over much of the supply enables the cartel to dramatically influence the price by either increasing or decreasing output.

The international nature of, and often extreme competition in, the airline industry may lend themselves to potential collusion. For example, in June 2006 British Airways became involved in a price-fixing scandal involving fuel surcharges on long-haul flights (Simpkins, 2006). Investigations by both British and American authorities uncovered the fact that calls were made to Virgin Atlantic concerning the timing and level of increases in fuel surcharges (Simpkins, 2006). On the other hand, there is a distinct problem when analyzing cartel and collusion issues in the airline industry, namely the fact that the prevalence of information in the industry makes it fairly simple for airlines to match the prices and output of competitors. Therefore, the fact that two airlines have price fluctuations that match exactly does not mean that they are in collusion, but more likely that they are competing fiercely.

Airline alliances create interesting issues related to airline collusion, which ultimately concern coordinating schedules and prices of flights. In order for alliances to be allowed, they must receive regulatory approval from the necessary bodies, although limits may be placed on their coordination. For instance, American Airlines and British Airways have several stringent restrictions placed on them, while the KLM–Northwest relationship was given extensive antitrust immunity by US regulators. Antitrust immunity is given to potential alliances, based on a variety of factors including the level of consolidation that would exist in the industry.

Again, however, there is fierce debate as to whether government antitrust actions against alleged collusion have been appropriate and beneficial to society. The abysmal rate of return for legacy airline investment, perhaps the lowest for all industries, suggests that there is no shortage of competition. Even if collusion were attempted, it is difficult to orchestrate high prices for any length of time. For one thing, such high prices invite new competitors to enter the market and undercut the cartel. Even before new entry there is always strong incentive for an individual firm to violate the collusive agreement, since one can potentially earn far greater profits by slightly undercutting one's partners.

In fact, the behavior of US airlines in the age of regulation illustrates the strong tendency to break a cartel agreement. Regulators severely limited price cuts, much as a conventional cartel would do, and even prevented any new entry for some 40 years. Yet airlines struggled to earn even normal profits. Airlines competed by improving quality—improving the food, giving away liquor, utilizing larger aircraft, providing roomier seating, and so on. The fact that even this quasi-cartel, run with the aid of government, still failed to suppress competition illustrates how difficult it is for firms to secretly suppress competition on their own.

Using regulation to prevent price-fixing is problematic because there is often, as in the case of alleged predatory pricing, no clear way of distinguishing innocent behavior from illegal collusion. For example, two airlines may constantly raise and lower their fares in tandem not because they are colluding, but because of logical, independent reactions to market conditions. One airline's product is normally a very close substitute for another's—in the leisure market it may approach being a perfect substitute. Inherently, it is not possible for the prices of two close substitutes to vary greatly; most consumers would flock to the one that is substantially cheaper. Thus, it is necessary to generally match any price cut by a competitor. The prevalence of online information in the industry also makes it particularly easy for airlines to monitor and quickly match the prices and output of competitors. Likewise, any airline attempting to increase price because of, say, higher fuel costs, will retreat from the price increase unless competitors follow suit. The net result is either a general increase in airline prices or no lasting price increase at all.

In the 1990s US regulators made much of the fact that airlines seemed to be constantly signaling each other to raise prices. For instance, an airline would normally announce its intention to raise prices several weeks in advance. If, following the announcement but before the scheduled price increase, the competition eventually announced that they, too, would increase fares by similar amounts, the announced price hikes would in fact materialize. On the other hand, if competitors left prices unchanged, the airline would cancel the previously announced fare increase.

In their defense, airlines pointed out that most announced increases were in fact cancelled, that the overall long-run trend in inflation-adjusted fares was downward, and that the lack of industry profitability indicated that prices were not "fixed," but rather "broken." Moreover, as already explained, it is understandable that prices for close substitutes will naturally either move in tandem or not at all. It is also understandable that firms sustaining substantial losses would make every effort to raise prices.

Nevertheless, antitrust regulators insisted on changes, and, among other things, forced airlines to agree to no longer announce fare increases in advance. Many travel agents and consumer groups complained about the change, since forbidding airlines from announcing planned price hikes seemed to make fare increases harder to predict and plan for. Airlines responded to the regulatory constraint by implementing fare increases on Saturdays, when relatively few people book flights. Thus, if competitors refused to join in the price increase, the new prices could be cancelled by the following Monday, before they had significant impact.

Critics of antitrust regulation argued that the whole episode seemed absurd, and that it was particularly ironic for regulators to force firms to actually raise prices as opposed to just announcing an intention to eventually raise prices. Still, it does seem at least theoretically possible that preventing airlines from signaling a desire to increase prices might ultimately benefit consumers, at least slightly. Of course, regulators claimed that their actions and general vigilance would make sure that airline profits and prices would never rise too much.

INDUSTRY CONSOLIDATION

The level of concentration in the industry helps determine the market's structure. Industries that are highly concentrated may be more prone to exhibit characteristics of monopolies and oligopolies, while industries with multiple players may tend to exhibit characteristics of perfect and monopolistic competition. Table 9.4 presents the market share of the major Internet browsers over the past few years. It is clear from the table that Internet Explorer is the primary product in this industry, although its market share has dropped significantly in recent years due to increased competition and government rulings. It should also be noted that Safari's market share has increased, primarily due to Apple's increasing popularity with consumers.

There are two widely used methods for evaluating industry consolidation: the four-firm concentration ratio and the Herfindahl Index.

Table 9.4 Market share of Internet browsers

Browser	Market Share (2004)	Market Share (2005)	Market Share (2006)
Internet Explorer	91.35%	87.23%	82.95%
Firefox	3.66%	7.81%	11.59%
Netscape	2.09%	1.66%	0.94%
Safari	1.50%	2.15%	3.46%
Mozilla	0.80%	0.54%	0.29%
Opera	0.51%	0.52%	0.62%
Other	0.02%	0.09%	0.15%

Source: Browser Market Share

The Four-firm Concentration Ratio

The concentration ratio is a measure of the total market share held by a certain number of firms in the industry. The general form of the concentration ratio formula can be described as:

$$CR_m = \frac{\sum_{m=1}^{n} Q_m}{n} \times 100$$

where: n is the number of firms measured and Q is output.

The concentration ratio is simply the summed output of *n* airline companies divided by the total industry output. The most commonly used concentration ratio is the four-firm concentration ratio, which measures the output of the four largest firms in the industry. The market can then be classified according to a continuum of the percentage share of the top four.

Consider the US airline data for 2005 shown in Table 9.5. To find the four-firm concentration ratio of the US domestic airline industry in 2005, the four largest airlines, in terms of output, need to be grouped together and compared to the industry total. In this example, American, Delta, Southwest, and United would be grouped together. These four firms' combined output is 382,879,255,000 ASMs. When divided by the total industry output, this produces a four-firm concentration ratio of 37.2 per cent. This means that the four largest airlines produce 37 per cent of the industry's total output.

Markets are separated into categories along the market continuum based on the four-firm concentration ratio. Markets where the four-firm concentration ratio is less than 20 per cent are assumed to mimic perfectly competitive industries. At the other end of the spectrum, markets where the concentration ratio is above 80 per cent are highly concentrated, and are assumed to be closer to the monopoly end of the market continuum spectrum. In the middle of the spectrum, where the majority of industries lie, monopolistic

Table 9.5 US airlines' available seat miles, 2005

Airline	ASMs
American	115,135,372,000
Continental	50,435,785,000
Delta	98,497,748,000
Northwest	55,284,927,000
Southwest	85,353,571,000
United	83,892,564,000
Industry Total	1,029,168,153,000

competition markets would have a four-firm concentration ratio between 20–50 per cent, while oligopolies would have a 50–80 per cent four-firm concentration ratio. Therefore:

- Perfect competition: a very low concentration ratio
- Monopolistic competition: below 40 per cent for the four-firm measurement
- Oligopoly: above 40 per cent for the four-firm measurement
- Monopoly: with near 100 per cent for the four-firm measurement.

Earlier in the chapter it was claimed that the airline industry is an example of an oligopoly market; however, the four-firm concentration ratio calculated above does not seem to support this claim. The reason for this is the fact that, for the airline industry, the four-firm concentration ratios should be calculated on an airport-by-airport basis: most airports have are only a few major carriers, and this produces an oligopolistic market in most of the industry. Since consumers are ultimately affected on an individual market basis, assessing the four-firm concentration ratio on an airport-by-airport basis provides a more realistic picture of the air transport industry.

Table 9.6 provides a synopsis of the four-firm concentration ratios when calculated on an airport-by-airport basis for five major airports in the United States. Although the industry's concentration ratio was only 37 per cent, all five of the airports analyzed had concentration ratios substantially above that. In the case of Atlanta and Dallas, one dominant hub carrier receives almost monopoly power because it effectively controls the market. Chicago, with both United and American operating hubs out of the airport, is essentially a duopoly. Finally, Los Angeles and Las Vegas both exhibit strong oligopolistic market tendencies with their four-firm concentration ratios equaling roughly 65 per cent. Similar statistics would be found if the analysis were applied to other airports.

Using an airport-by-airport analysis of the domestic aviation market, it is clear that the industry resembles a strong oligopoly rather than the monopolistically competitive market suggested by the industry-wide four-firm concentration ratio. While every market is unique, with varying levels of concentration, it is unlikely that any one airport would have a low four-firm concentration value. The four-firm concentration ratio enables the analyst to get a quick look at the amount of concentration in the industry, but it usually requires a more in-depth analysis to fully understand the specific market situation.

Table 9.6 Four-firm concentration ratio

Market Concentration	
Airport	Ratio
Atlanta	93.75%
Chicago	81.75%
Dallas	90.48%
Las Angeles	65.69%
Las Vegas	66.66%

Source: Compiled by the authors using Back Aviation Form41 data.

Herfindahl-Hirschman Index

Another method used to analyze the amount of concentration in an industry is the Herfindahl index (also known as the Herfindahl-Hirschman index, HHI). This is a popular measure that was used in Chapter 1 to analyze the amount of consolidation that exists in the industry. HHI is obtained by squaring the market share of each of the players, and then adding up those squares:

$$HHI = \sum S^2_m$$

where: S^2 is m-firms' market share.

The measure is simply the cumulative squared value of the market share for every firm in the industry. Therefore, the higher the index, the more concentration and (within limits) the less open the market. By squaring the market share values, firms with a large market share receive more weight in the calculation than firms with a smaller market share. The US Department of Justice considers a market with a result of less than 1,000 to be a competitive marketplace. A result of 1,000–1,800 is a moderately concentrated marketplace, and a result of 1,800 or greater is a highly concentrated marketplace. It should also be noted that market share can be calculated in terms of different products; therefore, unique HHI indices could potentially be created for the same market.

For example, in a duopoly market, if each of the two firms in the market has a market share of 50 per cent:

$$HHI= (50)^2 + (50)^2 =2500 + 2500 = 5000.$$

With two firms that have 80 per cent and 20 per cent respectively:

$$HHI = (80)^2+ (20)^2 = 6,800$$

Based on data, obtained through Form41, the industry HHI for the domestic US airline industry was calculated for the past several years and is presented in Table 9.7. The respective market share for each airline was based on the number of passengers enplaned.

Table 9.7 US Domestic Herfindahl Indices, 1998–2005

Year	HHI
1998	1016.02
1999	957.02
2000	932.28
2001	918.13
2002	940.59
2003	890.24
2004	838.02
2005	808.58

Source: Compiled by the authors using Form41 data.

It was found that the general trend in the US airline industry is toward deconcentration, as the HHI value has dropped over 200 points since 1998. The current HHI value of around 800 is fairly typical of an oligopolistic market, as values greater than 1,000 generally indicate a high degree of concentration in the market. However, just as in the four-firm concentration ratio calculation, a much higher degree of concentration exists at individual airports.

The Herfindahl Index is frequently used by the Department of Justice to determine if a proposed merger is acceptable for antitrust reasons. For example, if we consider the proposed 2006 Delta–US Airways merger, Table 9.8 presents the market share for the major carriers pre-merger, and Table 9.9 provides the new market share if the merger occurs.[12]

Based on the proposed merger, the Herfindahl Index would increase by slightly less than 200 points, which is well above the limit of 100 that the Department of Justice usually acts on. Furthermore, the Department of Justice usually does not like mergers that raise the industry Herfindahl Index above 1,000, as any point above that is deemed too monopolistic. In addition to these industry-wide Herfindahl measures, certain markets, particularly in the north-east, would experience far greater increases in the HHI. Therefore, based on statistical analysis alone, it would appear that the proposed merger would have trouble garnering antitrust immunity. However, since the Herfindahl Index is only one of many measures employed by the Department of Justice when evaluating mergers, the merger could be approved on other grounds.

BEYOND MARKET CONCENTRATION CONSIDERATIONS

Although the above calculations of market concentration lend a certain aura of rigor, it remains a very arbitrary decision criterion, partly because it completely ignores how mergers may increase efficiency and/or reduce prices through economies of scale, scope, and density impact. Of course, it is not possible for anyone, including the merging

12 Note that the Herfindahl values calculated included data from carriers not listed in the tables. Tables 9.8 and 9.9 merely provide an overview of the major carriers' market share.

Table 9.8 Pre-merger market share

Airline	Market Share
American	11.36%
Alaska	2.36%
JetBlue	3.61%
Continental	5.44%
Delta	9.74%
AirTran	3.05%
Northwest	6.71%
United	8.14%
US Airways	8.66%
Southwest	18.32%
Herfindahl Index	838.68

Source: Compiled by the authors using Back Aviation Form41 data.

Table 9.9 Post-merger market share

Airline	Market Share
American	11.36%
Alaska	2.36%
JetBlue	3.61%
Continental	5.44%
AirTran	3.05%
Northwest	6.71%
United	8.14%
US Airways	18.40%
Southwest	18.32%
Herfindahl Index	1007.38

Source: Compiled by the authors using Back Aviation Form41 data.

airlines themselves, to know with certainty exactly how efficient the newly combined airline will be. The mix of corporate cultures, merging of separate labor unions, and/or other factors creates uncertainties that become clear only long after the merger actually occurs. Proponents of strong anti-merger regulation argue that, with no guaranteed gains in efficiency, it makes sense to keep the number of competitors as high as possible for as long as possible.

On the other hand, opponents of such vigorous regulation maintain that the dismal rate of return for the airline industry shows that there is more than enough competition, implying that firms should generally be free to combine as they choose. The struggling industry should be allowed to repair itself. For instance, Ben-Yosef (2005, pp. 265–66) suggests that, had government regulators allowed them to merge, it is quite possible that United Airlines and US Airways might have avoided bankruptcy and been in a better position to keep fares low enough to profitably compete with the low cost airlines. Thus, rather than focusing on concentration ratios, regulators might better serve the public interest by generally allowing troubled firms to merge as they see fit, and allowing the industry to evolve as it will as long as there is no indication of higher than normal long-run profits being earned.

So far, regulators have allowed such free choice in mergers only if it becomes obvious that one firm is on its way to shutting down anyway. However, airline alliances, which might be viewed as a sort of partial merger, have often been allowed considerably more freedom. The KLM–Northwest alliance, for example, is quite extensive, having received antitrust immunity from US regulators. Sometimes governments severely restrict the action of alliance partners, such as in the case of American Airlines and British Airways, but at least some cooperation is allowed. The greater degree of freedom allowed in alliances seems to represent some compromise; regulators may be implicitly admitting that a rigid focus on concentration ratios is not appropriate in an industry largely floundering in bankruptcy. At the same time, from a pro-regulatory viewpoint, if alliances should prove to be somehow anti-competitive, they can more readily be altered or even completely undone.

ANTITRUST, MARKET EVOLUTION, AND COOPERATION

As mentioned above, some economists believe there is a an "empty core" problem where the high fixed cost/low marginal costs of the airline industry require more cooperation between firms to avoid competition that destroys profits and prevents efficient production and product innovation. The rivalry between Airbus and Boeing may also illustrate how some cooperation can benefit consumers. Consider, for example, the problematic production of the "jumbo aircraft," the A380. Suppose Boeing had decided to make its own version of the A380 and had then begun to encounter problems similar to those experienced by Airbus. Maybe, in these circumstances, both companies would have decided to simply abandon production. With each of them having to share demand with the other, the costs might have been prohibitive. Of course, it never came to this because Boeing chose not to enter the A380 market. Whether intentional or not, there was a sort of implicit cooperation in which Boeing stepped aside to make the project, perhaps, viable for Airbus. Governments may unwittingly facilitate such cooperation through patent laws, which can have the effect of segmenting the market for different producers.

Such implicit cooperation is less likely to take place, though, where the number of firms is greater. The airline industry may need explicit contracts to coordinate an efficient allocation of resources for consumers. Airlines might be able to offer more nonstop service or move to larger, more efficient and comfortable jets on more routes if they were able to cooperate explicitly. Some of this might be arranged through mergers and some through more limited alliances that might sometimes resemble cartels, but could be aimed at arranging efficient production rather than suppressing competition. Removal of the restrictions on cabotage and international mergers would facilitate this, and would also reduce entry barriers to help reduce the possibility of the cooperation taking an anti-competitive turn.

Of course, current regulators, and many economists, would oppose such a move. It might be, for instance, that the lack of industry profitability reflects simple overcapacity rather than the need for complex cooperation. Eventually, bankruptcy, capacity cuts, and liquidation may return the industry to normal profitability. In any case, many economists argue that airline consolidation in some form, both in Europe and the United States, is inevitable. According to this view, the record shows that the industry cannot be profitable, whether because there is simply too much capacity from too many airlines or because of a more complex lack of coordination. Since investors will ultimately require a reasonable rate of return to keep capital in the industry, some capital will be withdrawn, so the number of large airlines is probably bound to decline slightly. Regulators might slow this decline in numbers, but can not prevent it, and may, as outlined above, cause the adjustment to be less orderly and more severe than it would be if they simply got out of the way. Only time will tell.

SUMMARY

This chapter examined the models of monopolistic competition and oligopoly, the so-called hybrid markes. Most aviation industries fit the oligopoly model, but there are differencing views of what this implies. Some economists see oligopoly as inherently problematic, while others point to the lack of high profits in many oligopolies, particularly the airlines, and conclude that entry barriers are not so significant after all. The chapter then provided an overview of various theories of oligopoly; these included contestability, the kinked demand curve, and Cournot models. Empirical evidence indicates that the airline industry may be contestable when viewed as a network. From this viewpoint, it may even be that increased concentration and cooperation through more alliances can benefit consumers, primarily through economies of scale, and return the airline industry to a reasonable level of long-run profits. Finally, the chapter introduced various indices to measure the amount of concentration in the industry; these included the four-firm concentration ratio and the Herfindahl-Hirschman Index.

APPENDIX: A MATHEMATICAL EXAMPLE OF A PRICING DECISION

In order to understand the mathematics involved with determining the optimal price–output level for a firm in an oligopolistic market, consider the case of luxury airline DirectJet. In response to a request by DirectJet's management to study short-run pricing and production policy, DirectJet's financial managers were able to determine the airline's price, fixed cost, and variable cost functions. They were:

$$P = 10{,}000 - 8Q$$
$$FC = \$200{,}000$$
$$VC = 2{,}200Q + 5Q^2$$

Since the ultimate goal is to determine the optimal price–output decision for DirectJet, both the marginal revenue and marginal cost curves need to be found. On the revenue side, the total revenue function can be found by simply multiplying the price function by quantity (Q). This results in:

$$TR = 10{,}000Q - 8Q^2$$

The first-order derivative of the total revenue function will produce the marginal revenue function of:

$$MR = 10{,}000 - 16Q$$

On the cost side, the two cost components (fixed and variable) need to be combined together to create the total cost function. From there, the first-order derivative of the total cost function will yield the marginal cost function. The two functions are:

$$TC = 200{,}000 + 2{,}200Q + 5Q^2$$
$$MC = 2{,}200 + 10Q$$

The final step in determining the optimal price–output combination for DirectJet is to set the marginal revenue and marginal cost curves equal to each other in order to determine the optimal quantity. This is computed below and also displayed graphically in Figure 9.8.

$$MR = MC$$
$$10{,}000 - 16Q = 2{,}200 + 10Q$$
$$7{,}800 = 26Q \rightarrow Q = 300$$

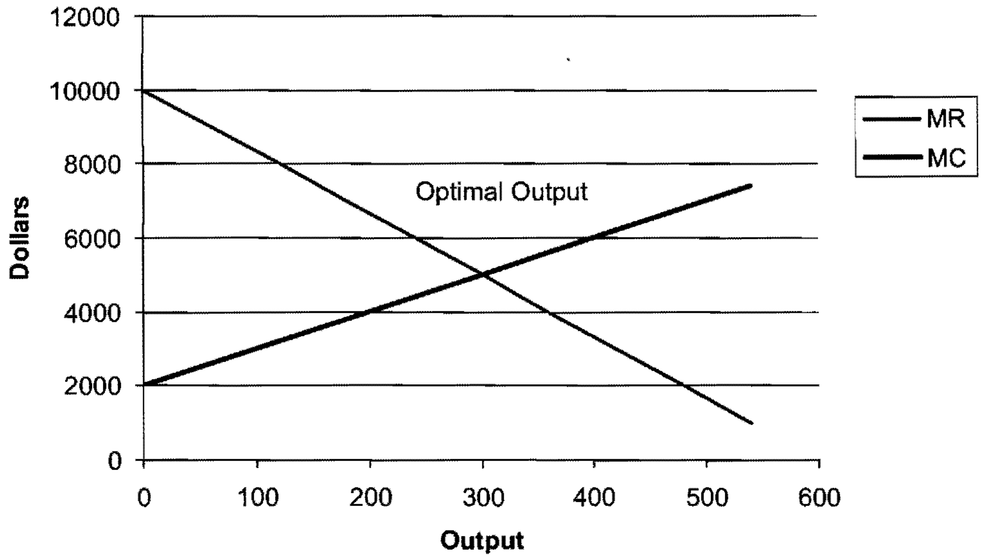

Figure 9.8 Marginal revenue and marginal cost for DirectJet

Based on this calculation, DirectJet's optimal output is 300 seats per day. At this optimal level, the average ticket price for DirectJet's flights and the total revenue are:

$$P = 10{,}000 - 8Q$$
$$P = 10{,}000 - 8(300)$$
$$P = \$7{,}600$$
$$TR = 10{,}000Q - 8Q^2$$
$$TR = 10{,}000(300) - 8(300)^2$$
$$TR = \$2{,}280{,}000$$

or

$$TR = P^*Q \rightarrow \$7{,}600 * 300 = \$2{,}280{,}000$$

While total revenue is not maximized at this point, total firm profit is maximized at this value. In order to determine total profit at the optimal level, the total cost at the optimal output is:

$$TC = 200{,}000 + 2{,}200Q + 5Q^2$$
$$TC = 200{,}000 + 2{,}200(300) + 5(300)^2$$
$$TC = \$1{,}310{,}000$$

This yields a total profit at the optimal output level for DirectJet of:

$$TP = TR - TC$$
$$TP = \$2{,}280{,}000 - \$1{,}310{,}000 = \$970{,}000$$

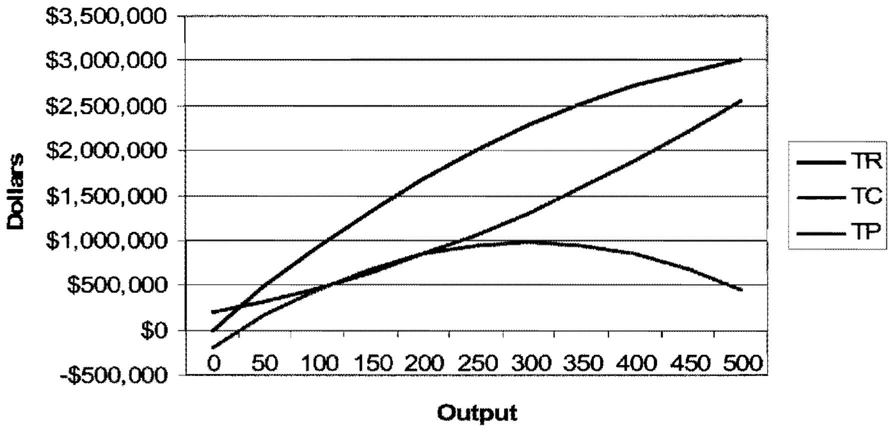

Figure 9.9 Total revenue, cost, and profit for DirectJet

No other combination of price and output based on the cost functions provided by DirectJet's financial managers will yield the airline with a greater profit. This is displayed graphically in Figure 9.9, which displays the total revenue, total cost, and total profit functions. Since the total profit curve is an upward "u-shape," the output level of 300 is a maximum point and is the point where total profit is maximized.

While the above scenario applies in the short term, it is very much different in the long run. As in perfectly competitive markets, the supernormal profits earned by monopolistic competitive firms in the short run will attract new firms to the market. These new firms will offer similar competing products, but with the increase in new firms and products entering the market, the degree of differentiation between products diminishes. As a result of the decreasing degree of product differentiation, the elasticity of the firm's demand decreases, causing the firm's demand curve to become more horizontal. This creates a situation similar to perfectly competitive markets as monopolistic competitive firms tend to become price-takers in the long run. This causes the supernormal profits to diminish, and zero economic profits are earned. In essence, monopolistic competition acts like perfect competition in the long term; therefore, the differentiating characteristic between the two is the length of the "short-run" period where supernormal profits are realized.

REFERENCES

Armentano, D. (1986) *Antitrust Policy The Case for Repeal.* April, Washington, DC: Cato Institute.

Bailey, E. (1981). Contestability and the Design of Regulatory and Antitrust Policy., *American Economic Review*, 71(2), May, pp. 178–83

Bailey, E. and Panzar, J. (1981). The Contestability of Airline Markets during the Transition to Deregulation. *Law and Contemporary Problems*, 44, Winter, pp. 125–45.

Ben-Yosef, E. (2005). *The Evolution of the US Airline Industry: Theory, Strategy and Policy.* Studies in Industrial Organization Series , Vol. 25, Dordrecht: Springer

Browser Market Share. Retrieved 28 September 2007, from Market Share by Net Application at: http://marketshare.hitslink.com.

Crandall, R. and Winston, C. (2003). Does Antitrust Policy Improve Consumer Welfare? *Journal of Economic Perspectives*, 19 October, pp. 3–26.

Galbraith, J. K. (1979). *Age of Uncertainty*. Boston, MA: Houghton Mifflin.

Hirschey, M. (2003). *Managerial Economics* (10th edn). Ohio: South-Western.

Oum, T., Zhang A., and Zhang, Y. (1993). Inter-firm Rivalry and Firm-specific Price Elasticities in Deregulated Airline Markets. *Journal of Transport Economics and Policy*,27(2), pp. 171–92.

Raghavan, S. and Raghavan, J. (2005). Application of Core Theory to the US Airline Industry. *Journal of the Academy of Business and Economics*, 5(3), pp. 116–25.

Simpkins, E. (2006). BA in the Dock—The Airline is being Investigated for Potential Collusion on Fuel Surcharges. *The Daily Telegraph*, 26 June, p. 6.

Strassmann, D. (1990). Potential Competition in the Deregulated Airlines. *Review of Economics and Statistics*, 72(4), pp. 696–702.

Whinston, M. and Collins, S. (1992). Entry and Competitive Structure in Deregulated Airline Markets: An Event Study Analysis of People Express. *RAND Journal of Economics*, 23(4), pp. 445–62.

10

Aviation Forecasting and Regression Analysis

So we went to Atari and said, "Hey, we've got this amazing thing, even built with some of your parts, and what do you think about funding us? Or we'll give it to you. We just want to do it. Pay our salary, we'll come work for you." And they said, "No." So then we went to Hewlett-Packard, and they said, "Hey, we don't need you. You haven't got through college yet.

Steve Jobs, founder of Apple Computer Inc.

Forecasting is one method for reducing the uncertainty in the business world, and, whether we realize it or not, we all forecast in one way or another. For example, if I take an umbrella out with me, I am obviously forecasting rain with some degree of probability. And, of all industries, the aviation industry stands out as one of the most uncertain and unpredictable. Since the variability of such inputs as jet fuel prices and passenger demand can dramatically affect airlines, forecasting is extremely useful in reducing the uncertainty associated with this volatility.

This chapter will look at the applications of forecasting in the aviation industry, and discuss many of the major forecasting methods used. Greater emphasis will be placed on quantitative tools, such as regression analysis, since it is the most powerful forecasting method to be discussed. The topics discussed in this chapter are as follows:

- Aviation forecasting applications
- Qualitative forecasting methods, including:
  - Focus group
  - Market survey
  - Market experiment
  - Barometric forecasting
  - Historical analogy
  - Delphi method
- Quantitative forecasting methods
- Descriptive statistics, including:
  - Mean
  - Variance
  - Standard deviation
- Time series analysis, including:
  - Trend analysis

- Seasonal variations
- Cyclical variation
- Random effect, including:
  - Moving average
  - Weighted moving average
  - Exponential smoothing
  - Trend analysis
- Forecast accuracy

- Regression analysis, including:
  - Goodness of fit
  - Performing regression analysis
  - Dummy or binary variables
  - Autocorrelation
- Data sources, including:
  - US Department of Transportation (DOT) / Bureau of Transportation Statistics (BTS)
  - Federal Aviation Administration (FAA)
  - International Air Transport Association (IATA)
  - International Civil Aviation Organization (ICAO)
  - Official Airline Guide (OAG)
  - Airports Council International (ACI)
  - Air Transport Intelligence (ATI)
  - Airline Monitor
  - UK Civil Aviation Authority (CAA)
  - Transport Canada (TC)
  - Euro-control
  - Aircraft Owners and Pilots Association (AOPA)
  - Bureau of Economic Analysis (BEA)
  - Bureau of Labor Statistics (BLS)
  - Organization for Economic Cooperation and Development (OECD).

AVIATION FORECASTING APPLICATIONS

> Airplanes are interesting toys but of no military value.
>
> Professor of Strategy, Ecole Supérieure de Guerre

Forecasting has many applications in the aviation industry, and probably the chief among these is the forecasting of demand. Since demand is not monolithic and varies for every flight, sophisticated forecasting tools need to be applied to help forecast the size and the nature of demand. As forecasting the nature of demand could include the mix between price-sensitive (leisure), time-sensitive (business) travelers and the expected booking rate for the flight, when these forecasting methods are applied to every flight, the forecasting operation becomes quite large. Moreover, demand forecasting is probably one of the most important applications of forecasting in the airline industry since strategic planning and yield management are dependent on it.

However, forecasting is not just limited to demand forecasting, and the methods employed can range from the rudimentary to the sophisticated. Planning of human resources, financial resources and needs, route developments, aircraft fleet, and infrastructure expansion are all based on some expectation of future events. In 2006 Airbus forecasted that between 2006 and 2025 world passenger traffic would increase by 4.8 per cent per year. This traffic growth, combined with fleet replacement, will require the production of 21,860 additional new passenger aircraft with more than 100 seats.[1] Other examples, important for airlines, might be forecasts concerning the amount of flying, crew requirements, training schedules, absenteeism, and employee turnover ratios. In addition, project viability and profit projections are all based on the expectation of future events. Since projects are analyzed over their lifespan, forecasts need to be created concerning future expected cash flows. On the basis of these forecasts, multi-million dollar projects are either approved or rejected. For example, evaluating whether to install AVOD (audio-visual on-demand) in-flight entertainment systems across the fleet could be based on forecasts concerning the installation schedule, future maintenance costs, and the passengers' opinions of the new system.

As the above examples highlight, forecasting spans multiple functional areas. Therefore, it is critical to understand the many aspects forecasting. To do this, four critical skills are needed:

- knowledge of the airline industry
- knowledge of economic principles and statistics
- computer applications
- communication.

Although all four skills are required to forecast, this book only covers the first two skills in detail. The overall text itself provides knowledge of the airline industry and basic economic principles, while this chapter will discuss the basic statistics used in forecasting. Of course, the chapter cannot replace a complete statistics textbook, but is meant to provide an introductory and applied overview. A variety of computer applications can be used to help forecast, including Microsoft Excel that can help perform basic regression and statistical analysis.

QUALITATIVE FORECASTING METHODS

> No flying machine will ever fly from New York to Paris ... [because] no known motor can run at the requisite speed for four days without stopping.
>
> Orville Wright, 1908

Forecasting methods can be broken down into two categories: qualitative and quantitative. Qualitative forecasting methods use subjective techniques to help forecast. Qualitative forecasts do not use statistical databases or provide measures of forecast accuracy since they are based on opinions, surveys and beliefs. On the other hand, quantitative

1 Airbus, *Global Market Forecast, 2006–2025.*

forecasts use statistical relationships to help forecast future events and, while they are more mathematical in nature, they may or may not be any more accurate than qualitative forecasts. Table 10.1 displays a few of the main advantages and limitations of qualitative forecasting.

One of the main advantages of qualitative forecasts is that they are flexible and can be easily altered to reflect any changes in the economy or environment. The flexibility of qualitative forecasting also enables early signals of changes and anomalies in data to be recognized.

On the other hand, one of the limitations of qualitative forecasting is that it can be difficult to track and isolate the primary variable that is causing changes in the dependent variable. In addition, the lack of tests for accuracy always creates a situation where there is no way of knowing how good the forecast is.

Of the numerous types of qualitative forecasts, six major methods will be discussed here. They are:

- focus group
- market survey
- market experiment
- barometric forecasting
- historical analogy
- Delphi method

Focus Group

A focus group is a relatively informal information gathering procedure in marketing research. It typically brings together 8–12 individuals to discuss a given subject. Usually focus group participants are brought into a room where a moderator asks questions to help move the discussion forward. Researchers observe the participants and their responses, providing a quick and relatively inexpensive insight into their research problem. Focus groups can be quite effective in evaluating new product options, such as new aircraft seats or in-flight entertainment systems. However, researchers need to be aware that participants may not produce completely honest responses and may feel pressured into accepting what everyone else believes. Furthermore, if the focus group is

Table 10.1 Advantages and limitations of qualitative forecasting

Qualitative Forecasting	
Advantages	**Limitations**
Flexibility—easily altered as the economy changes	Complex—hard to keep track of interactions in the primary variables
Early signals—can catch changes and anomalies in data	Lack of tests of accuracy—cannot easily test the accuracy in prior periods

not a representative sample of the target population, then the responses are likely to be inaccurate., By forming a focus group, an airline can reach out to its potential passengers for feedback and comment. Airlines generally use focus groups in planning, marketing, or serving a new destination domestically or internationally. Finally, a focus group can help airlines managers to identify the right level of in-flight and cabin crew services.

Market Survey

A survey is simply a method of acquiring information by asking people what they think will happen. The most common method of a market survey is by questionnaire, but there are many other methods. Depending on the nature of the questions asked, questionnaires can provide the researcher with both quantitative and qualitative results. One of the principal benefits of market surveys is that they are easy to use and do not require advanced theory or econometric analysis to interpret the results. One of their potential flaws is that the accuracy of the survey depends on the size and responsiveness of the sample. In the aviation industry, market surveys are often used to find out what improvements an airport can make, and airlines use them in the form of customer comment cards to find out ways of improving service.

Market Experiments

A more expensive method of qualitative forecasting is a market experiment, which involves testing new product factors, such as prices or packaging, in a few test markets (Hirschey, 2006). Market experiments use real-life markets, and this can be risky if the change is not accepted by consumers in that the change may permanently alienate them from the product. (Allen, Doherty, Weigelt, and Mansfield, 2005). Due to the costs involved, market experiments are rarely used in forecasting demand in the aviation industry. However, market experiments have been conducted with in-flight food and beverages. For example, an airline may wish to test the acceptance of new buy-on-board meals on a few flights to help forecast demand for the product.

Barometric Forecasting

Barometric forecasting involves using current values of certain variables, called indicators, to help predict future values of other variables (Truett and Truett, 1992). A leading indicator is a variable whose current changes give an indication of future changes in other variables. A lagging indicator is a variable whose changes typically follow changes in other economic variables. Depending on one's point of view, the relationship between any two variables can either be a leading or lagging indicator. For forecasting, a leading indicator is the most useful, since it provides an early signal of what may come. In aviation, GDP is a leading indicator of airline demand as changes in GDP are usually followed by changes in demand for air transportation. Therefore, if the GDP growth rates decline sharply over a period of time, airlines can expect a decline in demand shortly afterwards. Part of the reason for this lagging relationship is that consumers are generally slow to adjust their spending patterns to reductions in income.

A final type of indicator, a coincident indicator, is not tremendously useful for aviation forecasting, but does have other applications. Coincident indicators are variables whose changes roughly coincide with changes in other economic variables. In aviation, crude oil prices are a coincident indicator of jet fuel prices. Since changes in the price of crude oil occur at the same time as changes in jet fuel prices, crude oil is not a useful indicator for predicting jet fuel prices. However, crude oil could be used as a proxy variable in jet fuel hedging, since crude oil is more heavily traded than jet fuel.

Historical Analogy

Historical analogy is a simple forecasting technique where the future is forecast on the basis of historical events. While many quantitative forecasting methods use historical data to help predict the future, historical analogy is, by definition, on a qualitative level. The success of historical analogy largely depends on the depth of knowledge and history that the forecaster has. A forecaster who has seen all facets of an industry over a long period can probably better predict the future than a more recent employee. Therefore, historical analogy is only as good as the person making the forecast.

Delphi Method

The Delphi method is related to historical analogy in that the forecast is largely based on opinion, but differs in the fact that it collects forecasts and opinions from an independent panel of experts. Each expert provides their analysis and opinion independently, and then a consensus forecast is created, based on the analysis provided by each member of the panel. By having members independently submit their opinions, the Delphi method benefits from not having steamroller or bandwagon problems (Hirschey, 2006). In addition, the Delphi method benefits from having multiple experts analyze the issue rather than just one or a few people as in historical analogy. In theory, the accuracy of the forecast is based on the collective knowledge of the expert panel; however, because every opinion is equally weighted, the collective knowledge may not be as reliable as the forecast of just a few experts.

QUANTITATIVE FORECASTING METHODS

In contrast to qualitative methods, the quantitative methods use statistical data to analyze and forecast future behavior of specific variables. Statistical information is divided into time-series and cross-sectional data. Cross-sectional data are data compiled for different variables at a point in time; for example, the number of passengers over different geographically located airports, or the number of aviation accidents over different countries for one time period. Time series data, on the other hand, represent observations of a particular variable over a number of time periods—for example, the number of passengers at various past points in time at a given airport.

Although there are many methods of quantitative forecasting, we shall cover only two broad categories: time-series analysis and regression analysis. Time series analysis looks for patterns in data whereas regression analysis assumes a casual relationship between two or more variables. Both methods will be analyzed in greater detail below.

Quantitative forecasting has several advantages and limitations, which are summarized in Table 10.2.

One of the main advantages of quantitative forecasting is that tests of reliability can easily be performed to determine the accuracy of the forecast. In time-series analysis the most accurate forecasting method can be chosen, based on the test of reliability. In regression analysis forecasters are able to provide a not only a probability of how accurate the overall forecast is, but also the reliability of the individual variables in the forecast (Spirtes, Glymour, and Scheines, 2001). Two major drawbacks of using only quantitative forecasting are, first, that history is not always a correct predictor of the future, and, second, that, historically, that forecasts have required extensive data collection and processing. However, the introduction of advanced statistical computer software, has now made data collection and processing a much simpler task, making quantitative forecasting much easier. Another limitation of quantitative forecasting is the quality of the data that is used. Depending on how good the data are, quantitative forecasts may distort reality or model it perfectly. The garbage-in-garbage-out cautionary statement would apply here.

DESCRIPTIVE STATISTICS

Prior to analyzing various forecasting methods in detail, a fundamental understanding of elementary statistics is required, and, for most forecasting, that amounts to three basic statistics: the mean, the variance, and the standard deviation. Descriptive statistics are numerical estimates that organize, sum up, or present the data, and provide simple summaries about the sample and the measures. To meet that objective, a full range of indicators has been developed, and we will provide definitions and applications of the most important ones.

Mean

The mean—probably the most common indicator of a data set—is simply the average of the data. For students, a commonly used mean is the class average which measures the average scores of all the students in the class. For instance, assume that the class has three students. Student A received a 92 on the final exam, student B received a 65, and student

Table 10.2 Advantages and limitations of quantitative forecasting

Quantitative Forecasting	
Advantages	**Limitations**
Organize relationships	Economic changes may distort results
Behavioral relationships	Extensive data-mining of information
Tests of reliability determine forecast accuracy	Only a crude approximation of reality

C received a 77.[2] Based on this information, the class average, or mean, for the final exam was 78. This was calculated by taking the sum of the students' marks, and dividing by the number of students in the class. From this, the general form for the mean is:

$$Mean = \mu = \frac{\sum_{i=1}^{n} X_i}{n}$$

In statistics, the mean is usually denoted with the Greek letter, μ. In Microsoft Excel, the mean can simply be calculated by using the average function.

Variance

The variance of a sample measures how the observations are spread around the mean. Large variance means that the observations are widely scattered around the mean. The variance of variable x is simply the summed squared difference between the actual values, x, and the mean of x. The variance shows the dispersion of the data from the mean. In statistics, the variance is denoted as σ^2, with the general form being:

$$\sigma^2 = \frac{\sum_{i=1}^{n} (x_i - \mu)^n}{n-1}$$

The variance of a data set is important, since it gives some idea of the accuracy of the mean. For example, the variance in the previous class average example is 183.[3] The variance is computed in Table 10.3.

In this situation, the variance is relatively large, since there is a wide dispersion between the actual grades and the class average. In order to highlight how the mean is irrelevant without the variance, assume that student A received a 78 and student B received a 79 on their final exam. In this situation, the class average remains the same, but the variance drops dramatically to only 7. The calculation for the variance is displayed in Table 10.4.

Table 10.3 Variance calculation of class average (1)

	Grade	x - μ	(x - μ)²
Student A	92	14	196
Student B	65	-13	169
Student C	77	-1	1
μ	78	σ²	183

2 Sample mean is generally presented by $\bar{X}$ and population mean by the Greek letter μ.

3 The population standard deviation is represented by the Greek letter sigma σ, and the sample standard deviation is represented by S.

Table 10.4 Variance calculation of class average (2)

	Grade	x - μ	(x - μ)²
Student A	76	-2	4
Student B	81	3	9
Student C	77	-1	1
μ	78	σ²	7

This example displays how the variance is critical to understanding data. In the first scenario the variance was quite large, indicating that the mean did not properly represent the data. However, in the second scenario the extremely small variance indicated that the mean was a better representation of the data.

Standard Deviation

Standard deviation is directly related to variance since it is the positive square root of the variance. The standard deviation is a statistic that shows how tightly all the observations are clustered around the mean in a set of data. When the observations are spread around the mean, it indicates that we have a relatively large standard deviation. In order to avoid problems with the negative signs of some deviations from the mean (note that the sum of the values of the deviations from the mean in the above example is zero, and this will always be true by definition), the values of the deviations are displayed in squared terms. By taking the square root of the variance, the standard deviation returns the variance to a more easily interpretable number. However, it should be noted that, while the squaring procedure eliminates the problem of negative deviations canceling out positive ones, it also gives much greater weight to outlying observations. That is, the further away from the mean the observation is, the greater the difference between the observation and the mean, and therefore the greater the squared value of this observation.

In the first example, the standard deviation is simply 13.52. Standard deviation is usually denoted as, σ, with the general form being:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \mu)^n}{n-1}}$$

The concept of standard deviations is not difficult to understand. Assume that we have collected one month of ticket prices—for example, from New York to London, about 1250 observations—and entered them into a spreadsheet to calculate the average. Suppose the average price is calculated as $870. This number, by itself, is of limited value. By measuring the standard deviation of the ticket price, however, we can gain an idea of how volatile the ticket price really was (the larger the standard deviation the more volatile the ticket price). For a perfectly normal distribution, 68.4 per cent of all the observations fall within plus or minus one standard deviation of the average, 95.4 per cent fall within plus or minus two standard deviations of the average, and 99.7 per cent of the observations fall within plus or minus three standard deviations of the average .[4] To summarize:

About 68.3% of the data will be within: $X \pm 1\partial$

About 95.0% of the data will be within: $X \pm 2\partial$

About 99.0% of the data will be within: $X \pm 3\partial$

TIME-SERIES ANALYSIS

As mentioned earlier, time-series analysis measures the status of some activity, such as aviation accidents, number of aircraft operations, or number of enplanements over a period of time. Time series records the activity by means of measurements taken at equally spaced intervals using a consistent activity and method of measurement. Observations may be carried out annually, quarterly, monthly, weekly, daily, or every hour. All time-series data contain four components:

- trend analysis
- seasonal variations
- cyclical variations
- random effect.

Trend Analysis

The trend component accounts for the movements of time series over a long period of time. Any regular patterns of values above and below the trend component are probably attributable to the cyclical component of a time series. For the air transport industry, this shifting, or trend, is usually attributed to factors such as liberalization, deregulation, change in disposable income, introduction of new technology, population growth and/or privatization. The overall trend of demand has been consistently increasing (see Figure 10.1). The volume of traffic during the1968–2006 period grew at a healthy rate of 6.15 per cent per year.[5]

Seasonal Variations

The seasonal component accounts for regular patterns of variability within certain time periods, such as over a year. For an airline, the number of passengers may be very high in certain months or seasons and low in others.

4 A normal distribution is a distribution where the area to each side of the mean is equal to 0.5.

5 The growth rate in each period is the ratio of the absolute change in RPM to an earlier value: $g_{RPM} = \frac{RPM_t - RPM_{t-1}}{RPM_{t-1}}$

Cyclical Variations

The cyclical component refers to long-term fluctuation of time-series statistics over time.

The changes in traffic from 1968 to 2006 indicate a cyclical trend (Figure 10.1). We can see that the cyclical variation in the number of passengers during this time period is more irregular because of the business cycle. In general, it is harder to predict the cyclical components of time-series data than trend and seasonal variations. The most recent cycle was in 2000 when air traffic started to decline, and this cycle continued until late 2003.

Random Effect

Finally, random factors of a series are short-term, unanticipated and non-recurring factors that affect the values of the series; these factors are part of the natural variability present in all measurements. For example, events such as those of 11 September 2001 are impossible to predict. The trend component analyzes the data over a long period of time, the cyclical component occurs over a medium term, seasonal variations occur in the short term (say, over a year), while random events are unique incidents.

Although there are several different methods that can be used to forecast time series and to take into account the various components of a time series, we shall just discuss the following four major methods:

- moving average
- weighted moving average
- exponential smoothing
- trend analysis.

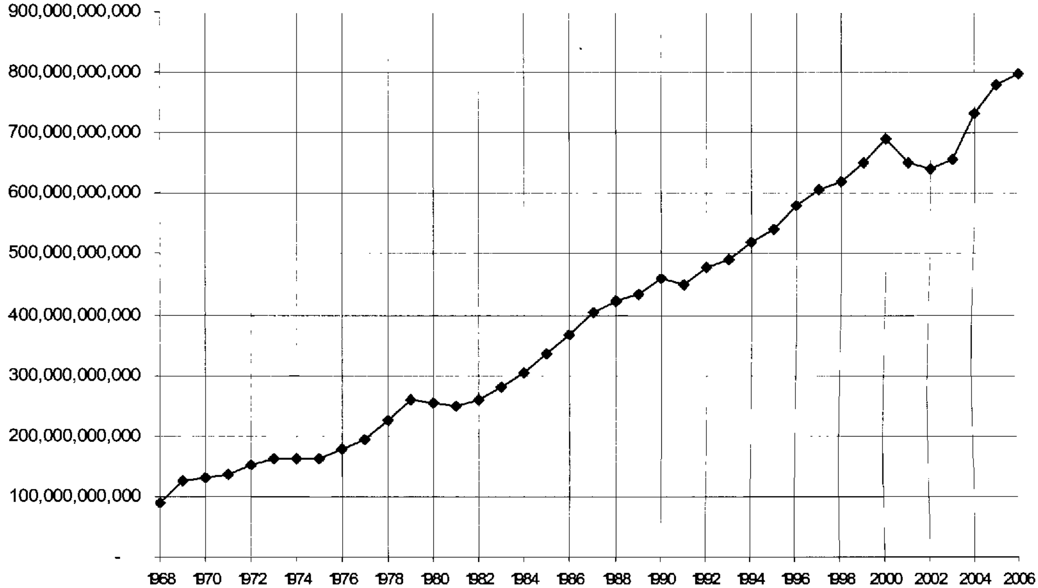

Figure 10.1 US airline industry: revenue passenger miles

Moving average Moving average is a smoothing technique that uses the average of the most recent data values to help forecast the next period. It is a very simple technique that contains the underlying assumption that the most recent values are the best representation of the future. Mathematically, the formula for moving average is:

$$MA = \frac{\sum_{i=1}^{n} X_i}{n}$$

In order to understand the applicability of moving average, consider the following data, which display the historical bookings for the last ten days of a DirectJet flight.

If the goal is to forecast the number of bookings on the eleventh day with a moving average, the first step is to determine n, which represents the number of recent data values to include in the forecast. Since every value is given equal weight in the moving average, the larger the n, the more weight historical values are given. Assuming that a three-day moving average is desired, the three most recent data values (days 8, 9, and 10) are used to help forecast DirectJet's bookings on day 11. The forecast for the eleventh day is:

$$F(11) = \frac{(150+130+160)}{3} = \frac{440}{3} = 146.67$$

Using the last three data values, the forecasted number of bookings on the eleventh day is 146.67, or 147 when rounded. If it is believed that a three-day moving average provides

Table 10.5 Time-series data of bookings for a DirectJet flight

Day	Bookings
1	115
2	100
3	105
4	120
5	135
6	130
7	145
8	150
9	130
10	160
11	?

too much emphasis on the most recent days, then a five-day moving average can be used instead. The five-day moving average provides a forecast for the eleventh day of:

$$F(11) = \frac{(130+145+150+130+160)}{5} = \frac{715}{5} = 143$$

Based on the five-day forecast, the forecasted value of 143 is less than the three-day forecasted value of 146.67. While results will vary depending on the data set, any moving average attempts to smooth out any distortions in the data; this can be extremely useful, especially with highly variable data.

The moving average can easily be calculated in Microsoft Excel using the moving average function in the "Data Analysis Toolpack." With the aid of Microsoft Excel, the three-day and five-day moving average forecasts were created for multiple days to compare the forecasted values and the actual values. These data are presented in Table 10.6.

Note the fact that the three-day moving average forecasts cannot be computed for the first three days and, likewise, the five-day average for the first five days, since these are needed to start the series. The accuracy of the forecasts can be determined by comparing the difference between the actual and forecasted value, and these techniques will be discussed in more detail later in the chapter.

Weighted moving average Weighted moving average (WMA) is similar to moving average. It still uses historical data to provide a forecasted value; however, instead of each value receiving equal weighting, as in moving average, values receive different weightings. For example, in a three-period moving average, each value receives an equal weighting of

Table 10.6 Three-day and five-day moving average forecasts

Day	Bookings	3-Day M.A.	5-Day M.A.
1	115		
2	100		
3	105		
4	120	107	
5	135	108	
6	130	120	115
7	145	128	118
8	150	137	127
9	130	142	136
10	160	142	138
11	?	147	143

1/3. However, weighted moving average enables the forecaster to weight the values as desired. Mathematically, the formula for WMA is presented as:

$$WMA = \sum_{i=1}^{n} W_i \times X_i$$

Using the same data contained in Table 10.5, a weighted moving average for day 11 can be created, assuming that the most recent value receives a 50 per cent weighting, the next most recent value receives a 30 per cent weighting, and the third value receives 20 per cent. On this basis, the forecasted value is:

$$F(11) = 0.5(160) + 0.3(130) + 0.2(150)$$
$$F(11) = 80 + 39 + 30 = 149$$

Based on the designed weightings, the forecasted value for day 11 is 149. However, the forecasted value can change, based on the assigned weightings. Assuming that the most recent value receives an 80 per cent weighting, the second value a 15 per cent weighting, and the final value a 5 per cent weighting, the forecasted value for the eleventh day would be:

$$F(11) = 0.8(160) + 0.15(130) + 0.05(150)$$
$$F(11) = 128 + 19.5 + 7.5 = 155$$

With this new weighting, the forecasted value for the eleventh day is considerably higher because the most recent value received a high weighting. Ultimately, the weightings that are assigned are based on the forecaster's judgment. Therefore, the more experience and expertise the forecaster has, the more likely that the assigned weightings will be accurate. Weightings can be assigned for any number of periods, as long as the total sum of the weightings equals 100 per cent or 1.0.

Exponential smoothing A third smoothing technique that can be used to forecast time-series data is exponential smoothing. Unlike moving average which uses multiple historical values to help forecast, exponential smoothing only uses data from the previous period. Exponential smoothing indirectly takes into consideration previous periods by using the previous period's forecasted value in helping determine the forecasted value. This creates a situation where the weighting for a value gets exponentially smaller as time moves on. The general formula for exponential smoothing is:

$$F_{t+1} = \alpha Y_t + (1 - \alpha) F_t$$

where: F_{t+1} is the forecasted value in the next period
Y_t is the actual value in the previous period
F_t is the forecasted value in the previous period
α is a smoothing constant with values between 0 and 1.

The smoothing constant helps determine what weighting of the forecast value should be based on the actual value from the previous period and the forecasted value. The higher the smoothing constant, the greater the weighting the actual value receives. Like the two previous forecasting methods, the forecaster must make a judgment in assigning the value for the smoothing constant. While higher smoothing constants usually provide more accurate forecasts, the overall objective of the forecast is to be as accurate as possible. Since the formula contains a term on the right-hand side that shows a previously forecast value, the question arises as to where that value will come from for the first observation. The answer to this is that that value comes from the actual value of the first period. This means that no matter what value is picked for the constant, the first value of the forecast will equal the first period of the series. Subsequent values will, of course, differ because the actual and forecast values will differ due to the smoothing constant.

Using the exponential smoothing function in the "Data Analysis Toolpack" from Microsoft Excel, forecasts can be created for DirectJet. Table 10.7 provides forecasts with two different smoothing constants—0.3 and 0.8.

From Table 10.7 it is clear that a smaller smoothing constant provides greater fluctuation in the forecasted value, while a larger constant provides less variability in the forecasts.

Trend analysis The fourth and final time-series method to be investigated is trend analysis. Scatter diagrams and line graphs provide a good first approximation in identify the existence of a trend line between independent and dependent variables. Depending on how closely the points group together, we may be able to identify a trend in the data. Unfortunately, trends are not always easy to see graphically, and there may also be a problem with units. A more quantitative method to identify a trend line is to use regression analysis which attempts to create a linear trend equation to describe the data (Anderson,

Table 10.7 Exponential smoothing forecasts using two different smoothing constants

Day	Bookings	$\alpha = 0.3$	$\alpha = 0.8$
1	115		
2	100	115	115
3	105	105	112
4	120	105	111
5	135	115	112
6	130	129	117
7	145	130	120
8	150	140	125
9	130	147	130
10	160	135	130
11	?	153	136

Sweeney, and Williams, 2006). Such equations can then be used to provide a forecast for a future value. The general form for these equations is as follows:

$$F_t = b_o + b_1 t$$

where: F_t is the forecasted value in period t
b_o is the intercept of the trend line
b_1 is the slope of the trend line.

Hence, in order to calculate the forecast value, the parameters b_0 and b_1 must first be calculated. The formulas for these values are presented here:

$$b_1 = \frac{\sum_{t=1}^{n} tY_t - \frac{(\sum_{t=1}^{n} t \sum_{t=1}^{n} Y_t)}{n}}{\sum_{t=1}^{n} t^2 - \frac{(\sum_{t=1}^{n} t)^2}{n}}$$

$$b_o = \overline{Y} - b_1 \overline{t}$$

where: Y_t is the actual value in period t
n is the number of periods
is the average value of the time series
$\overline{Y}$ is the average value of t.
$\overline{t}$

Based on these formulae, the linear trend line can be constructed. To complete the calculation, however, some additional information is required. Therefore, the original DirectJet problem is expanded in Table 10.8.

Table 10.8 Expanded data set for DirectJet

Day	Bookings	Day Squared
1	115	1
2	100	4
3	105	9
4	120	16
5	135	25
6	130	36
7	145	49
8	150	64
9	130	81
10	160	100
55	1290	385

From the expanded data set, the values required to determine the slope and the intercept can be found as follows:

$$\sum_{t=1}^{10} tY = 7{,}550$$

$$\sum_{t=1}^{10} t = 55$$

$$\sum_{t=1}^{10} Y_t = 1{,}290$$

$$n = 10$$

$$\sum_{t=1}^{10} t^2 = 385$$

$$\bar{Y} = \frac{\sum_{t=1}^{10} Y_t}{n} = \frac{1290}{10} = 129$$

$$\bar{t} = \frac{\sum_{t=1}^{10} t}{n} = \frac{55}{10} = 5.5$$

$$b_1 = \frac{\sum_{t=1}^{n} tY_t - \frac{(\sum_{t=1}^{n} t \sum_{t=1}^{n} Y_t)}{n}}{\sum_{t=1}^{n} t^2 - \frac{(\sum_{t=1}^{n} t)^2}{n}}$$

$$b_1 = \frac{7550 - \frac{(55*1290)}{10}}{385 - \frac{(55)^2}{10}}$$

$$b_1 = \frac{7550 - 7095}{385 - 302.5} = \frac{455}{82.5} = 5.52$$

$$b_o = 129 - 5.52(5.5)$$

$$b_n = 129 - 30.33 = 98.67$$

Using the values of the slope and intercept, the trend line is:

$$F_t = 98.67 + 5.52t$$

Based on this formula, forecasts for the number of bookings for DirectJet can be created by solving the equation. Note that trend analysis only forecasts the trend portion of a time series. Cyclicality, seasonality, and random factors can cause distortions from the trend line. Using the above equation, the forecast for the number of bookings on the eleventh day is:

$$F(11) = 98.67 + 5.52(11)$$
$$F(11) = 98.67 + 60.72 = 159.39$$

Trend analysis can be performed more quickly through computer programs such as Microsoft Excel and SPSS. By graphing the time-series data, a trend line can be fitted to the data and the equation can also be provided. Figure 10.2 displays the trend line, with the computer producing the exact same formula as was calculated above.

Forecast Accuracy

> We soon saw that the helicopter had no future, and dropped it. The helicopter does with great labor only what the balloon does without labor, and is no more fitted than the balloon for rapid horizontal flight. If its engine stops, it must fall with deathly violence, for it can neither glide like the aeroplane or float like the balloon. The helicopter is much easier to design than the aeroplane, but is worthless when done.
>
> Wilbur Wright, 1909

Forecasting is ultimately useful only if the forecasts are reasonably accurate. While the actual accuracy of the forecast is not known until the event has occurred, historical time-series data can be analyzed to give an indication of how well the technique works. In the aviation industry there are many examples of forecasting errors which caused significant loss to the company involved. For example, in 2006 Airbus's parent company, EADS, announced that the company needed to sell 420 A380s to break even—up from its initially announced 270 aircraft.

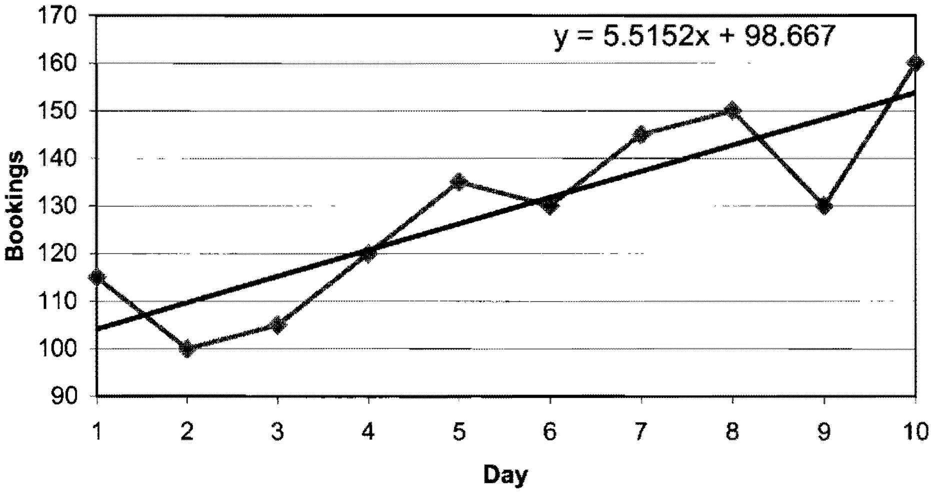

Figure 10.2 Graphical representation of trend line for the time series

The two principal methods of analyzing forecast accuracy are mean squared error (MSE) and mean absolute deviation (MAD).

Mean squared error averages the squared difference between the actual value and the forecasted value. The values are squared to eliminate the effect of negative errors canceling out positive errors (similar to the squaring of the deviations from the mean that was used to calculate the variance), and also to give greater weight to larger errors. The basic formula for mean squared error can be written as:

$$MSE = \frac{\sum_{t=1}^{n} (Y_t - F_t)^2}{n}$$

Tables 10. 9, 10.10, and 10. 11 provide the mean squared error calculation for the different forecasting methods employed in the DirectJet example. Table 10.9 provides the mean squared error calculation for both the three-day and five-day moving averages. Based on the results, the three-day moving average appears more accurate since its MSE value is less than the five-day MSE.

Table 10.10 provides the mean squared error for the weighted average forecasts used in the DirectJet example. Both were three-day moving averages; however, the first forecast used a 50/30/20 weighting, while the second forecast used an 80/15/5 weighting. Based on the mean squared error, the more evenly distributed forecast provides the most accurate forecast for this particular time series.

Table 10.11 provides MSE for exponential smoothing forecasts with smoothing constants of both 0.3 and 0.8. Based on all these calculations, the exponential smoothing forecast with a smoothing constant of 0.3 provided the most accurate forecast, since it had the lowest mean squared error.

Table 10.9 Mean squared error calculation for moving average forecasts

Day	Bookings	3-Day M.A.	Forecast Error	Squared Forecast Error	5-Day M.A.	Forecast Error	Squared Forecast Error
1	115						
2	100						
3	105						
4	120	107	13	178			
5	135	108	27	711			
6	130	120	10	100	115	15	225
7	145	128	17	278	118	27	729
8	150	137	13	178	127	23	529
9	130	142	-12	136	136	-6	36
10	160	142	18	336	138	22	484
			MSE	274		MSE	401

Table 10.10 Mean squared error calculation for weighted moving average forecasts

Day	Bookings	50/30/20 Forecast	Forecast Error	Squared Forecast Error	80/15/5 Forecast	Forecast Error	Squared Forecast Error
1	115						
2	100						
3	105						
4	120	109	12	132	112	8	60
5	135	106	30	870	102	33	1106
6	130	116	15	210	109	21	452
7	145	127	19	342	123	22	495
8	150	136	15	210	135	15	233
9	130	139	-9	72	133	-3	11
10	160	144	17	272	145	15	225
			MSE	301		MSE	369

Table 10.11 Mean squared error calculation for exponential smoothing forecasts

Day	Bookings	$\alpha = 0.3$	Forecast Error	Squared Forecast Error	$\alpha = 0.8$	Forecast Error	Squared Forecast Error
1	115						
2	100	115	-15	225	115	-15	225
3	105	105	1	0	112	-7	49
4	120	105	15	230	111	9	88
5	135	115	20	382	112	23	507
6	130	129	1	1	117	13	169
7	145	130	15	233	120	25	646
8	150	140	10	92	125	25	642
9	130	147	-17	293	130	0	0
10	160	135	25	618	130	30	913
			MSE	230		MSE	360

Another measure of forecasting accuracy is mean absolute deviation. Mean absolute deviation finds the average of the absolute value of the deviations. Since the deviations are not squared, large deviations are not given extra weight. The general formula for mean absolute deviation (MAD) is:

$$MAD = \frac{\sum |Y_t - F_t|}{n}$$

Mean absolute deviation can also be calculated for the various time-series forecasts in the DirectJet example. The absolute value of the forecast error is found in Microsoft Excel by using the abs function. Table 10.12 provides the mean absolute deviation for the moving average forecasts, Table 10.13 for the weighted moving average forecasts, and Table 10.14 for the exponential smoothing forecasts.

From the exponential smoothing calculations, the exponential smoothing forecast with a smoothing constant of 0.3 appears once again to be the most accurate forecasting method for these particular time-series data. In this situation, both mean squared error and mean absolute deviation picked the same forecasting method as the most accurate; however, this will not always hold true as it will depend on the data. Both are commonly used in practice: the MSE gives more weight to large errors, while the MAD is easier to interpret. Based on the measures of accuracy, an appropriate forecasting method can be chosen. Using this method, forecasts can be created for future periods. While the measures of accuracy highlight the most accurate forecasting method based on historical information, the data provide an ongoing repository of continually growing measurements that can be refined and updated. As this happens, new forecasting methods may be substituted for the original selection. This means

Table 10.12 Mean absolute deviation calculation for moving average forecasts

Day	Bookings	3-Day M.A.	Forecast Error	Absolute Forecast Error	5-Day M.A.	Forecast Error	Absolute Forecast Error
1	115						
2	100						
3	105						
4	120	107	13	13			
5	135	108	27	27			
6	130	120	10	10	115	15	15
7	145	128	17	17	118	27	27
8	150	137	13	13	127	23	23
9	130	142	-12	12	136	-6	6
10	160	142	18	18	138	22	22
			MAD	16		MAD	19

Table 10.13 Mean absolute deviation calculation for weighted moving average forecasts

Day	Bookings	50/30/20 Forecast	Forecast Error	Absolute Forecast Error	80/15/5 Forecast	Forecast Error	Absolute Forecast Error
1	115						
2	100						
3	105						
4	120	109	12	12	112	8	8
5	135	106	30	30	102	33	33
6	130	116	15	15	109	21	21
7	145	127	19	19	123	22	22
8	150	136	15	15	135	15	15
9	130	139	-9	9	133	-3	3
10	160	144	17	17	145	15	15
			MAD	16		MAD	17

Table 10.14 Mean absolute deviation calculation for exponential smoothing forecasts

Day	Bookings	$\alpha = 0.3$	Forecast Error	Absolute Forecast Error	$\alpha = 0.8$	Forecast Error	Absolute Forecast Error
1	115						
2	100	115	-15	15	115	-15	15
3	105	105	1	1	112	-7	7
4	120	105	15	15	111	9	9
5	135	115	20	20	112	23	23
6	130	129	1	1	117	13	13
7	145	130	15	15	120	25	25
8	150	140	10	10	125	25	25
9	130	147	-17	17	130	0	0
10	160	135	25	25	130	30	30
			MAD	13		MAD	16

that the forecaster must use some judgment in choosing which measure of accuracy to use, how much data should be incorporated, and how often to use it. Regardless of the choices, the ultimate goal of the forecaster is to provide the most accurate forecasts available.

REGRESSION ANALYSIS

The other principal quantitative forecasting method is regression analysis. Regression analysis assumes a casual relationship between the dependent and independent variables. The specific two-variable linear regression model is:

$$Y_i = \beta_0 + \beta_1 \times X_i + \varepsilon_i$$

where: Y_i is the dependent variable
β_0 *and* β_1 are the coefficients of the regression line (also known as the intercept and slope)
X_i is the independent variable
ε_i is the predictor error, or so-called residual.

A dependent variable is a variable that relies on other factors and variables, while an independent variable has a value which does not rely on any other factors (Ovedovitz, 2001). In order to understand the applicability of regression analysis, first consider the following data set: Table 10.15, with consumption (C) and income (Y) values, and the corresponding Figure 10.2.

Table 10.15 Data set for the relation between consumption and income (1)

C	Y
100	0
195	100
290	200
385	300
480	400
575	500
670	600
765	700
860	800
955	900
1050	1000
1145	1100
1240	1200
?	1300

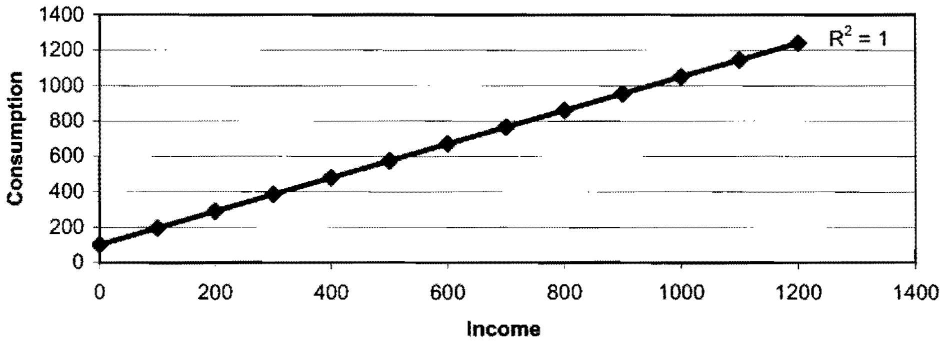

Figure 10.3 Graphical relation between consumption and income (1)

Using both the data set and the graph, it is very clear that there is a simple linear relationship between consumption and income. In this relationship, because consumption is dependent on income, consumption is the dependent variable and income the independent variable. Here, forecasting is fairly easy, since the pattern is obvious from the graph. In this case, for every 100-unit increase in income, consumption increases by 95 units. Based on the linear relationship, it is easy to forecast the level of consumption for an income level of 1,300. For this example, and using the graph above, consumption would be 1,335 for an income level of 1,300. Using the formulas from the trend analysis to calculate the slope and the intercept (the best-fit formulas are the same for both techniques because they both assume a linear relationship between the variables), it can be easily determined that the slope of the function is 0.95 (95/100) and the y-intercept value is 100. This creates the function below, which can be used for future forecasts:

$$C = 100 + 0.95Y$$

In this example, the relationship between consumption and income is very straightforward since they form a perfect linear relationship. However, real-world data are never perfectly correlated, since random events and other factors cause distortions in the relationship. Consider the same relationship where the values are slightly modified.

In this situation there is no clear linear relationship between consumption and income. However, when the data from Table 10.16 are plotted in Figure 10.4, the points lie somewhat randomly, but with a general upward trend. Since the previous trend analysis method would not provide an accurate forecast, regression analysis needs to be employed. As Figure 10.4 displays, regression analysis fits a trend line for the data points—or more simply, it calculates a quantifiable linear relationship for the various data points. It should be noted that regression analysis can also estimate exponential, quadratic, or other relationships, depending on the general trend of the data points, but such estimating techniques usually require a functional form transformation that is linear in the parameters.

In this example, since the data points have a general linear trend, linear regression analysis should be chosen. However, the linear trend line is not perfectly accurate because all the values do not lie directly on the trend line. Figure 10.4 is an example of a scatter diagram, which essentially plots the points on a Cartesian plane with the dependent variable plotted on the y-axis and the independent variable plotted on the x-axis. Prior to performing regression analysis, a scatter plot should be created to help understand the nature of the data points. From this the appropriate regression analysis (linear, exponential) can be chosen, in order to provide the most accurate forecast of the dependent variable.

Table 10.16 Data set for the relation between consumption and income (2)

C	Y
100	0
220	100
300	200
450	300
700	400
590	500
800	600
400	700
950	800
700	900
1020	1000
1200	1100
1240	1200
?	1300

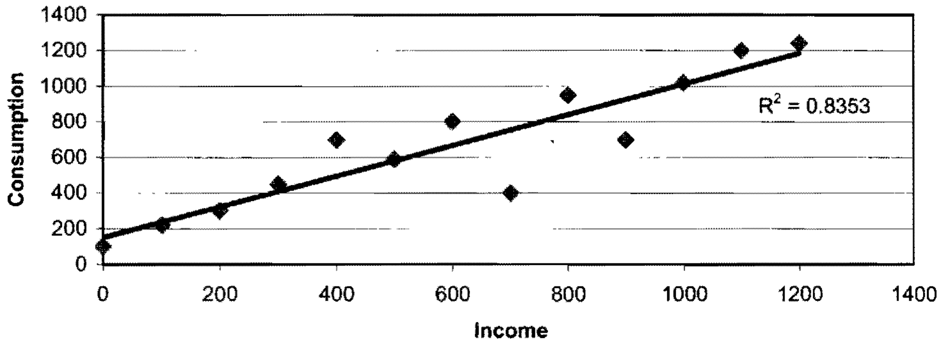

Figure 10.4 Graphical relation between consumption and income (2)

While all types of regression analysis perform the same function, there are many different methods for creating the trend line. The most common method is ordinary least squares, which minimizes the squared value of the residuals. Although it would seem more natural to minimize the sum of the errors, we run into the same problem that we observed earlier with the variance, namely that very large positive errors (deviations above the line) would tend to cancel out very large negative errors (deviations below the line). Therefore, the errors are squared to eliminate this problem. This, of course, means that the larger errors carry more weight in the procedure.

Thus, the formula for ordinary least squares regression can be stated as:

$$OLS = \min \sum_{i=1}^{n} e_i^2 = \min \sum_{i=1}^{n} (y_i - \hat{y})^2$$

where: e^2 is the residual
y is the actual value
$\hat{y}$ is the forecasted value.

Therefore, the residual is simply the difference between the actual value and the forecasted value along the trend line. In the first example, the residuals would equal zero since all the data points were located directly on the linear curve. However, in the second example, a residual value exists because there is a vertical gap between the data points and the trend line. Since the derivation of the ordinary least squares trend line (that is, the values of b_0 and b_1 that define the forecast line that minimizes the sum of the squared residuals) involves some elementary calculus, only the resulting values are presented here. In any event, the formulas that are derived are exactly the same as those given for the best-fit trend line discussed earlier in the chapter.

For example, by using the information in Table 10.17, the short-run consumption functions have been estimated as follows:

$$C = 150 + 0.8615Y$$

where: b0 = 150 and b1 = .8615
C= consumption
Y= income.

Since the basic formula for a residual is the difference between the actual and forecasted value, we can calculate the difference for each observation as follows:

$$\text{Residual} = e^2{}_i = (C_i - \overline{C})^2$$

These values are contained in Table 10.17.

Because ordinary least-square regression minimizes the sum of the residuals, no other linear trend line would produce a sum of the residuals less than 266,385. The accuracy of the forecast is ultimately determined by how large the residuals are. Forecasts with extremely large residuals imply that the spread between the forecasted value and the actual value is wide, and that the forecast is therefore not as accurate. However, depending on the nature of the data, any trend line may be extremely accurate or not accurate at all: measures of accuracy are discussed in the next section.

Goodness of Fit

The "goodness of fit" of the regression model evaluates the strength of the proposed relationship between the dependent and independent variables and can be measured in

Table 10.17 Residual values for forecasts of consumption

Y	X	$\hat{Y}$	e $(C - \hat{C})$	e2
100	0	150	-50	2,500
220	100	236	-16	261
300	200	322	-22	497
450	300	408	42	1,726
700	400	495	205	42,189
590	500	581	9	86
800	600	667	133	17,716
400	700	753	-353	124,644
950	800	839	111	12,277
700	900	925	-225	50,783
1020	1000	1012	9	
1200	1100	1098	102	10,476
1240	1200	1184	56	3,158
			$\sum$e2	266,385

several ways (Doane and Seward, 2007). The test that is most often used for the accuracy of any given regression is the coefficient of determination, or R-squared. The coefficient of determination measures the percentage of variability in the dependent variable that the independent variable(s) explains. It is usually used to help state the degree of confidence one has in the forecast. The coefficient of determination ranges from 0 to 1, with 1 being a perfectly accurate forecast (similar to the first example provided) and 0 being a completely inaccurate forecast. Using the following formula for the coefficient of determination, the R-square value is computed in Table 10.18.

$$R^2 = \frac{\sum_{i=1}^{n} (\hat{Y}_i - \overline{Y})^2}{\sum_{n=1}^{n} (Y_i - \overline{Y})^2}$$

where: R^2 is the coefficient of determination
$\hat{Y}_i$ is the forecasted value
Y is the actual value
$\overline{Y}$ is the mean of the actual values.

From this calculation, the coefficient of determination for the linear regression model is 0.8352. This shows that 83.52 per cent of the variability of the dependent variable, C, is explained by the independent variable, Y or, in simpler terms, the model is roughly 84 per

Table 10.18 Coefficient of determination calculation

C	I	$\hat{C}$	e (C – $\hat{C}$)	e2	$\hat{C}$ - Cμ	($\hat{C}$ - Cμ)2	C - Cμ	(C - Cμ)2
100	0	150	-50	2,500	-517	267,209	-567	321,402
220	100	236	-16	261	-431	185,565	-447	199,740
300	200	322	-22	497	-345	118,765	-367	134,633
450	300	408	42	1,726	-258	66,808	-217	47,056
700	400	495	205	42,189	-172	29,695	33	1,094
590	500	581	9	86	-86	7,426	-77	5,917
800	600	667	133	17,716	0	0	133	17,709
400	700	753	-353	124,644	86	7,418	-267	71,248
950	800	839	111	12,277	172	29,679	283	80,133
700	900	925	-225	·	258	66,784	33	1,094
1020	1000	1012	9	72	345	118,733	353	124,663
1200	1100	1098	102	10,476	431	185,526	533	284,171
1240	1200	1184	56	3,158	517	267,162	573	328,417
667	Cμ		∑e2	266,385	∑($\hat{C}$ - Cμ)2	1,350,772	∑(C - Cμ)2	1,617,277
							R2	0.8352

cent accurate. Although the process of calculating the coefficient of determination is lengthy, almost all statistical programs display the R-square value in the regression output.

In order to better understand the accuracy of regression analysis, consider the same example, but with modified values of consumption. Table 10.19 provides the new data set, and Figure 10.5 displays the new scatter plot with the trend line.

In this new scenario, the data points have less of a linear relationship, increasing the summation of the residual values. While ordinary least-squares regression provides the best possible trend line for the data points, the increase in the size of the residuals over the previous example indicates that this trend line is not as accurate as the previous one. This is confirmed through a coefficient of determination value of 0.1376, which is significantly less than the previous 0.8352. The forecast for this third data set yields only an R-square value of 0.1376, which means that this forecast is not very powerful and has very little usefulness. While the coefficient of determination determines the accuracy of the overall regression, it is also possible to determine the statistical relevance of separate independent variables when there is more than one independent variable in the regression. This topic is discussed in the next few sections.

Performing Regression Analysis

For ease of exposition, the previous examples contained only one independent variable; however, most regressions will include multiple independent variables, since the dependent variable can rarely be explained by one factor. For example, demand for

Table 10.19 Data set for the relation between consumption and income (3)

C	I
700	0
220	100
300	200
450	300
700	400
590	500
400	600
400	700
950	800
500	900
1020	1000
1300	1100
200	1200
?	1300

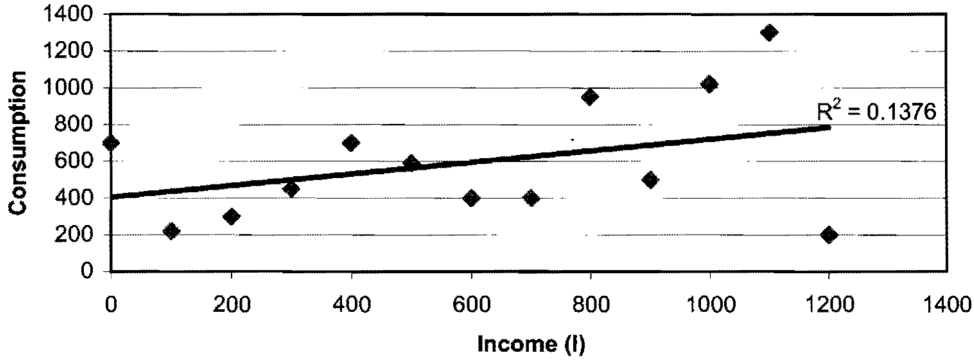

Figure 10.5 Graphical relation between consumption and income, with trend line

airline services includes a host of factors such as ticket price, income, competitor's price, seasonality, and customer service. It would be nearly impossible to perform ordinary least-squares regression by hand for multiple independent variables, but computer programs, such as Microsoft Excel and SPSS, allow regressions to be performed quickly and easily.

In order to understand applied regression analysis more completely, and to identify the important factors to analyze when interpreting regression results, we introduce a concrete example. Consider the demand for air travel between Orlando and Los Angeles. Four independent variables are used to help determine the demand. These are: average ticket prices, income, seasonality, and the presence of a random one-off event, such as 9/11. Prior to any forecasting, a hypothesis should be created to help identify the expected relationship between the dependent and independent variables. This hypothesis can then be used to help determine whether the results from the regression analysis are accurate. In this example, ticket price should have a negative coefficient, income a positive coefficient, seasonality could potentially have either positive or negative depending on the season, and 9/11 should have a negative value. The data set used in the regression analysis is contained in Table 10.20.

In order to forecast demand, historical data on the number of passengers flying between Orlando and Los Angeles need to be found. O & D data use a 10 per cent sample of total bookings to help quantify the total number of passengers flying the city-pair, regardless of whether they are flying on a nonstop flight or on connecting flights. Since it is difficult to determine which income needs to be measured, gross domestic product (GDP) is a reasonable proxy for income. Finally, the average ticket price for all travelers is determined through O & D data for each quarter. Quarterly data was used from 1998 in order to provide a sufficient number of observations. For regression analysis to be accurate, an appropriate number of observations is required. While controversial, a minimum of 30 observations is usually safe.

Dummy or Binary Variables

To capture the effects of a random event such as 9/11, or the impacts of qualitative events such as seasonality, we can apply dummy variables.[6] A dummy variable is an independent variable that takes on only two values: 1 or 0. Dummy variables, or binary or categorical variables, require no additional economic data (Barreto and Howland 2006).

Many studies, such as Anderson and Mittal (2000) and Brandt (1987), use regression analysis with dummy variables to identify the actual nature of the relationship between the dependent variable and independent (explanatory or exogenous) variables. The variable only determines if the presence of a factor exists or does not exist. In the case of seasonality, three unique independent variables can be created. The Q1 dummy variable takes on a value of 1 during the first quarter of every year, and the value of 0 for every other quarter. Similar dummy variables were created for the second and third quarters. A fourth seasonal dummy variable is not needed since the fourth quarter is acting as the baseline for all the other quarters,[7] so that, for example, the coefficient in the regression equation for the first quarter would measure the additional (or smaller) quantity demanded over the fourth quarter. With this set-up the first three quarters are being compared to the fourth quarter. Of course, we could have excluded any of the four quarters, and then the regression coefficients on the remaining three would be compared to the quarter excluded.

The other dummy variable used is to take into consideration the one-time shift in demand caused by the tragic events of 9/11. Since the events of 9/11 affected demand for air travel, all quarters following and including the third quarter of 2001 received a value of 1 to identify this impact. In this case, the excluded variable is the quarters that were not felt to be affected by the events of 9/11—that is, all prior quarters.

6 Also known as categorical variables.

7 In fact, if all the classes for a binary variable are included in a regression equation that includes a constant, the regression cannot be estimated since a linear dependence exists between the independent variables. This is the so-called dummy variable trap. See Hanushek and Jackson (1977, p. 104) for a more complete description of this.

Table 10.20 Data set for forecasting demand for the Orlando–Los Angeles flight

Quarter	Demand (# of Passengers)	Income (GDP in billions)	Avg. Price	Q1	Q2	Q3	9/11
1998 - Q1	50,060	$8,586.70	$242.83	1	0	0	0
1998 - Q2	57,710	$8,657.90	$226.55	0	1	0	0
1998 - Q3	61,910	$8,789.50	$213.19	0	0	1	0
1998 - Q4	56,290	$8,953.80	$214.97	0	0	0	0
1999 - Q1	55,200	$9,066.60	$228.14	1	0	0	0
1999 - Q2	63,680	$9,174.10	$214.77	0	1	0	0
1999 - Q3	61,560	$9,313.50	$208.83	0	0	1	0
1999 - Q4	60,000	$9,519.50	$195.95	0	0	0	0
2000 - Q1	58,930	$9,629.40	$212.52	1	0	0	0
2000 - Q2	69,190	$9,822.80	$203.16	0	1	0	0
2000 - Q3	66,240	$9,862.10	$201.28	0	0	1	0
2000 - Q4	68,740	$9,953.60	$208.42	0	0	0	0
2001 - Q1	60,230	$10,021.50	$237.66	1	0	0	0
2001 - Q2	63,090	$10,128.90	$209.95	0	1	0	0
2001 - Q3	60,670	$10,135.10	$181.01	0	0	1	1
2001 - Q4	46,470	$10,226.30	$180.81	0	0	0	1
2002 - Q1	45,330	$10,333.30	$220.36	1	0	0	1
2002 - Q2	49,780	$10,426.60	$210.86	0	1	0	1
2002 - Q3	51,950	$10,527.40	$191.95	0	0	1	1
2002 - Q4	53,220	$10,591.10	$208.60	0	0	0	1
2003 - Q1	50,610	$10,705.60	$211.52	1	0	0	1
2003 - Q2	59,590	$10,831.80	$192.99	0	1	0	1
2003 - Q3	62,300	$11,086.10	$175.84	0	0	1	1
2003 - Q4	63,750	$11,219.50	$176.68	0	0	0	1
2004 - Q1	63,980	$11,430.90	$181.95	1	0	0	1
2004 - Q2	76,780	$11,649.30	$165.59	0	1	0	1
2004 - Q3	76,930	$11,799.40	$158.13	0	0	1	1
2004 - Q4	74,620	$11,970.30	$162.18	0	0	0	1
2005 - Q1	74,480	$12,173.20	$178.73	1	0	0	1
2005 - Q2	79,780	$12,346.10	$178.10	0	1	0	1
2005 - Q3	78,670	$12,573.50	$179.97	0	0	1	1
2005 - Q4	69,640	$12,730.50	$191.06	0	0	0	1
2006 - Q1	70,530	$13,008.40	$192.57	1	0	0	1
2006 - Q2	81,080	$13,197.30	$208.06	0	1	0	1

Source: Compiled by the authors.

Once the data have been collected and placed in a statistical computer program, the program will return the values for the regression. Most programs require the user to define which variable is the dependent variable and which are the independent variables. In this example, the number of passengers is the dependent variable, and the remaining variables are all independent. The regression is then run, and the output is displayed. While the output varies from program to program, they all contain the same basic characteristics. For our example, all regression output is from the SPSS.[8]

The first major chart displayed in all regression output is a summary of the model. The model summary from the SPSS for the regression is given in Table 10.21. Probably the most important statistic contained in any model summary is the R-square value or coefficient of determination, which, as mentioned previously, determines the percentage of variation in the dependent variable that is explained by the independent variables. In this model, approximately 88 per cent of the demand for travel between Orlando and Los Angeles can be explained by the independent variables.

The adjusted R-square value is similar to the R-square value, but takes into consideration the degrees of freedom of the model. By way of definition, the degrees of freedom are the number of observations beyond the minimum needed to calculate a regression statistic (Hirschey, 2006). They are determined by taking the total number of observations minus the number of independent variables. Higher degrees of freedom are created through more observations or less independent variables. In this example, the degree of freedom is 28.[9] Since a forecast is usually more accurate with an increased number of observations or with a larger number of independent variables, the adjusted R-square value takes this into account. Therefore, the ordinary R-square value is adjusted downwards to account for the degrees of freedom in the particular model. Models that contain low degrees of freedom receive the greatest difference between the ordinary R-square and the adjusted R-square value. Since this model has a relatively high degree of freedom (28), the difference between the adjusted R-square and the ordinary R-square is not large.

Autocorrelation

Another major statistic to analyze in the model summary output is the Durbin–Watson statistic.[10] This measures autocorrelation, which can severely distort the accuracy and

Table 10.21 Model summary of the demand forecast for Orlando–Los Angeles flights from SPSS

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.940[a]	.883	.858	3702.89391	1.885

a = predictors: (constant), 9/11 dummy variable; Q2 seasonality; Q3 seasonality; Q1 seasonality; average ticket price; GDP.
b = dependent variable: number of passengers.

8 The SPSS (originally, Statistical Package for the Social Sciences) program is used for statistical analysis.
9 Degrees of freedom = number of observations—number of independent variables = 34—6 = 28.
10 It should be noted that the Durbin–Watson statistic is not displayed in the regression output obtained through Microsoft Excel.

significance of the regression model. Autocorrelation occurs when the residuals are not independent, and have an underlying trend that violates one of the major underlying assumptions used in performing regression analysis (that is, in the derivation of the parameters of ordinary least squares, it is assumed that the error terms or residuals are independent of each other). While the Durbin–Watson statistic detects autocorrelation, the residuals can also be plotted against time to detect whether any patterns exist in the residuals. Potential patterns that could exist include linear lines, fanning, or cyclical movements where the residual alternates from positive to negative. Figure 10.6 provides the residual plot for the Orlando to Los Angeles regression.

Since Figure 10.6 does not display any trend in the residuals, it is safe to say that autocorrelation does not exist in the regression. This is confirmed by a Durbin–Watson statistic of 1.885. Durbin–Watson statistics can range from 0 to 4, although only values less than 1.5 or greater than 2.5 suggest that autocorrelation may exist in the regression. Therefore, the Durbin–Watson statistic of 1.885 falls within the acceptable range of 1.5 to 2.5.

The second major table contained in all regression outputs is an ANOVA (analysis of variance) table (see the ANOVA table for the Orlando to Los Angeles demand forecast in Table 10.22). The ANOVA table provides the overall significance of the regression equation. As might be expected, there is a direct mathematical relationship between R-squared and the ANOVA F value for the overall significance of the regression.[11] The difference between them is the fact that the F statistic allows us to pick a level of significance for the overall equation and compare this to a predetermined F distribution. Put more simply, it helps determine if the model is a sound representation of reality, or if the sample data is just an abnormality. This is accomplished by comparing the F-statistic of the regression to a predetermined level of significance. Conventional levels of significance are ordinarily set at .90, .95 and .99 and these mean, respectively, that we can be 90 per cent, 95 per cent, and 99 per cent sure that our regression results are due to a true relationship between the independent and dependent

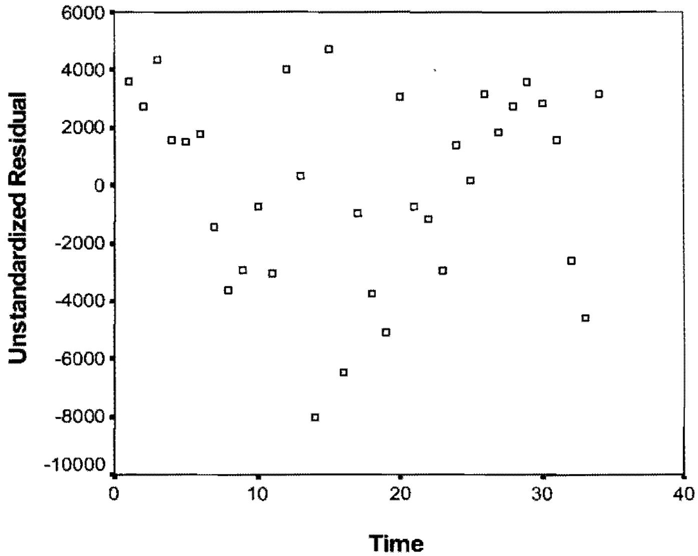

Figure 10.6 Residuals of the regression plotted against time from SPSS

11 $F = R^2 / K-1 / (1-R^2) / N-K$ where K stands for the number of independent variables and N stands for the number of observations. For a further discussion of this see Hanushek and Jackson (1977, pp. 127 and 128).

variables and not due to random chance. Most statistics textbooks contain complete tables of F distributions against which the regression F value can be compared; however, it is also true that most computer programs for regression contain (as part of the output) the level of significance of the independent variables for the given number of observations and degrees of freedom of the specific regression. For example, the level of significance might be reported at .001 or .02, and this means, respectively, that we can be 99.9 per cent and 98 per cent sure that our results are not due to chance.

The third major table contained in all regression output is a table of coefficients. This is displayed for the demand forecast from Orlando to Los Angeles in Table 10.23. The coefficients table allows the researcher to construct a linear equation that can be used for forecasting, and it also determines whether the individual variables are statistically significant. The first column of the coefficients table lists all the independent variables used in the analysis, plus the constant. The constant term is usually interpreted as the value of the dependent variable when all the other independent variables are set to zero. Columns 2 and 4 both display values for the coefficients. The standardized values (column 4) are generally used to compare the respective size of the impacts of the independent variables on the dependent variable. This is accomplished by calculating them in standardized units—that is, the standardized coefficient is the unstandardized value of the coefficient multiplied by the ratio of the standard deviation of the independent variable to the standard deviation of the dependent variable. Therefore, a standardized coefficient of 1.14, as the one for GDP, means that a 1.0 standard deviation change in the independent variable will lead to a 1.14 standard deviation change in the dependent variable. Similar interpretations apply to the other standardized coefficients. But, since the unstandardized values are the coefficients that are directly applicable to forecasting actual values, the unstandardized beta values are the coefficients that are used in the forecast equation. However, and as a final step prior to forming a demand equation, each independent variable needs to be tested to see if it is statistically significant.

The t-statistic is similar to the F-statistic discussed earlier, except that it applies to a single individual variable rather than to the whole (or some subset) of the independent variables. The t-statistic is a measure of how accurate a statistical estimate is. More specifically, a t value is calculated for each independent variable and this value is compared to a standardized t distribution. At this point, a probability statement can be made about the significance (at some predetermined level of confidence) of the independent variable. For example, if the predetermined level of significance is .90 or .95 and the t value for the independent variable selected exceeds the t value for the standardized table (at the degrees of freedom for the specific regression), then we can say that we are 90 per cent or 95 per cent sure that

Table 10.22 ANOVA for the demand forecast for the Orlando–Los Angeles flight, from SPSS

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2.81E+09	6	467649874.1	34.107	.000[a]
	Residual	3.70E+08	27	13711423.28		
	Total	3.18E+09	33			

a = predictors: (constant), 9/11 dummy variable; Q2 seasonality; Q3;seasonality; Q1 seasonality; average ticket price; GDP.
b = dependent variable: number of passengers.

the relationship between the individual variable and the dependent variable is not due to chance. Therefore, a significant t value indicates that the variable in question influences the dependent variable while controlling for other explanatory variables. Quantitatively, the t-statistic contained in column five is simply the unstandardized coefficient divided by the standard error of the coefficient. For example, the t-statistic for the independent variable GDP is found by dividing the beta value of 8.586 by the standard error of .869. This produces a t-statistic of 9.880.

$$t_i = \frac{b_i}{S_{b_i}} = \frac{8.586}{0.869} = 9.88$$

Generally speaking, if an independent variable passes the predetermined t-test, then it should be included in the model; however, if the variable fails the t-test, then it should be considered for exclusion from the model, unless there are strong theoretical reasons to include the variable or there is a clear problem of multicollinearity (discussed in the next section). As a rule of thumb, if the value of a parameter is more than twice the size of its corresponding standard deviation (error), we can conclude, under a two-tailed test, that the estimated coefficient is significantly different from zero at 5 per cent confidence level. Furthermore, if the estimated coefficient is greater than three times of the estimated standard error, we can conclude the estimated value is significantly different from zero at a 1 per cent level of significance.

Another potential major problem that Table 10.23, highlights is multicollinearity which occurs when two or more independent variables are highly correlated with each other. In the limit, if two independent variables are perfectly correlated, then the estimates of the coefficients cannot be computed. Intuitively, the problem arises because the regression cannot separate the effects of the perfectly correlated independent variables. Quantitatively, it arises because there is a term in the denominator for the variance of the individual independent variables that contains the correlation factor between the independent variables. As this term approaches 1 (perfect correlation), the variances of both of the independent variables approach infinity. And, as we have seen above, the t-statistic is calculated by dividing the numerical value of the coefficient by its standard deviation. Since the standard deviation is simply the square root of the variance, then the larger the variance, the larger the standard deviation and the smaller the t value. Thus, a high degree of multicollinearity between independent variables can cause a low level of significance for either one or both of the independent variables

Table 10.23 Coefficients' significance for the demand forecast for the Orlando–Los Angeles flight from SPSS

Model		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	23409.230	16123.163		1.452	.158		
	GDP	8.586	.869	1.149	9.880	.000	.319	3.131
	Average Ticket Price	-211.815	52.474	-.456	-4.037	.000	.339	2.952
	Q1 Seasonality	778.912	2032.159	.036	.383	.705	.502	1.993
	Q2 Seasonality	5218.855	1857.769	.238	2.809	.009	.600	1.666
	Q3 Seasonality	3840.672	1869.170	.169	2.055	.050	.642	1.559
	9/11 Dummy Variable	-19971.5	2216.165	-1.017	-9.012	.000	.339	2.950

a = dependent variable: number of passengers.

Two different methods can be used to detect for multicollinearity. SPSS contains collinearity diagnostics in the coefficients table, and these are contained in the last two columns of Table 10.23. The VIF statistic helps detect multicollinearity and, while the threshold of an acceptable VIF values varies (similar to confidence levels discussed earlier), a conventionally accepted level is that a VIF statistic above 4 indicates the presence of high multicollinearity.[12] The tolerance statistic is simply the inverse of the VIF score (1/VIF); therefore, smaller tolerance values indicate higher degrees of multicollinearity.

The other method for detecting a high degree of correlation between independent variables is to simply create a correlation matrix. Table 10.24 displays a correlation matrix for all the hypothesized independent variables in the example regression.

In Table 10.24 the correlations between all the independent variables are presented. The key statistic is the Pearson correlation statistic, and any correlations greater than 0.90 are of concern. While the correlation between GDP and the 9/11 dummy variable is sizeable at 0.783, both variables are still highly significant from their independent t-tests (see Table 10.23) so it is clearly not enough to reject either variable from the regression analysis. Therefore, based on both the collinearity diagnostics and the correlation matrix, multicollinearity does not appear to be a problem in this particular demand forecast.

If multicollinearity is found to be a problem in a particular regression, then conventional methods for dealing with the problem are acquiring more data or eliminating one or more of the highly collinear independent variables. Since it is rarely possible to acquire more data for a given regression (due to time constraints and so forth), attention shifts to the elimination of variables. If all the variables are still significant at conventional levels of significance, then it is generally advisable to retain the original model, since it was our best initial theoretical formulation of the relationship. If, on the other hand, one or more of the collinear variables are not significant at conventional levels, then consideration should be given to dropping the non-significant variable and rerunning the regression. In

Table 10.24 Correlation matrix for independent variables from SPSS

		GDP	Average Ticket Price	Q1 Seasonality	Q2 Seasonality	Q3 Seasonality	9/11 Dummy Variable
GDP	Pearson Correlation	1	-.659**	-.023	.042	-.039	.783**
	Sig. (2-tailed)	.	.000	.895	.811	.828	.000
	N	34	34	34	34	34	34
Average Ticket Price	Pearson Correlation	-.659**	1	.370*	.062	-.272	-.668**
	Sig. (2-tailed)	.000	.	.031	.729	.119	.000
	N	34	34	34	34	34	34
Q1 Seasonality	Pearson Correlation	-.023	.370*	1	-.360*	-.333	-.040
	Sig. (2-tailed)	.895	.031	.	.036	.054	.823
	N	34	34	34	34	34	34
Q2 Seasonality	Pearson Correlation	.042	.062	-.360*	1	-.333	-.040
	Sig. (2-tailed)	.811	.729	.036	.	.054	.823
	N	34	34	34	34	34	34
Q3 Seasonality	Pearson Correlation	-.039	-.272	-.333	-.333	1	.041
	Sig. (2-tailed)	.828	.119	.054	.054	.	.816
	N	34	34	34	34	34	34
9/11 Dummy Variable	Pearson Correlation	.783**	-.668**	-.040	-.040	.041	1
	Sig. (2-tailed)	.000	.000	.823	.823	.816	.
	N	34	34	34	34	34	34

* Correlation is significant at the 0.05 level (2-tailed).
** Correlation is significant at the 0.01 level (2-tailed).

12 Variance inflation factors (VIF) is a statistics used to measuring the possible collinearity of the explanatory variables.

this case the researcher is implicitly assuming that the two highly collinear variables are providing the same information with respect to the dependent variable.

Once all the regression issues have been checked, the final step is to quantify the demand function so that forecasts can be created. Although computer statistical packages provide a wide array of regression results, the default regression performed is the linear ordinary least-square regression discussed earlier in the chapter. Therefore, the forecasted demand function is a typical linear equation. Using the unstandardized coefficients for statistically significant variables, the forecasted demand function for air travel between Orlando and Los Angeles is:

$$D_{MCO\text{-}LAX} = 8.586(GDP) - 211.815(P) + 5{,}218.855(Q2) + 3{,}840.672(Q3) - 19{,}971.5(9/11)$$

Based on this equation, the demand for Orlando to Los Angeles air transportation can be estimated. Moreover, the forecast demand function also displays the impact that a change in one of the independent variables has on the demand. For example, a $1 increase/decrease in the average ticket price will cause demand to decrease/increase by over 200 seats. As we might expect, this kind of information is extremely useful to aviation managers of all types. The seasonality dummy variables also have a large impact on demand. For example, if the flight is in the second quarter, then the demand for the flight will increase by over 5,200 passengers as compared to the fourth quarter. Again, this information is critically important to successful fleet mix planning. The other variables in this equation can be analyzed in a similar fashion.

Therefore, and by way of summary, in the air transportation industry, demand forecasting is critical to strategic planning and the ultimate success of the airline. Regression analysis is a powerful tool that can be extremely useful in forecasting and other strategic decisions.

Although this chapter has merely provided an overview of various methods for forecasting, and a somewhat more detailed presentation of regression analysis, in-depth discussions of all the topics can be found in the References section at the end of the chapter.

DATA SOURCES

In order to perform successful forecasting in the aviation industry, various data are required. This section outlines some of the data sources commonly used in aviation applications. It also indicates whether the data is freely accessible or can only be obtained through subscription fees, and, where appropriate, the web addresses are provided.[13] The majority of data sources described are from the United States, but data sources for international aviation are also provided. The major data sources discussed are:

- US Department of Transportation (DOT) / Bureau of Transportation Statistics (BTS)
- Federal Aviation Administration (FAA)
- International Air Transport Association (IATA)
- International Civil Aviation Organization (ICAO)
- Official Airline Guide (OAG)

13 Website addresses were current as of 2007.

- Airports Council International (ACI)
- Air Transport Intelligence (ATI)
- Airline Monitor
- UK Civil Aviation Authority (CAA)
- Transport Canada (TC)
- Eurocontrol
- Aircraft Owners and Pilots Association (AOPA)
- Bureau of Economic Analysis (BEA)
- Bureau of Labor Statistics (BLS)
- Organization for Economic Cooperation and Development (OECD)

US Department of Transportation (DOT)/Bureau of Transportation Statistics (BTS)

One of the best sources for aviation-specific data for US aviation activity is the US Department of Transportation (DOT), through the Bureau of Transportation Statistics. There are multiple DOT databases that provide a wealth of information for the airline industry.

One database that is used throughout this book is Form 41 which provides a wealth of information about US airlines, ranging from general airline financial data, specific airline cost data, general traffic data, and airport activity statistics. All US-registered airlines are required to provide the data to the DOT, and they can be useful for evaluating airlines.

Another useful database is O & D, which stands for Origin and Destination. Using a 10 per cent sample of actual tickets, various statistics are provided for individual US domestic city-pairs. The O & D database shows on what airline the passengers traveled, the average ticket price, and a large amount of other data. As might be expected, the O&D data is very useful for demand estimation.

The T100 database is similar to the O & D, but covers international city-pairs. However, the data is presented in a slightly differently format and is not as extensive. These data enable demand estimation for international routes.

In addition to these three major databases, the DOT also provides other databases such as Schedules, Fleet, and Commuter.[14] While DOT statistics are technically public information and can be obtained free of charge, unless the user has advanced Excel and Access skills, the data is very difficult to access. Therefore, in order to use most of the DOT data, airline database packages such as Back Aviation are required. Unfortunately, these products require a paid subscription.

Federal Aviation Administration (FAA)

The Federal Aviation Administration is another good source for US data, particularly for information and data concerning aviation accidents and safety. The FAA also provides data about aviation forecasts and other issues such as terminal space usage, passenger

14 It should be noted that additional aviation data is provided through the Department of Transportation (http://www.dot.gov) and the Bureau of Transportation Statistics (http://www.bts.gov) websites.

facility charges, and airline service indexes. All these data can be obtained without charge through the Federal Aviation Administration's website, http://www.faa.gov.

International Air Transport Association (IATA)

The International Air Transport Association highlights issues and provides information concerning issues affecting airlines globally. The free economic analysis section provides information concerning the industry outlook, cost comparisons, traffic analysis, and fuel prices. In addition to the free data, IATA provides a wealth of additional subscription information that compares international carriers and provides airline rankings in terms of a variety of statistics. Through IATA's website, http://www.iata.gov, a wealth of information, (particularly concerning global aviation issues) can be collected.

International Civil Aviation Organization (ICAO)

The International Civil Aviation Organization, an arm of the United Nations, is the source for pertinent legal issues, particularly international air service agreements. However, probably the most valuable source from ICAO is ICAOdata—a subscription database that provides international data, including origin and destination passenger statistics, airline financial data, and airport activity statistics. ICAOdata is a useful back-up source for filling in any data not covered by DOT O&D, and T100 databases. Information concerning ICAO and ICAOdata can be obtained through the website, www.icao.int.

Official Airline Guide (OAG)

The Official Airline Guide is a compilation of over 1,000 airline schedules, creating the definitive source on airline schedules. Users can access date-specific schedule information through www.oag.com without charge. However, for airlines, a complete historical OAG database is more useful. Through this database, ASMs can be easily determined for a large number of city-pairs.

Airports Council International (ACI)

Airports Council International is a community of international airports that collectively lobbies on various issues concerning airports. Through ACI's website, www.airports.org, data and rankings can be obtained concerning the number of passengers handled by various airports, the cargo movements through the airports, and the number of international passengers, to name just a few. ACI helps collate information concerning airports worldwide.

Air Transport Intelligence (ATI)

Air Transport Intelligence is a database encompassing a wealth of information on the aviation industry. ATI provides a database of aviation-specific journal articles from such publications as *Airline Business* and *Flight International*, and these can be quite helpful in any qualitative analysis. ATI also provides searchable databases on information concerning airlines, airports, aircraft, suppliers, and schedules. While ATI does not provide quantitative data, it is a valuable resource when initially researching specific areas. ATI is only available to subscribers, and more information can be gathered at www.rati.com.

Airline Monitor

Another subscription database is Airline Monitor, which reviews trends in the airline and commercial jet aircraft industries. Airline Monitor provides a variety of reports, in a variety of formats, over issues such as block-hour operating costs, airline financial results, and commercial aircraft production. It also provides historical data, and this is especially helpful in constructing time-series data with numerous observations. More information concerning the products offered by Airline Monitor can be found at www.airlinemonitor.com.

UK Civil Aviation Authority (CAA)

The UK Civil Aviation Authority provides a function similar to the FAA, except in the United Kingdom. Using www.caa.co.uk, information covering the entire UK aviation industry can be obtained. Through the economic regulation and statistics portion of the CAA's website, a wealth of statistical data can also be accessed.

Transport Canada (TC)

Transport Canada is the governing body for all transportation related activities in Canada. Statistics, data, and regulations concerning the commercial aviation industry can all be obtained through Transport Canada and StatsCan. More information concerning Transport Canada can be found at www.tc.gc.ca.

Eurocontrol

Eurocontrol, standing for the European Organization for the Safety of Air Navigation, is the primary provider of air traffic control services throughout Europe. Although specific data can be difficult to obtain from Eurocontrol, its website, www.eurocontrol.int, does provide a variety of information concerning the aviation industry in Europe. More specifically, Eurocontrol can provide detailed information pertaining to airport traffic, delays, and capacity management initiatives.

Aircraft Owners and Pilots Association (AOPA)

The Aircraft Owners and Pilots Association is a membership community that promotes and advocates for the general aviation industry. Recently, AOPA has been involved in the fight over fuel surcharges and restrictions concerning the use of general aviation aircraft in congested airspace. The AOPA website, www.aopa.org, is split into two sections: public and members. While the general public can receive basic information from AOPA, members can obtain a more thorough investigation of issues facing the general aviation community. In addition, members receive information pertaining to weather and flight planning.

Bureau of Economic Analysis (BEA)

The US Bureau of Economic Analysis is an essential source when forecasting demand for air transportation services. The BEA provides detailed statistics of the state of not only the US economy, but also regional economies. Since GDP is a suitable proxy for consumer income, data from the BEA can help in any regression analysis. The BEA also provides additional macroeconomic indicators such as balance of payments, unemployment, and industry-specific economic accounts. Data can be freely obtained at www.bea.gov.

Bureau of Labor Statistics (BLS)

The US Department of Labor's Bureau of Labor Statistics is the definitive source concerning the labor force in the United States. The BLS provides data on such factors as unemployment, consumer price indices, wages, and labor demographics. The level of data can be quite detailed, with the various statistics broken down into industries and regions. For any analysis involving labor, www.bls.gov should be consulted.

Organization for Economic Cooperation and Development (OECD)

The Organization for Economic Cooperation and Development comprises 30 member countries which have active relationships with over 70 countries and multiple non-governmental organizations (NGOs). The OECD is primarily concerned with social and macroeconomic issues, and statistics are sorted into various industries, enabling comparisons between countries. Unfortunately, the OECD does not publish any reports concerning the aviation industry, so much of the useful data from the OECD will be general macroeconomic data, usually displayed on a monthly or quarterly basis at www.oecd.org.

SUMMARY

This chapter presented an elementary discussion of the methods and techniques that can be used for forecasting, including qualitative and quantitative methods. The qualitative methods covered were focus groups, market surveys, market experiments, barometric forecasting, historical analogy and the Delphi method. The quantitative methods contained a brief overview of descriptive statistics followed by an explanation of the moving average, weighted moving average, exponential smoothing, and, finally, trend analysis. All the techniques were illustrated with concrete numerical examples. The next part of the chapter contained a more in-depth discussion of regression analysis followed by more numerical examples. Finally, a number of good data sources for forecasting were listed.

REFERENCES

Allen, W., Doherty, N., Weigelt, K. and Mansfield, E. (2005). *Managerial Economics* (6th edn). New York: W.W. Norton & Company.

Anderson, D. Sweeney, D. and Williams, T. (2006). *Quantitative Methods for Business* (10th edn). Mason, OH: Thomson Higher Education

Anderson, E.W. and Mittal, V. (2000). Strengthening the Satisfaction–Profit Chain. *Journal of Service Research*, 3(2), pp. 107–20.

Barreto, H. and Howland, F. (2006). *Introductory Econometrics: Using Monte Carlo Simulation with Microsoft Excel*, Cambridge: Cambridge University Press.

Brandt, R. (1987). A Procedure for Identifying Value-enhancing Service Components using Customer Satisfaction Survey Data. in C. Surprenant (ed.), *Add Value to Your Service*, Chicago: American Marketing Association, pp. 61–65.

Doane, D. and Seward, L. (2007). *Applied Statistics in Business and Economics*. New York: McGraw Hill.

Durbin, J. and Watson, G. (1951). Testing for Serial Correlation in Least Squares Regression, II. *Biometrika*, 38, pp. 159–79.

Hanushek, E.A. and Jackson, J.E. (1977), *Statistical Methods for Social Scientists*. New York: Academic Press.

Hirschey, M. (2006). *Managerial Economics* (11th edn). Mason, OH: Thomson Higher Education.

Ovedovitz, C. (2001). *Business Statistics in Brief*. Cincinnati, OH: South Western College Publishing.

Spirtes, P., Glymour, C. and Scheines, R. (2001). *Causation, Prediction, and Search* (2nd edn). Boston, MA: MIT Press.

Truett, J. and Truett, B. (1992). *Managerial Economics* (4th edn). Cincinnati, OH: South-Western College Publishing.

11

Pricing Policy and Revenue Management

I believe that revenue management is the single most important technical development in transportation management since we entered the era of airline deregulation in 1979.

We estimate that revenue management has generated $1.4 billion in incremental revenue in the last three years by creating a pricing structure that responds to demand on a flight-by-flight basis.

Robert L Crandall, Chairman and CEO, AMR, 1992

This chapter will introduce the reader to the concepts of airline pricing policy and revenue management. Revenue management is essentially the combination of methods, analysis, and techniques that an airline applies to the types of service it offers in order to maximize the aircraft revenue. Airlines employ revenue management not only to sell as many high-priced seats as efficiently as possible, but to also keep airplanes full. A short section on airlines' past pricing practices is followed by a discussion of current pricing structure. Further, we show that segmenting the market or "price discrimination," based on the elasticity of demand for different types of passenger, can increase revenues. We then discuss strategies that airlines can use to segment their markets directly and indirectly, including such practices as advanced purchase restrictions and Saturday night stay requirements. The topics covered in the chapter are as follows:

- The importance of revenue management
- Pricing policy before and after deregulation
- Price discrimination
- Revenue management "fences, " including:
  - Advance purchase restrictions
  - Saturday night stay requirement
  - Frequent-flyer mileage
  - Refundability
  - Change Fees
  - Airline schedule
- Revenue management control types
- Spoilage and spillage

- Leg-based EMSR model
- Overbooking
- Other issues associated with revenue management.

THE IMPORTANCE OF REVENUE MANAGEMENT

The financial performance of airlines, like most other businesses, depends mainly on their sales strategy within a competitive industry. The knowledge of varying demand conditions, different classes of passengers, degrees of price sensitivity (elasticity of demand) among various groups of passengers, and the significance of the stochastic nature of demand by the traveling public (for example, the number of reservations and actual trips may differ) will influence the airlines' ultimate performance. Therefore, it is not surprising that airlines, recognizing all these dynamic factors, charge different fares in order to respond effectively to varying elasticities, different passengers' income, competitors' pricing policy, and market conditions. This practice is termed revenue (or yield) management.

Revenue management is a quantitative technique which allows an airline manager to handle the supply of aircraft seats and passenger demand to maximize revenues. The basic theory behind revenue management is that it may be beneficial not to sell something today at a low price if it can be sold tomorrow at a higher price, or allowing something to be sold today at a low price if it is otherwise likely to remain unsold. In essence, revenue management is a game of probabilities with the goal of extracting the maximum revenue that a passenger is willing to pay. An effective revenue management system requires:

- the establishment of a differential fare structure
- a system of constraints (or fences) on the use of lower-fare seats to limit their availability to passengers who might otherwise be willing to a pay a higher fare
- a system of seat allocation which maximizes expected revenue in the face of stochastic demand
- forecasts of demand, no-shows, cancellations, go-shows, overbooking, and inventory limit.

The importance of revenue management cannot be overstated. The example of People Express in the 1980s is probably one of the starkest examples of the importance of revenue management to the industry. Donald Burr, former CEO of People Express explains:

> We were a vibrant, profitable company from 1981 to 1985, and then we tipped right over into losing $50 million a month. We were still the same company. What changed was American's ability to do widespread Revenue Management in every one of our markets. We had been profitable from the day we started until American came at us with Ultimate Super Savers. That was the end of our run because they were able to under price us at will and surreptitiously. There was nothing left to defend us. What you don't know about revenue management could kill you!
>
> (Cited in Loveman and Beer, 1991)

People Express was a fledgling discount airline that was born out of deregulation in 1978. Initially, the discount airline flew niche markets that competed mostly with buses and cars—markets that the major carriers were happy to leave it to. However, benefiting from a cost structure that was $1 billion below other airlines such as American Airlines, People Express began to undercut the fares of the major US domestic carriers. Operating at a 75 per cent load factor and a 72 per cent break-even load factor on some of America's busiest routes in 1983, the discount airline, with its simple pricing structure and extensive cost advantage, seemed unstoppable, yet it had one weakness in its corporate structure. In order to save money during start-up, People Express had installed a simple reservations system that was unable to practice revenue management. Its information technology system could offer peak and off-peak fares, but each flight had to be either one or the other (peak or off-peak); multiple fares were not possible. This meant that, on each flight, People Express was only able to offer only one fare (Cross, 1995).

On 17 January 1985 American Airlines became the first airline to expose People Express's weakness when it launched "Ultimate Super Saver" fares that were priced at People Express's lowest prices (Cross, 1995). American placed 21-day advance purchase restrictions on the "Ultimate Super Saver" fares in order to allow only the most price-sensitive travelers to be eligible for the discounted fares (Cross, 1995). In addition, American controlled the number of "Ultimate Super Saver" fares available on each flight in order to save space for high-revenue passengers. In essence, American Airlines was able to generate revenue from both low-revenue and high-revenue passengers, while People Express could only accommodate low-revenue passengers with its single-fare class reservations system. As a result of American Airlines revenue management practices, People Express's load factor dropped from 70 per cent in 1984 to 57 per cent in 1986. Furthermore, as Table 11.1 shows, People Express's break-even load factor jumped 10 per cent from 1985 to 1986. Ultimately, with a break-even load factor significantly above its actual load factor, the discount airline hemorrhaged and lost substantial sums of money.

The example of People Express not only displays the importance of revenue management to an airline, but also introduces the concept of revenue management to other businesses. While modern revenue management has its roots in the airline industry, it is also widely used in the car rental, hotel, and cruise ship industries, to name a few. The same theory and practice of revenue management applies to all these industries, and it is therefore a very powerful tool.

PRICING POLICY BEFORE AND AFTER DEREGULATION

The Civil Aeronautical Act of 1938 established a policy of economic regulation of the domestic US airline industry. This Act created the Civil Aeronautical Board (CAB), which had authority over the level and structure of airfares within the United States. Prices

Table 11.1 Load factor for People Express

People Express						
	1981	1982	1983	1984	1985	1986
Load Factor	58%	61%	75%	70%	61%	57%
Break-even Load Factor	71%	60%	72%	70%	62%	72%

Source: Compiled by the authors using Form41 data.

were set by the CAB according to industry average costs, which disbarred lower-cost airlines from offering lower prices, since it was deemed unhealthy for the industry. The only exceptions to this rule were in the states of California and Texas where airlines were able to set their own prices on intrastate routes where the CAB did not have authority. This led to the rise of low-cost airlines AirCal in California and Southwest Airlines in Texas.[1] Within the regulated environment, airlines were provided with a protected route structure and guaranteed revenues that exceeded costs. This meant that airlines could rarely fail in the domestic marketplace, and in the event that an airline incurred losses, federal subsidies were available to bail them out (Spiller, 1981).

The Airline Deregulation Act of 1978 abolished the CAB's authority over airlines, and the market was permitted to decide airline fares and routes. Throughout the deregulation process airlines were eventually allowed to elect to fly any domestic route, without legal restrictions, and could offer any fare on any flight. Deregulation also allowed low-cost carriers, like Southwest, to expand beyond its Texas market and new discount carriers, like People Express, to start up. At the beginning of deregulation, about 50 per cent of total traffic traveled on a discount fare; by 1990 the figure was nearly 90 per cent. Increased competition and the liberalized pricing structure have led some industry analysts to claim that today's airfares are 20–30 per cent below what they would have been had regulation remained in place.

As the People Express example highlighted, revenue management in the airline industry was initiated in 1985 by American Airlines with its "Ultimate Super Saver" fares. Since then, almost every airline in the world has adopted a revenue management scheme to some degree. The benefits of doing so are immense: Delta Air Lines attributed $300 million in total bottom line to revenue management when it first started implementing revenue management. Another way of highlighting the benefits of revenue management is by presenting a simple, fictitious example.

Assume that DirectJet Airlines operates a short-haul route where the maximum daily demand for the flight is 100 passengers and the maximum any passenger is willing to pay for the flight is $250. This information helps us construct a demand curve for the flight, which, in this case, is assumed to be linear. A depiction of the demand curve is shown in Figure 11.1.

Under a uniform pricing strategy, DirectJet Airlines sets one single price for all passengers on a single flight. Recall that People Express was an airline that operated with a uniform pricing strategy. In our example, let's assume that DirectJet Airlines charges a uniform price of $100 for this particular short-haul flight. Based on the estimated demand function and at $100 airfare, 60 passengers are willing to buy tickets from the airline. This would generate $6,000 in total daily revenue for DirectJet Airlines. This is graphically represented in Figure 11.2, with the shaded area under the curve representing the total daily revenue for the flight.

The second pricing scenario available to DirectJet Airlines is a multiple-pricing strategy in which the airline uses segmental (differential) pricing to maximize revenue. In our example, DirectJet Airlines decided to adopt a new four-tier pricing structure whereby it offered fares ranging from $200 to $50. Based on the estimated demand function for this particular flight, 20 passengers are willing to pay the $200 fare, 40 passengers are willing to pay for the $150 fare, 60 passengers for the $100 fare, and 80 passengers for the $50 fare. Figure 11.3 displays graphically the revenue potential for the DirectJet Airlines flight

1 AirCal was a California-based airline that was eventually bought out by American Airlines.

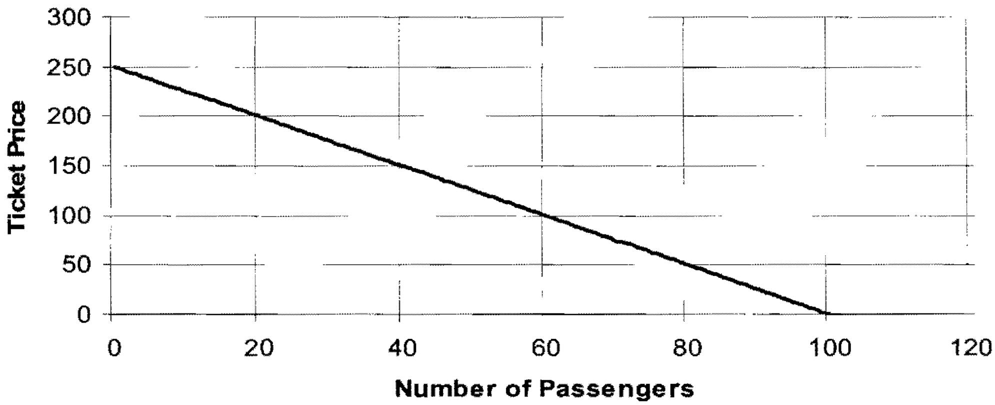

Figure 11.1 Demand curve for a DirectJet Airlines flight

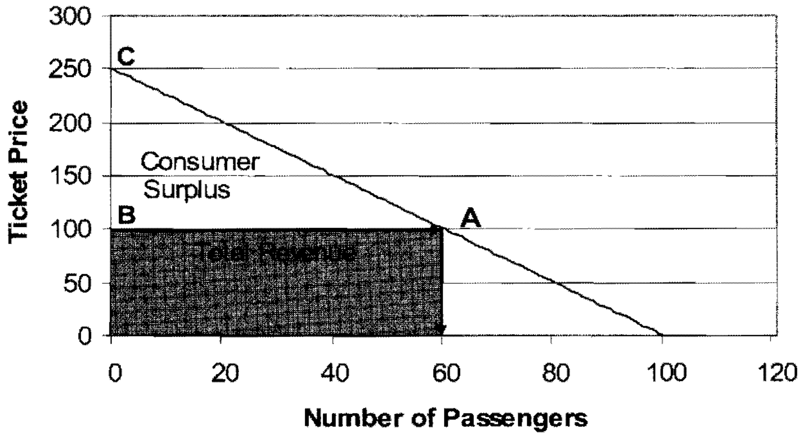

Figure 11.2 Uniform pricing for a DirectJet Airlines flight

using a multiple pricing strategy. It is immediately apparent that the shaded area under the multiple-pricing policy is greater than the shaded area in Figure 11.2 with the uniform pricing policy. This is confirmed numerically in that the four-tier pricing structure generates $10,000 in total daily revenue, which is greater than the uniform pricing policy:

New aircraft revenue = 20 seats($200) + 20 seats($150) + 20 seats($100) + 20 seats($50) = $10,000

This shows that a multiple-pricing policy brings a major benefit of increasing total flight revenues, but also that it enables DirectJet Airlines to offer cheap, discounted airfares that could undercut the competition. If we assume that DirectJet Airlines is operating a 100-seat aircraft, then price-sensitive passengers would simply be occupying an otherwise empty seat. This is exactly what American Airlines was able to achieve with its "Ultimate Super Saver" fares. The end result of a multiple-pricing strategy is that total revenue and, possibly, total passengers would increase compared to what would be achieved by a uniform pricing strategy. By using a multiple-pricing strategy with four fare groups in our fictitious example, revenues increased from $6,000 to $10,000.

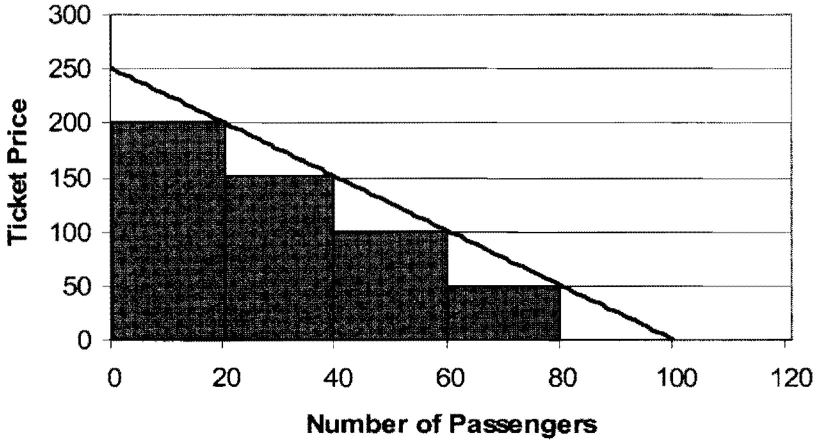

Figure 11.3 Multiple pricing for a DirectJet Airlines flight

Another way of looking at uniform versus multiple pricing is through the eyes of the passengers. In our example, under uniform pricing, many passengers who purchased the $100 airfare were willing to pay more than that. There were many passengers who were willing to pay over $200, but ended up only having to pay $100. This difference between what a passenger was willing to pay and what the passenger actually paid is called consumer surplus. Consumer surplus can also be easily calculated by finding the area of the unshaded triangle region that lies beneath the demand curve (the triangle A,B,C in Figure 11.2). Under uniform pricing, consumer surplus amounted to:

$$\$4{,}500 \text{ or } [\frac{60\,seats \times \$150}{2}]$$

Conversely, under a multiple-pricing strategy, the amount of consumer surplus is the area of the multiple unshaded triangles that lie beneath the demand curve. From Figure 11.3 we are able to determine that there exists only $2,000 in consumer surplus for DirectJet's four-tier pricing structure.[2] Since the goal of revenue management is to extract the largest amount of revenue from every passenger, it can also be said that revenue management's objective is to also minimize consumer surplus. Of course, if DirectJet could sell each customer a ticket at the maximum fare that the customer would pay, then DirectJet would get the maximum possible revenue, and there would be no consumer surplus. However, in the real world this is never possible since the information requirements are too large, and there is too much uncertainty in consumer behavior itself. Nevertheless, it is possible to present a menu of prices to consumers based on estimates of their likely price sensitivity and that leads us directly to our next topic—price discrimination.

2 Consumer surplus = .5(20*50) + .5(20*50) + .5(20*50) + .5(20*50) = $2,000.

PRICE DISCRIMINATION

Price discrimination is the practice of charging different prices to different customers for the same product. The different prices charged are based on the consumers' various price elasticities of demand. Although the practice may seem unfair, price discrimination is legal and common in modern business. For example, grocery stores practice price discrimination by offering coupon discounts to consumers who are not time-sensitive but price-sensitive, and who are willing to search out and bring the coupons to the store. US universities, especially state universities, practice price discrimination by offering different tuition levels for international, out-of-state, and interstate students. The telephone companies' practice of offering discounted calling rates during evenings and weekends is price discrimination. Typically, in many flea markets there are really no set prices for the goods offered, but customers bargain with the merchants. In fact, bargaining is the oldest form of price discrimination and has existed since commerce began.

As we saw above, price discrimination is essentially the flip-side of yield management and is a requirement for the practice of yield management in the airline industry. And, as is now commonly appreciated, every flight has numerous fare classes for essentially the same seats and service. Today, the airline industry is one of the industries that spend the most effort on practicing price discrimination.

As mentioned earlier, consumer surplus is the difference between the amount the passenger is willing to pay and the amount he or she actually pays. In essence, it is the perceived "deal" that consumers receive when they purchase a good or service. The goal of price discrimination, and therefore yield management, is a reduction in consumer surplus.

Consider a flight with six passengers who are all willing to pay various prices for the same flight. Their maximum willingness to pay is contained in Table 11.2. Based on this data, a demand curve can be created for this flight, which is depicted in Figure 11.4. If an airline charges a single fare of $250, consumer surplus—the difference between the maximum willingness to pay and the actual ticket price—would exist for five of the six passengers. Only passenger F (or 6) would receive no consumer surplus. The ultimate goal of both price discrimination and revenue management is to minimize consumer surplus; therefore, six individual fare categories would have to be created that maximize airline revenue and minimize consumer surplus. The catch for the airlines is that ascertaining every passenger's willingness to pay and the flight's demand curve can be difficult, if not impossible.

Table 11.2 Consumer surplus

Passenger	Demand	Ticket Price	Ticket Price	CS*
A	1	250	500	250
B	2	250	450	200
C	3	250	400	150
D	4	250	350	100
E	5	250	300	50
F	6	250	250	0

* *CS = Consumer surplus.*

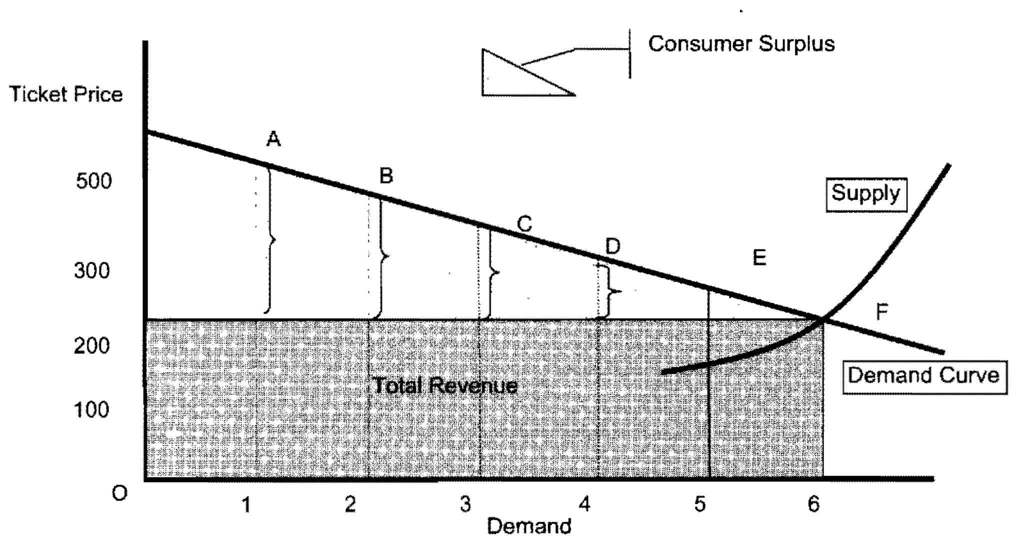

Figure 11.4 Demand curve

In practice there are a number of ways in which a business can institute at least some form of price discrimination. However, in order for it to be successful, three necessary conditions need to apply to the market:

- market segmentation
- different elasticities in different submarkets
- market separation.

The first requirement for price discrimination is that the markets must be segmented. By this we mean that there should exist different groups of consumers who do not have the same interests. In the aviation industry a common method of market segmentation is leisure and business travelers. Since there maybe extensive overlap in these categories another more accurate segmentation would be time-sensitive or price-sensitive travelers. Time-sensitive travelers are typically business travelers who demand to travel on certain days and at certain times. These passengers will typically ignore the ticket price in order to satisfy their demand for traveling at a certain time and date. Certain types of leisure travelers may also be contained in this category, especially vacationers who may be leaving and returning on a set schedule. Price-sensitive travelers are the opposite in that their selection of flights is based on the ticket price. These are travelers who are willing to travel at inconvenient times and by longer routings if this results in lower fares.

The second requirement for price discrimination is that different elasticities must exist for different submarkets. This requirement is closely related to the first one in that the market can be segmented by price elasticity, but the first requirement deals with how the passengers can be grouped, while this requirement deals with the passengers' willingness to pay. If all passengers had the same price elasticity, then the airline would be unable to charge different prices. In the air travel industry both these requirements are easily met, since every market contains a variety of different people who are willing to fly at different times and at different prices.

The third requirement of price discrimination, market separation, is that the airline must be able to effectively isolate the market and be successful at charging different prices to different passengers. Airlines achieve market separation through pricing "fences." This practice is covered in the next section, but, briefly, it includes non-transferability of tickets and Saturday night stay requirements; these allow the airline to prevent the customer from reselling or using the ticket in some other way than for the flight—hence the term "fences".

There are three degrees of price discrimination. First-degree price discrimination, also called perfect price discrimination, involves charging different prices for every unit up to the point where consumer surplus does not exist. Bartering is the classic case of perfect price discrimination, as are car dealerships to a certain extent. In both these cases, every consumer pays a different and unique price for the same product. Another example is an auction where consumers will keep bidding up until they reach their maximum willingness to pay. With regard to the airline industry, perfect price discrimination is practically non-existent for reasons mentioned earlier.

Second-degree price discrimination is simply quantity discounts. Wal-Mart is successful in obtaining low prices from its suppliers through massive quantity discounts. The airline industry has limited experience with second-degree price discrimination although charter fights, cruise ship companies, consolidators, and corporate travel deals are examples of second-degree price discrimination. In all of these situations, the companies can receive discounted prices by agreeing to buy a large proportion of seats on a flight. Airlines typically favor this practice, as it provides them with a certain amount of guaranteed revenue for a flight, albeit at a reduced rate.

Finally, there is third-degree price discrimination, and this is the type that is typically practiced by the airline industry. It involves dividing consumers into different groups, based on a set of certain characteristics, and estimating their respective demand curves. At this point, each group is charged a different price. The group with the most inelastic demand (typically the most time-sensitive group) is charged the highest price. The different fare classes that can be observed in the market correspond to the groups. With third-degree price discrimination, a certain amount of consumer surplus will exist, since the prices are not set for every individual but for the group as a whole. However, the aim of creating additional fare classes is to reduce the amount of consumer surplus and ultimately increase revenues for the airline.

Price discrimination, despite the negative connotation, is a common and generally efficient procedure. To see this, consider the impact of a uniform pricing policy as shown in Figure 11.5. With the typical cost structure of imperfect competition, we see long-run equilibrium where P* = ATC at output Q*.[3] Note that this leaves a huge segment of demand unsatisfied and results in a great degree of wasted capacity (that is, an awful lot of empty airline seats). Many consumers would be willing to pay a price higher than the marginal cost of serving them, though lower than P*. However, with uniform pricing, the price can not be set below ATC in the long run because the firm is barely breaking even at a price equal to ATC. The only option available to airlines to increase their revenues is selective price cuts. If it were possible to read consumers' minds, we could reduce price just enough below P* to induce them to buy a ticket. In this case, each consumer pays exactly the highest price they are willing to pay; MR is now equal to this personalized price so that every customer willing to pay a price greater than marginal cost can be served. Under this "perfect price discrimination" regime, output can be increased to Q'.[4]

3 ATC = average total cost.

4 Beyond Q', marginal cost is greater than the price that could be charged, so that segment of demand will not be serviced.

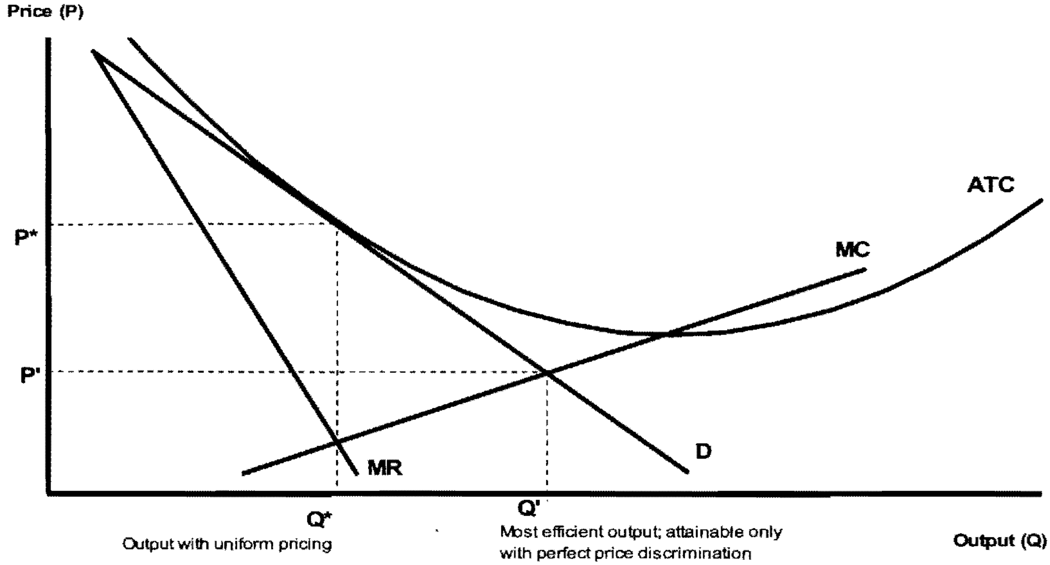

Figure 11.5 Pricing policy and price discrimination

For airlines, perfect price discrimination is not possible, but sophisticated third degree price discrimination can allow airlines to move closer to the efficient output level of Q′ as they fill otherwise empty seats with selective price cuts. For instance, if most students are on the demand segment below P*, then a discount for students will bring in new revenues. Similarly, demographically-based price cuts for senior citizens or families can also achieve the desired effect. The airlines can also keep the price cuts limited through the use of revenue management "fences", which will be discussed in the following section.

Of course, in moving from uniform pricing to the more tailored approach of price discrimination, it may be that some segments of demand will face a price higher than P*. An interesting question is whether price discrimination leads to an average fare lower than P*, the theoretical uniform price. Most economists agree that price discrimination typically does lead to a lower average price, producing fewer empty airline seats and greater economic efficiency. The key reason for this result is competition. Because of the threat of being undercut by a competitor, it is always easier to cut prices than it is to raise them.

Even if price discrimination reduces average price, it may still raise price for particular consumers in the upper portion of the demand curve. A commonly asserted complaint is that certain business travelers end up paying high fares to "subsidize" consumers receiving discounts. However, there are problems with this theory. As already explained, a uniform price would drive many discount customers completely out of the scheduled airline market. This would reduce airline revenues, and ultimately lead to both price increases and reductions in available seat miles, both of which would be unpleasant to business travelers. Thus, in a sense, one could just as easily argue that discount flyers "subsidize" business travelers. In order to continue the present service standards, airlines need every penny of revenue they can get, just as a restaurant may be financially viable only with revenues from regular, full-paying customers combined with revenues from patrons who dine there only occasionally when they have a discount coupon. Each set of consumers benefits from the other, since only their combined revenues are enough to sustain the product they both enjoy.

Price discrimination based on characteristics such as gender, either when firms offer policies at a fixed price or when they charge according to some consumption variable that is correlated to costs has been studied by Buzzacchi and Valletti (2005). For instance, consider an airline serving only leisure vacationers (therefore, no business or time-sensitive travelers). This airline would have a very different product design from other carriers. Their consumers care little about the exact time of departure, are willing to commit to a schedule way in advance, and are very price-sensitive. This type of airline would look pretty much like today's charter airlines: they would operate a very infrequent service, would employ large aircraft in a high-density seating configuration with very high load factors (probably over 90 per cent), and would routinely cancel any flight well in advance if it was substantially undersubscribed. Under such conditions costs, and therefore average prices, could be kept very low. However, if the airline wanted to accommodate business travelers, it would have to offer multiple flights with varying departure times and mostly stick to a schedule published a few months in advance. The airline would also keep some seats open for late, even last-minute, travelers. Since business travelers require the design of a more expensive product, it seems reasonable to argue, as Frank (1983) does, that it is philosophically appropriate to charge them more. In essence, the appearance that business and leisure travelers are sometimes paying very different prices for the same service is an illusion. In reality, the typical time-sensitive traveler demands a very different, and much more expensive, sort of service than price-sensitive travelers.

REVENUE MANAGEMENT "FENCES"

One of the most important factors for the implementation of a revenue management system is the effective use of "fences," or barriers that limit the use of discounted seats to passengers who might otherwise be willing to pay a much higher fare. The airlines do not want a business traveler who is willing to pay full fare actually obtaining a deeply discounted fare. The way this is accomplished is through the use of "fences". In practical airline pricing policy there are six principal "fences," each of which will be discussed in greater detail:

- advance purchase requirements (restrictions)
- Saturday night stay
- frequent-flyer mileage
- ticket refundability
- change fees
- schedule-driven "fences."

Advance Purchase Restrictions

Advance purchase restrictions—one of the oldest "fences" implemented in the airline industry—simply limit the amount of time before the day of departure that a ticket can be purchased. American Airlines' "Ultimate Super Saver" fares, for example, had a 21-day advance purchase restriction on them. Advance purchase restrictions were implemented in the belief that passengers who were more price-sensitive (and less time-sensitive) would book further in advance. Conversely, passengers who show up at the airport wishing to travel on the next flight are clearly extremely time-sensitive and price-insensitive, and should therefore, be willing to pay quite a high fare. A typical fare class structure relating to advance purchase restrictions is illustrated in Table 11.3.

Saturday Night Stay Requirement

One of the most infamous revenue management "fences" is the Saturday night stay requirement which was implemented to try to keep business travelers from obtaining cheaper airfares. Since most business travelers are time-sensitive and want to depart Monday morning and return Friday evening, the Saturday night stay requirement was used to help segment the business travelers from the leisure travelers. Many low-cost carriers have eliminated the Saturday night stay requirement from their pricing policy because business traveler trends have changed slightly, and the rule seems archaic. Low-cost carriers have used the abolition of this "fence" in many marketing campaigns as well. Table 11.4 expands our sample fare structure to include Saturday night stay requirements.

Frequent-Flyer Mileage

Although frequent-flyer programs have been around for quite some time, only recently have they begun to be used as a revenue management "fence." Because frequent-flyer programs have been successful at attracting and retaining loyal customers, the number of miles offered for a fare class can be an important factor for passengers. For instance, a passenger may be willing to purchase-up a fare class if the next fare class offers more frequent-flyer miles. However, in order for this fence to be effective, full transparency of

Table 11.3 Fare class advance purchase restrictions

Fare Class	Advance Purchase Restrictions
Y	No
M	7-day
Q	14-day
T	21-day

Table 11.4 Fare class Saturday Night stay restrictions

Fare Class	Advance Purchase Restrictions	Saturday Night Stay?
Y	No	No
M	7-day	No
Q	14-day	Yes
T	21-day	Yes

the fare classes/options is required. In other words, if only one fare option appears when a passenger wishes to purchase, this "fence" will not be effective since the passenger will not know about other options. The most effective marketing use of frequent-flyer mileage as a yield management "fence" is a matrix approach, with various fare types available to the passenger. Alaska Airlines' and Air Canada's websites are good examples of airlines offering matrices of fare types to their customers. Another related frequent-flyer benefit is complimentary first-class upgrades. Depending on the fare class booked, a passenger may be entitled to a complimentary first-class upgrade or an upgrade with additional miles. For most airlines the passenger must be booked above a certain fare class level to be eligible for these perks. Table 11.5 updates the fare structure to include a percentage of actual miles flown that a passenger would receive as frequent-flyer mileage.

Refundability

Ticket refundability is another important "fence" implemented by airlines worldwide to help segment the market. Usually higher fare classes will have full ticket refundability; enabling a passenger to cancel a reservation and receive a full refund. Thus, the refundable ticket provides the passenger with greater flexibility—something that, of course, is usually more desired by time-sensitive travelers. Lower fare classes usually do not provide a refund unless there are extenuating circumstances. Table 11.6 updates the fare structure to include a refund option.

Table 11.5 Fare class frequent-flyer mileage

Fare Class	Advance Purchase Restrictions	Saturday Night Stay?	Frequent-Flyer Mileage
Y	No	No	150%
M	7-day	No	100%
Q	14-day	Yes	100%
T	21-day	Yes	50%

Table 11.6 Fare class refundable restrictions

Fare Class	Advance Purchase Restrictions	Saturday Night Stay?	Frequent-Flyer Mileage	Refundable?
Y	No	No	150%	Yes
M	7-day	No	100%	Yes
Q	14-day	Yes	100%	No
T	21-day	Yes	50%	No

Change Fees

Similar to ticket refundability, change fees are another important revenue management "fence" used to differentiate travelers based on their time sensitivity. Travelers who require great time flexibility like to be able to change flights at will, or for a nominal fee. Usually the highest fare class allows full flexibility and is desired by business passengers whose schedule can change at short notice or who simply do not want to wait around at the airport. At the other end of the spectrum, some of the lowest fare classes may not even permit schedule changes. However, most fare classes require a change fee to be paid, in addition to the difference in fare. While the change fee may be minimal, the difference in fare could be extensive, especially from lower fare classes. In essence, the difference in fare charge is the difference between the fare class paid by the passenger and the lowest available fare class on the flight the passenger wants to change to. Since higher-fare passengers have fewer fare classes above them, the difference in fare charge is usually not as large. Table 11.7 updates the fare structure to include various change fees.

Airline Schedule

The final "fence" to discuss relates to the timing of an airline's schedule. Since different types of passenger have different traveling patterns, airlines can more profitably allocate high- and low-fare seating if they are aware of the likely composition of the passengers for a flight. For instance, time-sensitive travelers might want an early morning departure and an evening return so that they can conduct a full day's business. In this case, the airline will choose to limit the number of low-fare flights for a same-day round trip. Leisure passengers, on the other hand, exhibit different travel patterns, and these might include the ability to be flexible with regard to departure and return dates. Hence we observe mid-week sale specials and last-minute discounts to various locations.

REVENUE MANAGEMENT CONTROL TYPES

Before the various fare class allocation methods (or control types) are presented, two key terms need to be explained. Booking limit is the maximum number of seats that can be purchased for each fare class, and protection level is the number of seats that are left unsold so that they may be purchased by a higher fare class. Depending on the control type implemented, there may only be a few seats distinctively retained, or protected, for

Table 11.7 Fare class change fee restrictions

Fare Class	Advance Purchase Restrictions	Saturday Night Stay?	Frequent-Flyer Mileage	Refundable?	Change Fee
Y	No	No	150%	Yes	Free
M	7-day	No	100%	Yes	$25
Q	14-day	Yes	100%	No	$150
T	21-day	Yes	50%	No	No Changes

higher fare classes or there may be a substantial number. For a yield management analyst, booking limits and protection levels are extremely important concepts as they attempt to maximize revenue for every flight.

There are two main types of control limit used in yield management: distinct and nested. In distinct control, a fixed number of seats are allocated to each fare class, and the fare can only be purchased if there remains inventory in the fare bucket (see Figure 11.6). Under distinct control, protection level and booking level are equal since there is no provision for shifting fare classes. From the airline's point of view this is obviously an inefficient scheme since it amounts to a rather inflexible form of price discrimination. For example, there may be many passengers who are price-sensitive and would purchase a lower-priced fare if it were available. If the airline has guessed wrong on the number of seats allocated at the lower fare, and if the passengers are unwilling to pay the next higher fare, then there are likely to be unsold seats on the flight. This, of course, results in lower overall lower revenue. Because of these inefficiencies, distinct control is very rarely implemented in airline yield management.

The predominant scheme utilized in yield management is some derivation of nested control. Nested control schemes can be customized to suit the individual characteristics of the flight, but the basic principle is that lower fare classes are embedded in higher fare classes' booking limit. Therefore, under a pure (or serial) nested control scheme, a higher fare bucket will never be closed out prior to a lower fare bucket. Figure 11.7 highlights a serial nesting scheme for the same 300-seat aircraft, with the number representing the booking limit for each class. Under this scenario, the total aircraft capacity could be booked in Y-class, but only 100 seats are protected for Y-class. In this case, protection level is calculated by simply finding the difference between each fare class. Another example of a nested control structure is a parallel nesting scheme as presented in Figure 11.8. While similar to serial nesting, a parallel structure allows for the M-class fare to be closed prior to the T-class, yet still allows the entire aircraft to be booked in full Y-class. Such a structure, or derivation thereof, may be used to provide a set inventory reserved for frequent-flyer mileage redemption or corporate travel arrangements (Vinod, 1995).

A major type of nested control is virtual nesting. Virtual nesting deals from a total revenue and total network perspective as it helps determine whether selling a seat in a high fare class on a single sector might be sub-optimal relative to selling that same seat to a connecting passenger in a lower fare class. In essence, with most airline itineraries seemingly involving a change of planes through a hub, virtual nesting looks at the total revenue the booking would generate. For instance, if a full unrestricted economy fare on a short-haul sector generates less revenue than a discounted fare on a long-haul international

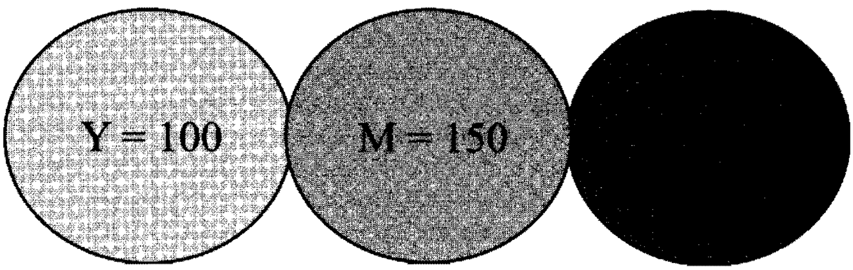

Figure 11.6 Distinct control

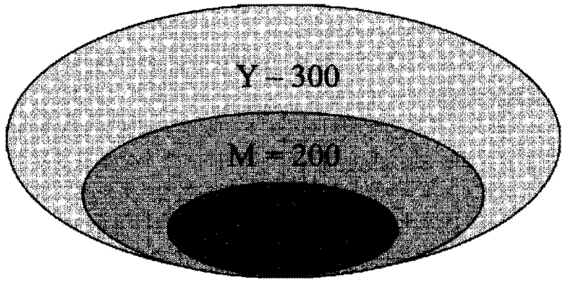

Figure 11.7 Serial nesting

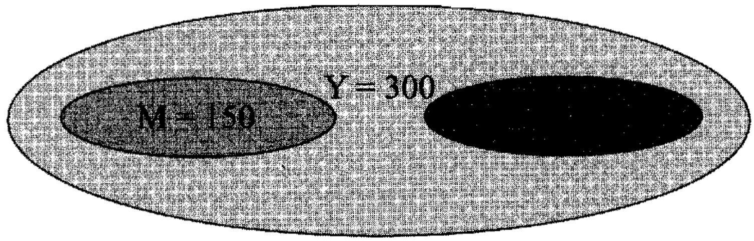

Figure 11.8 Parallel nesting

flight, the longer itinerary would have priority. This is accomplished by "clustering the various itinerary fare classes that flow over a flight leg into a manageable number of buckets, based on customer value" (Vinod, 1995, p. 462). Thousands of potential itineraries can be grouped into a few virtual buckets, but the variance in each of these buckets can be considerable (Vinod, 1995).

SPOILAGE AND SPILLAGE

Using the various control types, revenue management analysts are able to open and close fare buckets to adjust to the demand for the flight. Prices for flights are adjusted with respect to the normal booking curve for the flight, which is based on historical demand. A normal booking curve is when the last seat of the aircraft is purchased just before the time of departure.[5] If such a situation were to occur, this would represent complete revenue management effectiveness and maximize the airlines' revenues (Littlewood, 2005). Figures 11.9, 11.10, and 11.11 each display different situations that may occur with respect to the normal booking curve. In Figure 11.9 the actual booking curve results in the number of bookings at the date of departure being less than the capacity of the flight. This difference between capacity and actual bookings is called "spoilage," and is visually represented as empty seats on an aircraft. Airlines want to reduce spoilage since an empty seat provides no additional revenue for the airline, and, in all likelihood, that seat could have been sold if the price was right. Spoilage is a result of prices being too high for the market, and if the booking rate is less than the normal booking curve for the flight, the revenue analyst can lower the average ticket price, or open higher fare classes for sale, to reduce potential spoilage.

5 If there is no overbooking.

Figure 11.10 depicts the reverse situation of spoilage, which occurs when all the seats are purchased prior to the departure of the flight. This situation, called "spillage," is also a problem because the airline generally wants to hold a few seats available for last-minute travelers who are willing to pay full fare for the flight. By already having the flight fully booked, the airline is incurring a potential loss of revenue for the flight. Spillage is the result of having too low average fares for the flight, which leads in a booking curve that lies above the normal booking curve. Both spillage and spoilage are of concern for revenue analysts, and they must balance the fine line between both problems to reach an ideal normal booking curve.

The third figure of the series, Figure 11.11, depicts a situation similar to spillage when the airline books more passengers for the flight than capacity. This situation, called "overbooking," is a normal occurrence, since a probabilistic percentage of passengers do not show up for their flight. Airlines routinely set booking limits that exceed the capacity of the aircraft in order to maximize revenue. Further analysis of the overbooking issue is discussed later on in this chapter.

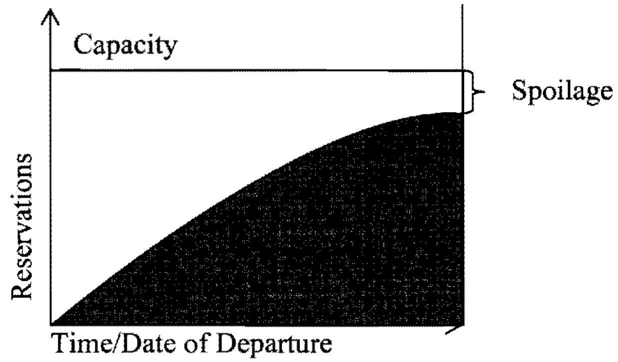

Figure 11.9 Spoilage

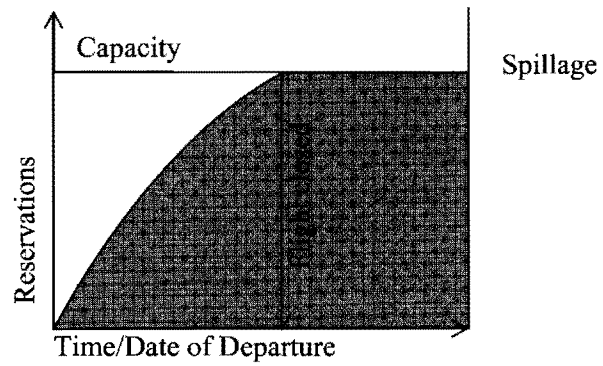

Figure 11.10 Spillage

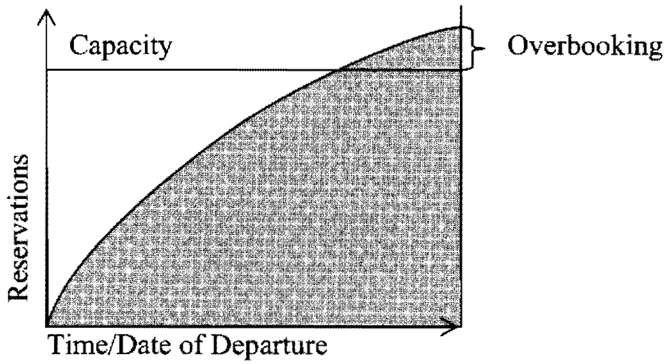

Figure 11.11 Overbooking

LEG-BASED EMSR MODEL

One of the main methods used to determine the desired booking limits for a flight is the Expected Marginal Seat Revenue (EMSR) model. It was developed by Littlewood (1972); later, Belobaba (1987) used this concept to address a single-leg flight with multiple fare classes. Simply put, the expected marginal seat revenue means the expected revenue contribution of one additional seat. In an EMSR model the number of seats allocated to each fare class is determined by using historical information about fares and current and past booking figures. The expected marginal seat revenue of the *i*th seat sold is:

$$EMSR_i = f_i \times P(S_i)$$

where: EMSR is the product of the fare level, f_i, and the probability that there will be at least *n* passengers willing to buy *i* class tickets for the flight under consideration.[6]

Figures 11.12 and 11.13 provide the cumulative probability distribution for two unique fare classes. Since the underlying assumption is that the probability of booking is based on a normal distribution, the average probability of demand for each fare class is 50 per cent. In both fare classes there is close to a 100 per cent probability that at least one seat can be sold at the given fare classes, while it is unlikely that more than 25 and 50 seats will be demanded for class 1 and 2, respectively.

Using the formula, the expected marginal seat revenue for every seat can be calculated by simply multiplying the ticket fare by the cumulative probability of demand for that seat. For instance, assuming that a ticket costs $500 and the cumulative probability of demand for that seat is 50 per cent (or 0.5), the EMSR for that seat is $250. This formula is applied to every set of seats for every fare class, creating the ability to graph the EMSR curve. This has been done for the two cumulative probability distributions that were

6 The probability for a passenger's willingness to buy is assumed to be a normal distribution. The cumulative normal distribution is essentially the one-sided probability of a passenger's willingness to purchase a seat. Values for a cumulative normal distribution can be calculated by using the normdist function in Microsoft Excel, or similar functions in other spreadsheet packages.

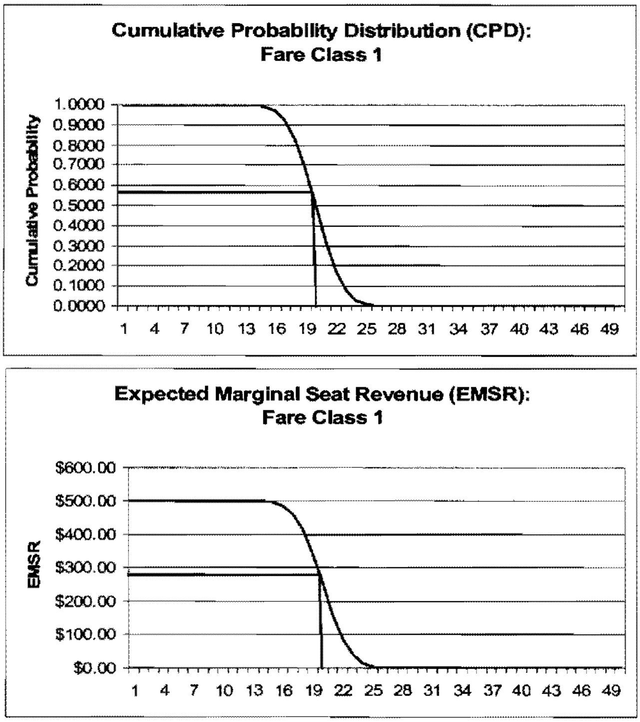

Figure 11.12 CPD and EMSR: fare class 1

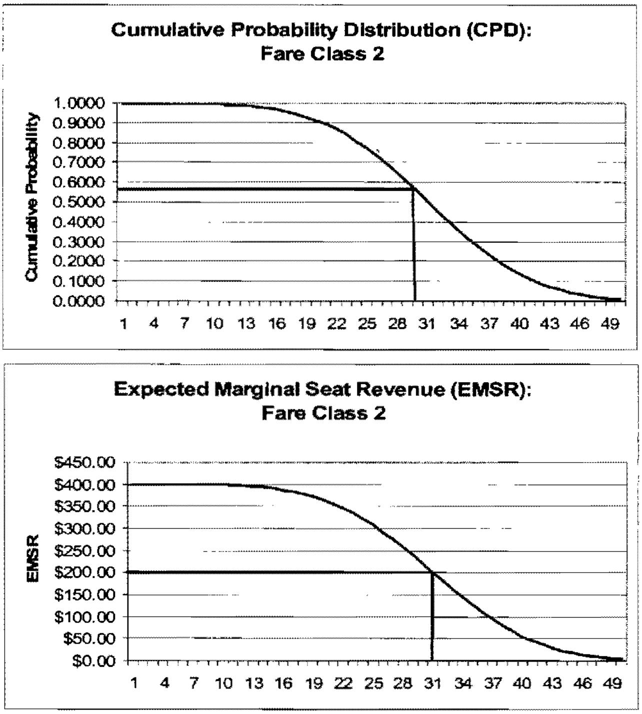

Figure 11.13 CPD and EMSR: fare class 2

presented above, assuming the fare in the first fare class is $500 and the fare in the second fare class is $400.

The EMSR curves for both fare classes appear similar to the cumulative probability distributions, except that the vertical axis is no longer probabilities, but actual dollars. The average expected marginal seat revenue for fare class 1 occurs at $250 and roughly 20 seats, while in fare class 2 the average expected marginal seat revenue is $200 at 31 seats.

The final step in the analysis is to combine the two EMSR curves together in order to determine the protection level and booking limits for the fare classes. In Figure 11.14 the two EMSR curves intersect at a point close to 19 seats. This point represents the protection level for fare class 1 over fare class 2. If we assume that these are the only two fare classes for the flight, the booking limit for the higher fare class would be the capacity of the aircraft, since if everybody wants to purchase the highest ticket, the airline would be more than glad to accept. The protection level would be 19 seats because it would be prudent for the airline to reserve this amount of seats for the highest fare class; therefore, the corresponding expected marginal revenue from the first 19 seats is greater for the first fare class than for the second fare class.

The application of the expected marginal seat revenue to revenue management can be presented through the use of a decision tree, as in Figure 11.15. In essence, every seat has various probabilities of being booked, and an airline revenue management analyst must choose the option that provides the airline with the greatest expected seat revenue. Since demand is not deterministic, revenue management analysts must use probability to foresee the future and be able to protect various quantities of seats for higher-paying customers. This protection of seats is in addition to the previously discussed "fences," such as a Saturday night stay requirement.

The decision tree scenario presented in Figure 11.15 assumes that the airline is presented with the situation that it can either sell a fare, or open up a fare bucket, for a discounted price of $200, or not sell the ticket at the discounted price. The key in determining which situation to choose is the probability of selling the discounted ticket and the full price ticket. In our scenario, the probability of selling the discounted ticket at $200 is 100 per cent, and the probability of selling the full-price ticket at $500

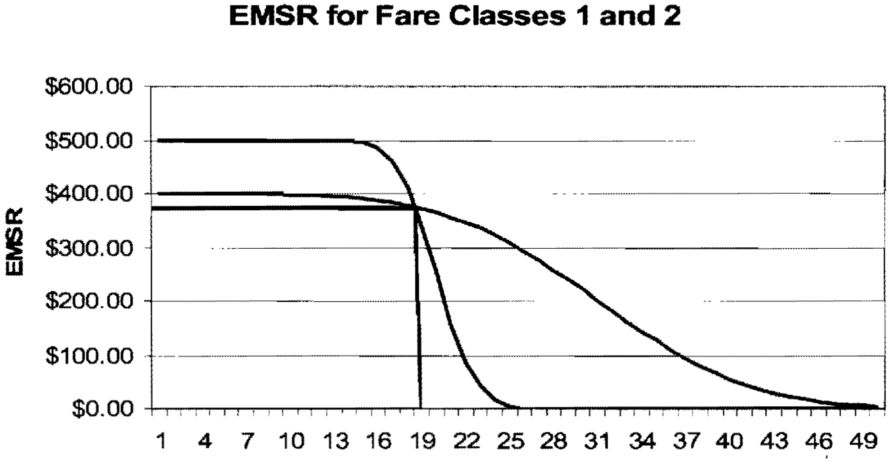

Figure 11.14 Optimal booking limit and protection level in two nested fare classes

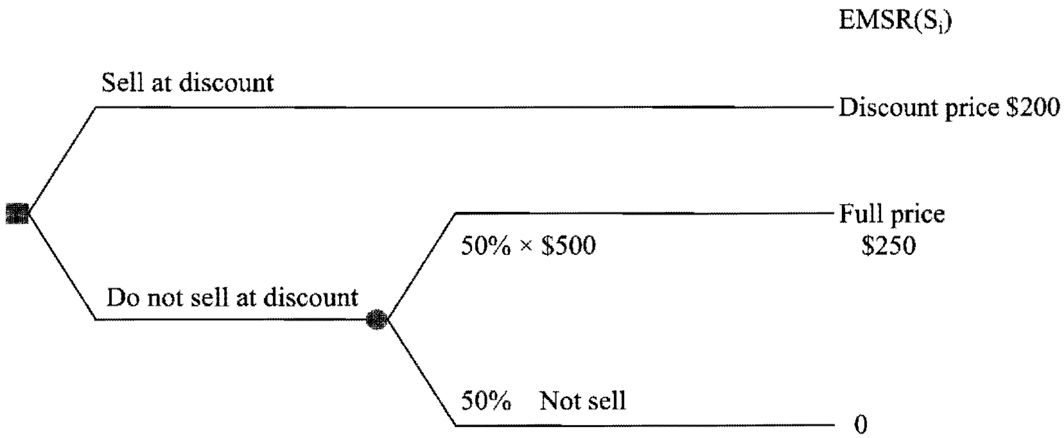

Figure 11.15 Decision tree

is 50 per cent. In addition there is a 50 per cent chance that a full-price ticket will not be sold. In order to make a decision, the expected revenue needs to be computed. This is done by simply multiplying the probability of selling the ticket by the ticket price. This provides expected revenue of $200 for the discounted ticket and expected revenue of $250 for the full-price ticket; therefore, the discounted ticket should not be sold in the hope of selling the full-price ticket. Thus, the formula for computing EMSR can be stated as:

$$EMSR(S_i) = f_i * P(S_i) + 0*[1-P(S_i)]$$

where: EMSR(S_i) is the product of the average fare level, f_i, and the probability of selling the i-th seat P(S_i) plus the fare associated with the alternative event of not selling the seat (0) times the probability of that event [1-P(S_i)].

In a situation where the capacity of the aircraft is increased by one seat, the revenue management analyst must choose the highest marginal EMSR that the seat would generate.

In order to better understand the EMSR approach, we will walk through a simple exercise with the goal of determining the appropriate booking limits and protection levels for every fare class.

DirectJet is a new airline that operates an 80-seat regional jet aircraft. The airline utilizes a nested three-tier fare structure ($300, $400, $500), and demand for all three fare classes is assumed to be normally distributed. The airline has only been in operation for a few months, but has historical demand data for the past 30 days for one of its routes, which is presented in Table 11.8.

The first step required in determining the optimum booking limit and protection level for DirectJet's flight is to determine the mean and standard deviation of the demand for each fare class. This data will be required when determining the probabilities of purchasing the ticket in each fare class. Both the mean and standard deviation for the three fare classes are contained in Table 11.9.

The next step in the EMSR process is to create a normal distribution for each fare class in order to assign a probability that a least certain number of seats would be purchased

Table 11.8 Historical demand for DirectJet's three fare classes

Flight History	$500	$400	$300
Day	Fare	Fare	Fare
1	20	30	36
2	17	40	34
3	18	35	33
4	22	25	32
5	24	18	31
6	20	45	40
7	20	32	42
8	21	22	29
9	22	29	32
10	19	34	38
11	18	38	40
12	20	31	38
13	19	22	36
14	22	24	30
15	18	29	26
16	23	30	26
17	24	36	30
18	23	35	31
19	17	26	33
20	20	42	34
21	19	22	37
22	20	18	35
23	17	34	45
24	18	33	27
25	23	48	28
26	21	16	34
27	20	26	36
28	20	42	33
29	18	38	37
30	17	32	29

Table 11.9 Mean and standard deviation

	$500 Fare	$400 Fare	$300 Fare
Mean	20	31.07	33.73
Standard Deviation	2.13	8.17	4.7

for each fare class. Knowing the mean and standard deviation, a cumulative probability distribution for each fare class can be created using the NORMDIST function in Microsoft Excel. The cumulative probability of selling each seat is then multiplied by the fare to produce the EMSR for each fare class. The data is presented in Table 11.10.

The final step is to determine the appropriate booking limit and protection level for each fare class, based on the EMSR values calculated. The goal when choosing the appropriate level is to select the highest EMSR value, regardless of which column the value lies in. This process continues until all 80 seats (assuming no overbooking) are allocated. The shaded region of Table 11.10 represents the greatest 80 EMSR values for the flight. For DirectJet, the appropriate protection levels for the flight would be 20 $500 fares, 30 $400 fares, and 30 $300 fares. The appropriate booking limits for each class would be 80 for the $500 fare, 60 for $400 fares, and 30 for the $300 fares.

The EMSR curves can also be displayed graphically, as in Figure 11.16. Graphing the EMSR curves provides a visual method of determining the protection level and booking limit for each fare class. The initial intersection point between the $500 fare EMSR and the $400 fare EMSR represents the protection level between the two highest fare classes and occurs at roughly 18 seats. This point, however, does not represent the complete protection level for the $500 fare class as computed in the exercise. This exists because there is a second protection level for the $500 fare class between the $500 fare EMSR and the $300 fare EMSR, which is represented by the inflection point between the $500 fare EMSR and the $300 fare EMSR, and occurs at 20 seats. Therefore, the true protection level for the $500 fare, as computed in the sample problem, is the inflection point between the $500 fare EMSR curve and the $300 fare EMSR curve. Mathematically, the protection level for the $500 fare is:

$$\text{Protection level of \$500 fare class} = S^{\$500}_{\$400} + S^{\$500}_{\$300}$$

where: $S^{\$500}_{\$400}$ represents the protection level between the $500 fare class and the $400 fare class and $S^{\$500}_{\$300}$ represents the protection level between the $500 fare class and the $300 fare class.

Similar methodologies can be used to find the protection level and booking limit for the $400 fare class.

The example of DirectJet helps show how booking limits and protection levels are determined in revenue management. Of course, airlines utilize many more fare classes than the three used by DirectJet, making the process all the more complicated, but the principles behind it are exactly the same. The one glaring omission from the DirectJet example is the presence of overbooking. Overbooking is a real issue faced by revenue management analysts and is usually taken into account when setting the appropriate protection levels for a flight. The issue of overbooking is explored in the next section where the DirectJet example will be further expanded.

Table 11.10 Optimal booking limit for DirectJet's three fare classes

Seat	$500 Fare Probability	$500 Fare EMSR	$400 Fare Probability	$400 Fare EMSR	$300 Fare Probability	$300 Fare EMSR
1	1.0000	$500.00	0.9999	$399.95	1.0000	$300.00
2	1.0000	$500.00	0.9998	$399.92	1.0000	$300.00
3	1.0000	$500.00	0.9997	$399.88	1.0000	$300.00
4	1.0000	$500.00	0.9995	$399.81	1.0000	$300.00
5	1.0000	$500.00	0.9993	$399.71	1.0000	$300.00
6	1.0000	$500.00	0.9989	$399.57	1.0000	$300.00
7	1.0000	$500.00	0.9984	$399.35	1.0000	$300.00
8	1.0000	$500.00	0.9976	$399.04	1.0000	$300.00
9	1.0000	$500.00	0.9965	$398.61	1.0000	$300.00
10	1.0000	$500.00	0.9950	$398.01	1.0000	$300.00
11	1.0000	$499.99	0.9930	$397.18	1.0000	$300.00
12	0.9999	$499.96	0.9902	$396.06	1.0000	$300.00
13	0.9995	$499.75	0.9865	$394.58	1.0000	$300.00
14	0.9976	$498.79	0.9816	$392.64	1.0000	$300.00
15	0.9905	$495.27	0.9753	$390.13	1.0000	$299.99
16	0.9698	$484.90	0.9673	$386.94	0.9999	$299.98
17	0.9205	$460.25	0.9574	$382.94	0.9998	$299.94
18	0.8261	$413.06	0.9450	$378.01	0.9996	$299.88
19	0.6806	$340.32	0.9300	$372.02	0.9991	$299.74
20	0.5000	$250.00	0.9121	$364.84	0.9982	$299.47
21	0.3194	$159.68	0.8909	$356.37	0.9966	$298.98
22	0.1739	$86.94	0.8663	$346.53	0.9937	$298.11
23	0.0795	$39.75	0.8381	$335.25	0.9887	$296.62
24	0.0302	$15.10	0.8063	$322.53	0.9807	$294.22
25	0.0095	$4.73	0.7710	$308.40	0.9683	$290.49
26	0.0024	$1.21	0.7323	$292.92	0.9499	$284.97
27	0.0005	$0.25	0.6906	$276.23	0.9238	$277.14
28	0.0001	$0.04	0.6462	$258.49	0.8885	$266.55
29	0.0000	$0.01	0.5998	$239.92	0.8428	$252.85
30	0.0000	$0.00	0.5519	$220.76	0.7863	$235.88
31	0.0000	$0.00	0.5033	$201.30	0.7194	$215.81
32	0.0000	$0.00	0.4545	$181.82	0.6437	$193.12
33	0.0000	$0.00	0.4065	$162.61	0.5619	$168.58

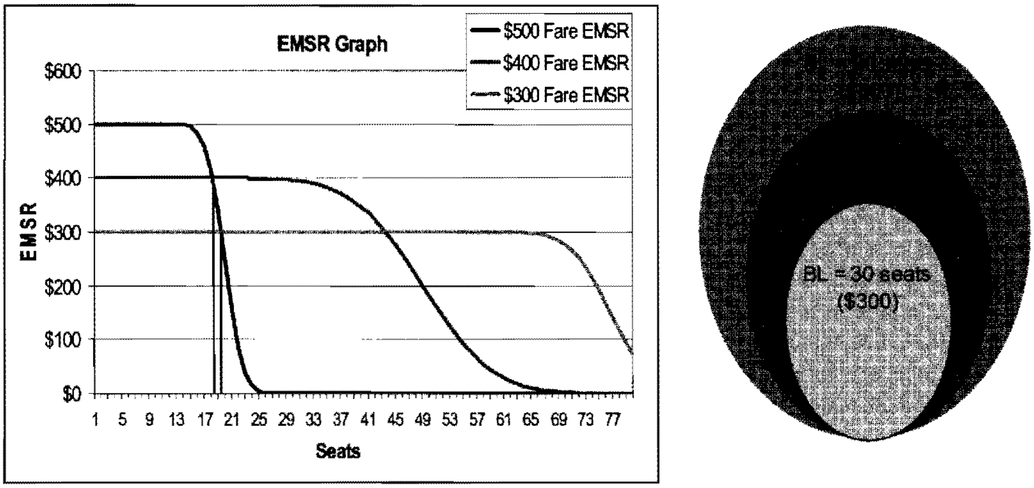

Figure 11.16 Optimal booking limit and protection level in three nested fare classes

OVERBOOKING

If anyone has ever experienced an involuntary denied boarding, they probably do not appreciate the airline's policy of overbooking. Overbooking is practiced by airlines to combat spoilage as invariably some passengers do not show up for a flight or miss connections, especially for flights departing from hubs. Although overbooking may cause headaches for a few passengers, the benefits to the airline include increased seat availability, more access to the flight of first choice, and the reduced overall cost of travel through the more efficient use of airline seats (Dunleavy, 1995). Overbooking is also practiced in other industries, such as car rentals and hotels (Netessine and Shumsky, 2002).

Airlines are able to predict to some degree the "no-show" level, or percentage of passengers who will fail to show up at the gate for the flight, based on the probabilistic nature of overbookings. Furthermore, different passenger groups have various patterns of not showing up for a flight. Typically, business passengers have a higher probability of missing their flight than leisure passengers. Therefore, it is not surprising that flights to leisure markets have lower authorization levels for overbooking than flights to business-heavy markets. Revenue managers must also take into consideration the probability of passengers cancelling itineraries close to departure, misconnecting, or having ticket issues where they show up for the flight but do not have confirmed reservations. These ticket issues are known as a "go-show" in the airline industry (Dunleavy, 1995).

While the benefits of overbooking for the airline is reducing spoilage and increasing potential revenue, the trade-off is that overbooking can also be terribly costly for the airline. Costs associated with overbooking include meal and hotel vouchers, flight coupons for future flights, departure delays, passengers being rolled over to other flights, staffing issues, and loss of goodwill. The level of these costs vary from flight to flight, since the costs associated with a daily (or less frequent) international flight are invariably higher than those associated with a short-haul flight which runs ten times a day.

If more passengers show up for a flight than there are seats, the airline will ask for volunteers to give up a seat. In order to entice passengers to do so, the airline will usually offer some form of compensation package that may include future travel discounts, meal vouchers, hotel accommodation, or first-class upgrades. Ideally, airlines want to solve their overbooking problem by asking for volunteers, as the next step, involuntary denied boarding, can be much more costly. If there are not enough passengers willing to give up their seat, some passengers will be denied boarding. This situation creates extremely negative goodwill against the airline, and the airline is still required to transport the passenger to their destination. Since airlines want to avoid overbooking situations as much as possible, effective forecasting is necessary.

In order to understand what level of overbooking the airline should set for a flight, we will look at the DirectJet example once again. By looking at historical data, DirectJet has been able to determine that the average number of no-shows for one of its flights is normally distributed with a mean of 20 and a standard deviation of 10. Moreover, DirectJet estimates that it costs $900 to "bump" a passenger (the airline receives no revenue from the passenger—$900 is the cost of accommodating the passenger for his/her next flight). On the other hand, if the seat is not sold, then the airline loses revenue equal to the price of a ticket at the discounted rate. In this example, the lowest discounted air fare is $300. The EMSR analysis can actually be described as a news vendor problem. From the news vendor analysis, the optimal protection level is the smallest value Q* such that:

$$F(Y) \geq \frac{B}{B+C}$$

where: Q* = the optimum number of seats to overbook
B = the opportunity cost of flying an empty seat
C = the negative goodwill and penalties associated with denying a passenger boarding (Winston, 2007).

Applying this formula to DirectJet's analysis, the corresponding Q* would be:

$$F(Q) \geq \frac{300}{300+900} = 0.25$$

From the normal distribution table the z value of 0.25 is approximately —0.675, creating the optimal number of seats to overbook of:

$$S_{oberbook} = \mu - Z \times \sigma$$
$$S_{overbook} = 20 - 0.675 \times 10 = 13.25$$

Based on this formula, DirectJet should overbook 13 passengers for this flight.

The other, and simpler, option, is to use the NORMINV function in Microsoft Excel (Netessine and Shumsky, 2002). Either way, the optimum number of seats for DirectJet to overbook is equal to 11.25. Therefore, the total authorization level for the flight would be 93 seats (capacity of aircraft plus overbooking level) if we round the overbooking level down. This solution computes the trade-off between overbooking passengers and the additional revenue they would generate against the costs associated with overbooking. Overbooking by 13 additional seats would allow the revenue management analyst to authorize six additional seats to be sold in both the $300 and $400 fare levels, and one additional seat in the $500 fare level, based on the EMSR values contained in Table 11.20. This would generate an additional $2,018 in potential revenue for DirectJet.

OTHER ISSUES ASSOCIATED WITH REVENUE MANAGEMENT

Since revenue management was first introduced by American Airlines in 1985, it has grown considerably in scope and complexity in order to encompass more issues and better minimize consumer surplus. While current revenue management practices have been effective in helping airlines' bottom lines, revenue management is still largely based on historical, probabilistic demand rather than forward-looking, future demand. While this is not an ideal situation, it probably is the best way to forecast demand for revenue management purposes. In addition, the creation of fare classes is done by grouping customers together based on their price elasticity, but even with multiple fare classes, customers are still going to be grouped into classes to which they ideally do not belong. This can increase consumer surplus, which reduces revenue for the airline. A potential solution to both these problems is dynamic pricing where seats are priced based on passengers' demand and other factors such as competitors' revenue management strategy (Burger and Fuchs, 2005). In essence, dynamic pricing is forward-looking and allows a carrier to closely match its normal booking curve. A few airlines, mostly low-cost carriers, have implemented dynamic pricing into their revenue management models.

A recent trend in airline revenue management that was started by low-cost carriers such as Southwest is to reduce the complexity of their fare structures. In the realization that the average passenger has long been confused by a seemingly infinite number of booking classes and fares for a particular flight, airlines such as Delta Air Canada are reducing the number of booking classes and simplifying fare structures. These new simplified fare structures, termed restriction-free pricing, are forcing revenue management to use "weak" market segmentation through active management of fare availability instead of "strong" market segmentation, such as "fences" (Ratliff and Vinod, 2005). Although this move has proved generally more appealing to passengers, it remains to be seen if the increase in consumer surplus from simpler fare structures can be offset by any incremental increase in bookings. Moreover, the move by legacy carriers to allow one-way tickets has introduced further complexities into revenue management systems.

Revenue management, because it is responsible for ensuring that customers purchase the airline's product, is one of the most important business units in an airline organization. A revenue management analyst can, on a daily basis, have more impact on the company's bottom line than almost any other employee. Ever since revenue management was introduced by American Airlines to combat People Express, it has been an effective tool used by airlines to maximize revenues and enhance their bottom line.

SUMMARY

This chapter has presented an in-depth discussion of revenue management as it is now practiced in the airline industry. The topic of price discrimination was reintroduced as the basis for all revenue management schemes. This was followed by a thorough discussion of how revenue management is implemented to include "fencings", other control types and the leg-based EMSR model. Finally, other issues associated with revenue management were discussed in some detail.

REFERENCES

Belobaba, P. (1987). *Air Travel Demand and Airline Seat Inventory Management*. Cambridge, MA: Flight Transportation Laboratory, Massachusetts Institute of Technology.

Burger, B. and Fuchs, M. (2005). Dynamic Pricing—A Future Airline Business Model. *Journal of Revenue and Pricing Management*, 4(1), pp. 39–53.

Buzzacchi, L. and Valletti, T. (2005). Strategic Price Discrimination in Compulsory Insurance Markets. *Geneva Risk and Insurance Review*, 30(1), pp. 71–96.

Cross, G. (1995). An Introduction to Revenue Management. In D. Jenkins (ed.), *Handbook of Airline Economics*. New York: McGraw-Hill, pp.443–58.

Dunleavy, N. (1995). Airline Passenger Overbooking. In D. Jenkins (ed.), *Handbook of Airline Economics*. New York: McGraw-Hill, pp. 469–76.

Frank, R. (1983). When are Price Differentials Discriminatory? *Journal of Policy Analysis and Management*, 2(2), pp. 238–55.

James, G. (1988). The Critical Importance of Airline Revenue Enhancement: A U.S. View. Paper presented at Airlines Group of International Federation of Operational Research Societies, AGIFORS Annual Symposium, Cape Cod, MA.

Littlewood, K. (1972). Forecasting and Control of Passenger Bookings. *12th AGIFORS Symposium Proceedings*, pp. 103–5.

Littlewood, K. (2005). Forecasting and Control of Passenger Bookings. *Journal of Revenue and Pricing Management*, 4(2), pp. 111–23. Originally written in 1972.

Loveman, G. and Beer, M. (1991). People Express Airlines: Rise and Decline. *Teaching Note*. Harvard Business School 5-491-080.

Netessine, S. and Shumsky, R. (2002). Introduction to the Theory and Practice of Revenue Management. *INFORMS Transactions on Education*, 3(1), pp. 34–44.

Ratliff, R. and Vinod, B. (2005). Airline Pricing and Revenue Management: A Future Outlook. *Journal of Revenue and Pricing Management*, 4(3), pp. 302–7.

Spiller, T. (1981). The Differential Inputs of Airline Regulation on Industry, Firms, and Market. *Journal of Law and Economics*, 24, pp. 655–84.

Vinod, B. (1995). Origin-and-Destination Revenue Management. In D. Jenkins (ed.), *Handbook of Airline Economics*. New York: McGraw-Hill, pp. 459–68.

Winston, W. (2007). *Operations Research: Applications and Algorithms* (4th edn). Pacific Grove, CA: Duxbury Press.

12

Low-cost, Start-up Airlines: A New Paradigm

> ...the twentieth century largely belonged to the traditional, high-cost airlines (with a few snipers, like me, upsetting their cozy cartel). The twenty-first century will be the preserve of the no-frills airlines...
>
> Sir Freddie Laker, cited in Calder (2002)

One of the more recent developments in the aviation industry has been the emergence of a new breed of airlines—low-cost carriers. While only the last decade has seen tremendous growth in this sector of the industry, the roots of low-cost carriers can be traced back to 1971 when founder Herb Kelleher mapped out the route and cost structure for Southwest Airlines. Since the US aviation industry was still regulated, Southwest Airlines was able to fly only intra-Texas routes where the Civil Aeronautics Board (CAB) did not have authority. From these humble beginnings, and based on its simple low-cost strategy, Southwest Airlines was able to grow into one of the most successful airlines in the United States. The overall objective of this chapter will be to introduce the various strategies used by low-cost airlines to gain competitive advantages over legacy airlines. The chapter starts by providing information on "legacy" airlines and how the emergence of low-cost airlines affected the market share of legacy airlines. The general outline for this chapter is:

- The emergence of low-cost carriers
- Characteristics of low-cost carriers, including:
  - Lower labor costs per hour of productivity
  - Lower ticket distribution costs
  - No-frills service
  - Common fleet type
  - Origin and destination route structure
  - Use of secondary airports
  - Increased aircraft utilization
- Cost structure comparison
- Incumbent carriers' response to low-cost carriers, including:
  - LCC creation
  - Cost-cutting
- The future of low-cost carriers

THE EMERGENCE OF LOW-COST CARRIERS

Following airline deregulation in the United States, many new airlines entered the market, causing airfares to plummet 40 per cent in real terms between 1978 and 1997, and more than doubling the number of passengers. Airlines like Southwest were able to successfully expand, while new low-cost carriers like America West, Reno Air, and People Express emerged with varying degrees of success. Today there are many successful low-cost carriers (LCCs) in the United States, such as AirTran, Frontier, and JetBlue, which have taken away market share from the traditional legacy carriers. On the other hand, despite a huge demand for low-cost travel, many low-cost carriers have failed. Predictably, the remaining low-cost carriers have been those with the lowest cost base. For example, an airline like Southwest has been tremendously successful at retaining a low cost structure while still expanding aggressively.

The emergence of low-cost carriers is not just a North American phenomenon, but a global trend in the airline industry: today almost all markets contain at least some low-fare carriers. After Southwest founded the low-cost carrier model, the idea quickly spread to the UK in the early 1970s when Sir Freddie Laker was able to secure the necessary route authorizations to launch cheap transatlantic flights between Gatwick and New York. With North America experiencing a wave of low-cost carrier start-ups following US airline deregulation, Europe experienced a "second-wave" of low-cost carriers following the liberalization of European airspace. Low-cost carriers such as Ryanair and easyJet expanded rapidly and, in the process, acquired market share from Europe's large established carriers.

Not only have these airlines been successful at acquiring market share, but they have also been profitable. Table 12.1 provides a brief synopsis of the financial situation for some of the world's major low-cost carriers. As the table shows, the founding airline, Southwest, is still the most profitable airline in the low-cost category. It is followed by European discounter Ryanair and fledging Brazilian low-cost carrier, Gol.

Table 12.1 Low-cost carriers: operating profit

2004			2005		
Rank	**Airline**	**USD million**	**Rank**	**Airline**	**USD million**
1	Southwest	554.00	1	Southwest	820.00
2	Ryanair	397.08	2	Ryanair	445.75
3	Gol	247.56	3	Gol	265.46
4	Virgin Blue	168.79	4	Virgin Blue	117.09
5	JetBlue	110.89	5	JetBlue	85.85
6	easyJet	89.02	6	easyJet	52.20
7	Air Tran Holdings	32.84	7	Air Tran Holdings	47.61
8	AirAsia	16.01	8	AirAsia	35.38
9	WestJet	(8.49)	9	WestJet	13.39
10	SkyEurope	(16.24)	10	SkyEurope	(7.90)
11	Frontier	(26.45)	11	Frontier	(40.44)

Source: ATW World Airline Report 2005.

As mentioned above, while many low-cost carriers have been successful, the list of failed low-cost carriers is long. In North America, major carriers, like US Airways, Delta, Continental, and Air Canada, all attempted low-cost subsidiary airlines and all failed. Moreover, it was not only the major airlines whose attempts to compete with LCCs failed, but also independent LCCs like Reno Air, People Express, and Independence Air. In Europe, KLM divested itself of Buzz while British Airways did the same with its low-cost carrier Go. Furthermore, the list of failed low-cost carriers in Europe include Airlib Express, BerlinJet, Fresh Aer, and Goodjet, to name a just few. This raises the obvious question of why some LCCs succeeded, while others failed.

Part of the problem stems from the fact that many airlines (particularly major carriers) that tried to imitate the low-cost model never truly adopted it. That is, they tried to establish the LCC with much of their existing cost structure. This inevitably led to their failure. Moreover, regardless of the type of carrier, any new airline entrant will face tremendous competition from the incumbent airlines, making any new start-up airline's probability of success slim. Although every carrier is unique, there are certain common characteristics that have enabled certain LCCs to succeed where others have failed. In general, these characteristics, discussed in more detail below, allowed the carriers to maintain a low cost structure.

CHARACTERISTICS OF LOW-COST CARRIERS

> Most people think of us as this flamboyant airline, ... but we're really very conservative from the fiscal standpoint ... We never got dangerously in debt and never let costs get out of hand.
>
> Herb Kelleher

Despite the fact that LCCs operate all over the globe in different environments, they all exhibit a few basic general characteristics.

Lower Labor Costs per Hour of Productivity

Since labor costs are one of the largest costs for any airline, it is imperative for low-cost carriers to keep their labor costs under control and/or increase labor productivity. While many LCCs simply pay lower than industry average wages, Southwest has proved that low-cost carriers can pay competitive rates, yet still have low labor costs per hour of productivity. By having high employee productivity, Southwest has been able to pay high salaries, yet remain very competitive on a per block hour basis. For example, Southwest pilots are some of the highest paid in the industry, yet they fly many more hours per month than their counterparts at other airlines and they also pitch in to help do things such as clean the aircraft or carry bags.

Although employee productivity tells some of the story, the importance of low labor costs cannot be overstated. Any airline with high labor rates and moderate to low productivity will ultimately be unsuccessful. Both Air Canada and Delta launched low-cost carriers that utilized employees from the mainline carrier. In essence, the new carrier was supposed to be low-cost, but had the same employee group as the "high-cost" airline. Since these new low-cost carriers were uncompetitive on the labor front and had no new productivity increases, it is not surprising that they disappeared shortly after their start-up.

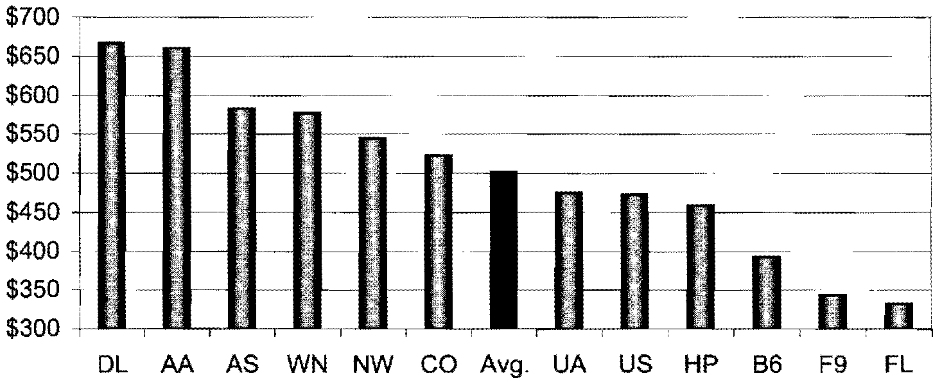

Figure 12.1 Crew costs per block hour, 2006

Source: Compiled by the authors using Back Aviation Form41 data.

Figure 12.1 provides a comparison of labor costs per block hour for various US airlines, including both legacy carriers and low-cost carriers. While JetBlue, AirTran, and Frontier are the three leading carriers in terms of crew cost per block hour, Southwest Airlines is in the middle of the pack.[1] The relatively low labor costs for the first three airlines represent a significant competitive advantage, while Southwest Airlines, a more mature airline, must focus on employee efficiency to help offset the relatively higher crew costs per block hour. Not surprisingly, Delta and Northwest had the highest crew costs in 2005, which happens to coincide with the year that both airlines entered bankruptcy protection.

Lower Ticket Distribution Costs

Ticket distribution costs are another major area where the entire airline industry is attempting to reduce costs. The initial step that airlines took to reduce ticket distribution costs was to cut travel agent commissions. Then, through the Internet, airlines moved to electronic ticketing and pushed ticket sales through their online websites. However, most of the major carriers still rely on GDS (Global Distribution System) providers, such as Sabre and Worldspan, to distribute their tickets worldwide. While GDSs provide an airline with a global reach, it costs close to $13 to distribute a ticket through a GDS as opposed to a mere few dollars through Internet e-ticketing (Ionides and O'Toole, 2005).

LCCs have been far more successful at selling tickets through their online websites than the major carriers. For example, Southwest books about 60 per cent of its revenues through its website, while JetBlue brings in 75 per cent of their revenues through jetblue.com (Field and Pilling, 2005). Compare those percentages to Continental (25 per cent) and Delta (28 per cent), and it is obvious that the LCCs achieve clear cost savings in ticket distribution costs (Field and Pilling, 2005; Ionides and O'Toole, 2005).

One successful strategy used by LCCs is to initially align themselves with multiple GDS providers and, then, as their brand becomes stronger, to slowly end their agreements with them. This enables the carrier to have a wide distribution network initially, and then narrow its distribution (and reduce costs) as it pushes ticket sales towards its website.

1 Note that the appendix contains the two-letter airline code for various airlines.

Ryanair in Europe was very successful with this strategy (Field and Pilling, 2005). On the other hand, low-cost carrier Independence Air decided to begin operations without any GDS distribution, but because its website had no brand awareness, the airline ultimately went out of business (Field and Pilling, 2005).

It is worth noting that a universal push toward online ticket distribution by LCCs is not typical of them all. Since the Internet is not as readily available to consumers in South America and Asia, airlines in these areas still need to rely on travel agencies for ticket distribution.[2] In general, though, LCCs have been the catalyst in the e-ticketing/website distribution world, and this has provided them with a significant ticket distribution cost advantage. Now, with IATA's announcement that it would not accept paper tickets as of 2008, issuing e-tickets has become one of the most important tasks of all airlines, not just LCCs.

No-frills Service

Historically, one of the clearest examples to consumers of the difference between LCCs and legacy airlines was a "no-frills" service. In the United States on a legacy carrier's flight, passengers received a complimentary hot meal with an extensive beverage service whereas on a Southwest flight a passenger would receive peanuts and a soda. However, with the cost-cutting measures implemented by legacy carriers, all economy class service in North America has turned into "no-frills." In Europe, LCCs have gone one step further where everything, including beverages, is on a buy-on-board basis. Therefore, the in-flight food service that used to easily distinguish low-cost airlines from "full-service" carriers is no longer applicable.

However, no-frills service does not just pertain to in-flight service. Many LCCs also do not have frequent-flyer programs or expensive business lounges; these amenities are not offered in order to cut costs. Another cost-cutting measure that has recently been implemented by LCCs is the restriction of luggage allowances. In Europe in particular, LCCs have strict rules concerning luggage allowance weights per passenger; this conserves fuel and generates extra marginal revenue.

The underlying premise behind the LCCs' no-frills service strategy is ultimately a "pay as you go" approach, where the ticket price entitles you to just a seat on the aircraft. As a result of this strategy, LCCs can offer attractive airfares. While these service cuts may seem minimal, when they are compounded over the number of flights, it can actually make the difference between profit and loss.

Common Fleet Type

Another major characteristic of successful low-cost carriers is the use of a common fleet type. Southwest Airlines was the pioneer of this strategy, focusing its entire fleet around the Boeing 737. A single fleet type provides many advantages for an airline; these include a reduction in maintenance spare parts inventories, reduced flight crew training expenses, and increased operational flexibility. In addition, bulk purchase discounts from suppliers (including aircraft manufacturers) can be received when using a single fleet type. However,

2 According to the International Air Transport Association (IATA), only 13 per cent of airlines in the Middle East issued electronic tickets in 2006.

it is economies of scale that are the most important cost reduction elements underlying the common fleet type strategy, in that the airline is required to spend fixed fleet costs—for example, all the specialized equipment that might be needed for a 737—only once.

In addition to economies of scale savings, a single fleet provides increased operational flexibility. In the event of irregular operations, a single fleet type makes it easier to find a replacement aircraft or usually, more importantly, a replacement flight crew. Since airlines usually have a reserve pilot pool for each fleet type, restricting the number of fleet types limits the number of reserve pilots the airline requires.

Using one fleet can also have advantages and disadvantages concerning markets served. Depending on aircraft choice, the aircraft used by the airline may not be the optimal aircraft for some markets. Thus, if the aircraft has a relatively short range, many intercontinental markets will not be feasible. AirTran, for example, had this problem with the 717s and therefore had to purchase another fleet of 737s. Conversely, a single fleet contains aircraft that have the same pilot requirements and maintenance standards. For LCCs, the two most widely used generic aircraft types are the 737NG and the A32X. Both these aircraft types enable a carrier to have planes with as few as 120 seats all the way up to close to 200 seats. This enables them to switch aircraft sizes interchangeably to better meet demand on any given day.

Although a single common fleet has been the LCC standard, both JetBlue and easyJet bucked the trend by creating fleets with two aircraft types because they felt that the economies of scale benefits on their initial fleet had reached a threshold. This occurs when the benefits of the first large fleet type are outweighed by the benefits of a second large fleet type. Based on these examples, it appears that the minimum number of aircraft needed to achieve full economies of scale benefits is probably slightly under 100.

Table 12.2 shows clearly that LCCs in North America have less diverse fleets than the legacy carriers. Part of the reason for the legacy carriers' more complex fleets is that international flying requires larger aircraft, and legacy carriers have undergone more mergers, thereby combining fleets. In general, LCCs are also younger companies, and have emphasized a single fleet strategy. And, while Southwest Airlines is shown as having two fleets (B737CL and B737NG), the airline still operates only one aircraft type since these are two different generations of the same aircraft.

Regardless of aircraft type, low-cost operators configure their aircraft in a high-density all-economy configuration. In some cases, closets and washrooms are removed in order to squeeze more passengers onto the flight. The seats on easyJet and Ryanair also are all non-reclining in order to accommodate more passengers. Obviously, since every flight is largely a fixed cost (once it has been launched), the more people on board, the more revenue the airline can obtain. This in turn enables the airline to offer a few seats at highly discounted prices. In the North American market, the LCCs have pursued different marketing practices—for example, JetBlue and WestJet have removed seats from their aircraft to provide additional legroom in an effort to attract more business clientele. Also, very few discount carriers operate any sort of a premium cabin in the belief that additional economy revenue will always exceed any premium cabin revenue. Tables 12.3 and 12.4 provide seating capacities of airlines based in North America. All the LCCs put more seats in their aircraft than the industry average, thereby spreading costs per seat over a greater number of seats. For example, a Southwest 737-300 has 7 per cent more seats than the industry average, while a JetBlue A320 has roughly 6 per cent more seats than the North American industry average.

Table 12.2 Aircraft fleets for major North American operators, 2006

Carrier	Fleet Types	No. of Fleet Types
Low Cost		
AirTran	B717, B737NG	2
Frontier	A32X	1
JetBlue	A32X, Emb 190	2
Southwest	B737CL, B737NG	2
Westjet	B737NG	1
Legacy		
Air Canada	Emb 175/190, A32X, B767, A330/340-300, A340-500	5
Alaska	MD80, B737CL, B737NG	3
American	MD80, B737NG, B757, B767, B777, A300-600	6
America West	B737CL, A32X, B757	3
Continental	B737CL, B737NG, B757, B767, B777	5
Delta	MD80, MD90, B737NG, B757, B767, B777,	6
Northwest	DC9, DC10, A32X, B757, A330, B747	6
United	A32X, B737CL, B747, B757, B767, B777	6
US Airways	A32X, B737CL, B757, B767, A330	5

Table 12.3 Seating capacity of North American Boeing operators

Aircraft Type	Southwest	AirTran	Westjet	Continental	United	US Airways	Alaska	America West	American	Delta	Industry Average
737-300	137	-		124	120	124	-	134	-	-	128
737-400	-	-		-	-	144	144	-	-	-	144
737-500	122	-		114	104	-	-	-	-	-	113
737-700	137	137	136	124	-	-	124	-	-	-	132
737-800	-	-	166	156	-	-	160	-	142	150	155
737-900	-	-		167	-	-	172	-	-	-	170

Source: Compiled by the authors using seatguru.com and southwest.com.

Table 12.4 Seating capacity of North American Airbus and Embraer operators

Aircraft Type	JetBlue	Frontier	Ted	United	US Airways	America West	Northwest	Air Canada	Industry Average
A319	-	132	-	120	120	124	124	120	123
A320	156	-	156	138	142	150	148	140	147
Embraer 190	100	-	-	-	-	-	-	93	97

Source: Compiled by the authors using seatguru.com.

Origin and Destination Route Structure

Since deregulation, the legacy carriers have adopted a hub-and-spoke route structure; this means that all spoke flights come into one hub airport, and this airport provides the connecting feed for the spoke flights that depart shortly thereafter. While the hub-and-spoke system has been effective for legacy carriers in providing a large number of city-pair connections, a hub is also an extremely expensive operation. Hubs usually have peaks to minimize passenger transfer time, but also downtimes where it is not fully utilizing many of its facilities. This is extremely costly as employees may be idle for extended periods of time and assets (such as gates and ground equipment) may be left unused. Since employees and gates must be paid for over the full working day (and not only when flights are arriving and/or departing) this represents some level of inefficiency. Moreover, the flip-side of this is equally true, and that is that the hub carrier must also have adequate staffing for its peak number of flights, thereby further increasing the cost of unproductive time.

Besides the undoubted revenue advantage of many city-pair choices and the ability to increase load factors by consolidating passengers at the hub, one of the main cost benefits underlying a hub is the ability to realize certain economies of scale. Consolidating operations in one place reduces fixed overhead costs such as required reserve labor pools, maintenance operations, and terminal-related expenses. The problem is that the peak flight scheduling that is necessary for passenger convenience also means that these economies of scale are not always achieved. Moreover, once the number of flights reaches a critical level, diseconomies of scale will set in—that is, any additional flight will in fact increase average costs rather than reduce them—because the added congestion increases costs. For example, as the airport becomes busier and busier, aircraft have to wait longer to land and take off, and this is reflected in higher fuel and labor costs. Arguably, many of the hubs in the United States have surpassed the critical inflection point so that economies of scale have turned into diseconomies of scale. Unfortunately, since there are many different and interactive decision-makers, it is nearly impossible to determine an optimized flight level for a hub airport.

The prevalence of diseconomies of scale mentioned above is one of the main reasons why LCCs typically operate a point-to-point or origin and destination (O & D) route structure. Under a point-to-point route structure, the airline will operate a more spread-out route network and typically will offer nonstop flights between city-pairs. Under this route structure, airlines will still operate bases where economies of scale are realized, but will not have any peak level of flights. This allows the airline to continually use airport facilities and more evenly utilize employee services. This increased utilization of airport assets allows a point-to-point airline to operate more flights with fewer facilities and personnel, and this ultimately reduces costs. Southwest Airlines has sizeable operations at many airports across the United States, but these bases have not grown to the size of the legacy carriers' hubs. Also, Southwest Airlines generally operates at least 8–10 flights out of any city in order to experience some level of economies of scale, spread fixed costs over a greater number of flights, and increase the frequency of flight choice for the passengers.

In North America, while airlines like Southwest generally operate a point-to-point route structure, a good number of passengers still connect on Southwest flights through some of Southwest's larger bases. In Europe by contrast, LCCs typically do not allow any connecting flights, thereby relying solely on origin and destination demand for all its flights. By not connecting passengers, the airline does not have to worry about transferring luggage between aircraft or compensating passengers for misconnections, and this further reduces operating costs. The European low-cost carrier model generally bases a few aircraft at one airport and then flies to various destinations from there. This enables the carrier to receive some of the benefits of economies of scale at these bases. Thus, European LCCs operate a base-and-spoke network with no connections or synergies with the airline's other bases. Both Ryanair and easyJet have been successful using this strategy.

Use of Secondary Airports

Similar in vein to using a point-to-point route structure, the use of secondary airports is another characteristic of LCCs. Congested primary airports usually mean more time on the ground and higher airport fees, so LCCs avoid them where possible. For example, Southwest does sometimes fly into busy airports, such as Los Angeles, but it has avoided flying into Chicago O'Hare (instead, serving Midway airport) and all three of the major New York City airports (choosing instead the Long Island airport). While Southwest pioneered the utilization of secondary airports (saving costs by being able to turn aircraft around quicker), Ryanair has been the most aggressive in serving secondary airports (Doganis, 2001). Examples of Ryanair using secondary airports include its use of Hahn airport for Frankfurt, Charleroi airport for Brussels, Beauvais airport for Paris, and Weeze airport for Dusseldorf. These secondary airports are miles away from the city center, but are practically deserted; therefore the LCC is able to operate at the airport efficiently and cost-effectively.

Since airports benefit immensely from an LCC starting service, they have become very aggressive in wooing new low-cost carrier services. The classic example of this is an agreement Charleroi airport made with Ryanair whereby Ryanair received a reduction in airport charges of around €2 per passenger, a reduction in ground-handling charges to €1 per passenger, one-off incentive bonuses for starting new routes, and marketing promotion (Smith, 2005). In addition, airports have begun to design facilities that cater specifically to the needs of an LCC (that is, with low operating costs). Marseille airport has designed a dedicated low-cost terminal with cheaper passenger service charges, while Geneva airport has opened a terminal for "simplified aviation" (Buyck, 2005). These actions obviously lead to lower overall costs, and these can in turn be passed on to passengers in the form of lower fares

Increased Aircraft Utilization

Another central focus of success for low-cost airlines is a high level of aircraft utilization. Since an aircraft is not earning money while sitting on the ground, the more an aircraft is flying, the more passengers the airline can carry. There are two central ways in which an airline can increase its daily average aircraft utilization: turn the aircraft around quicker or fly longer routes. Figure 12.2 provides a comparison of average daily block hours per aircraft between LCCs and network carriers in the United States for domestic operations only.[3]

Since high aircraft utilization rates are one of the major strategies for any LCC, it is not surprising that the top four airlines in terms of aircraft utilization are all LCCs. Airlines

3 The legacy carriers' aircraft utilization statistics are slightly distorted since their international flights would increase aircraft utilization.

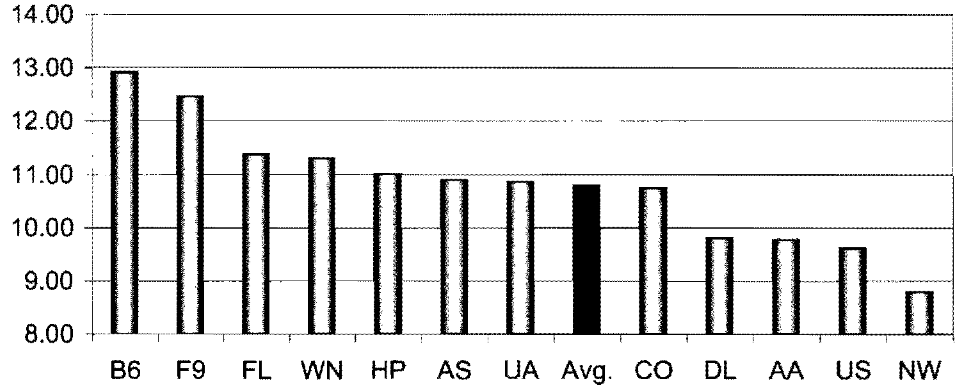

Figure 12.2 Aircraft utilization (block hours per aircraft per day), 2006

Source: Compiled by the authors using Back Aviation Form41 data.

such as Southwest and JetBlue are able to achieve high aircraft utilization rates because they focus on quick turnarounds. Because seats are not pre-allocated, passengers tend to enplane and deplane faster, and by having considerably fewer hub-and-spoke operations, the ground baggage-handling situation is less complex. In addition, using secondary, less congested airports allows the airlines to schedule more flights, since there is less delay in the overall schedule as is the case with the more heavily congested main airports . These efficiencies enable LCCs to operate more flights, and thereby earn more revenue. The downside to increased utilization is that maintenance costs increase since the aircraft are being flown harder—a trade-off that most airlines, however, are willing to make.

The other method of increasing aircraft utilization is to fly longer routes. Figure 12.3 provides a comparison of the average domestic stage length for major US carriers. It is interesting to note that two LCCs, JetBlue and Southwest, are on opposite ends of the average stage length chart, but both are able to maintain high levels of aircraft utilization. JetBlue accomplishes high aircraft utilization rates by flying transcontinental and Florida flights from its New York JFK base, while Southwest obtains high aircraft utilization by operating short flights with quick turnarounds (such as intra-Texas flights). The fact that a Southwest aircraft will be landing and departing more frequently in a day than a JetBlue aircraft highlights Southwest's tremendous efficiency in its ground-handling operations. Only recently has Southwest ventured into the transcontinental flying market in an effort to increase its average stage length and increase aircraft utilization.

Figure 12.4 displays the results of a correlation analysis between operating cost per available seat mile (CASM) and aircraft utilization. CASM is expressed in cents to operate each seat mile offered, and is calculated by dividing operating costs by available seat miles (ASMs). The regression line is plotted in the figure, and it is evident that a strong negative relationship exists between increased aircraft utilization and reduced operating costs. This trend is confirmed statistically by having a significant R-squared value of 0.59. Airlines that lie below the trend line have a lower operating CASM for their level of aircraft utilization than the industry trend. Not surprisingly, LCCs such as JetBlue and Southwest lie at the bottom end of the trend line, while legacy carriers are generally at the top. More specifically, Figure 12.4 shows clearly that legacy airlines, such as Continental and United, need to reduce their operating CASM for their level of aircraft utilization, as they lie significantly above the trend line.

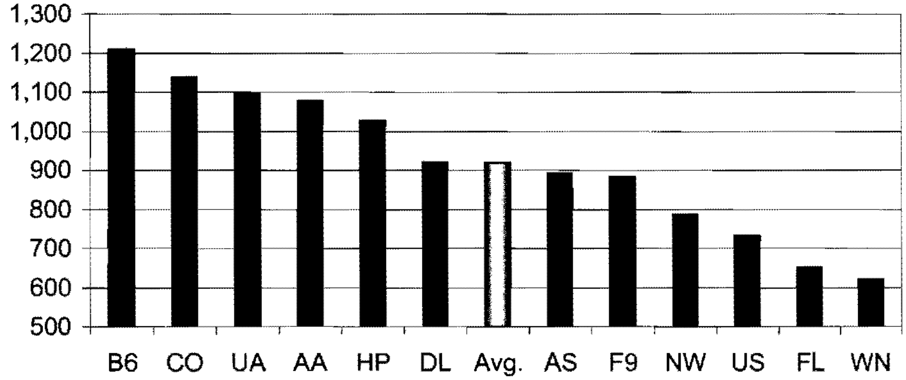

Figure 12.3 Average aircraft stage length in miles, 2006

Source: Compiled by the authors using Back Aviation Form41 data.

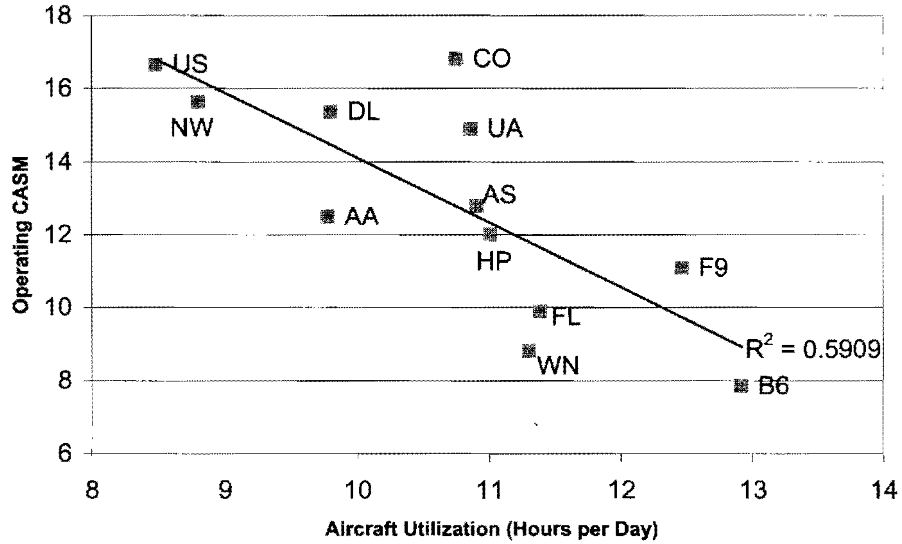

Figure 12.4 Correlation between operating CASM and aircraft utilization, 2006

Source: Compiled by the authors using Back Aviation Form41 data.

The number of aircraft departures per day is also directly related to aircraft utilization—that is, the more departures per day per aircraft, the higher the utilization and the lower the CASM. Figure 12.5 shows this negative relationship between operating CASM and departures per aircraft per day. While this relationship is not as strong as the previous one, it still has a significant R-squared value of 0.39. Carriers such as Southwest and AirTran have low operating costs per ASM with a relatively high number of departures per aircraft. Conversely, the majority of the legacy carriers are all grouped in the upper-left quadrant above the trend line. The major outlier in this analysis is JetBlue. However, this can be explained by the fact that Jet Blue has pursued a strategy of lowering operating CASM by flying fewer departures per aircraft but with much longer average stage lengths.

A third correlation to operating CASM is fuel efficiency—that is, the number of available seat miles flown on one gallon of fuel. As expected, the general relationship between operating CASM and fuel efficiency is downward-sloping, since the more fuel-efficient an airline is, the lower its operating costs. Figure 12.6 displays the relationship, which contains an R-square value of roughly .21. Based on the trend line, all the LCCs outperform the market in terms of fuel efficiency since their observations all fall below the trend line. Of all the legacy carriers, American Airlines is the only carrier whose fuel efficiency–operating CASM relationship falls below the industry trend line.

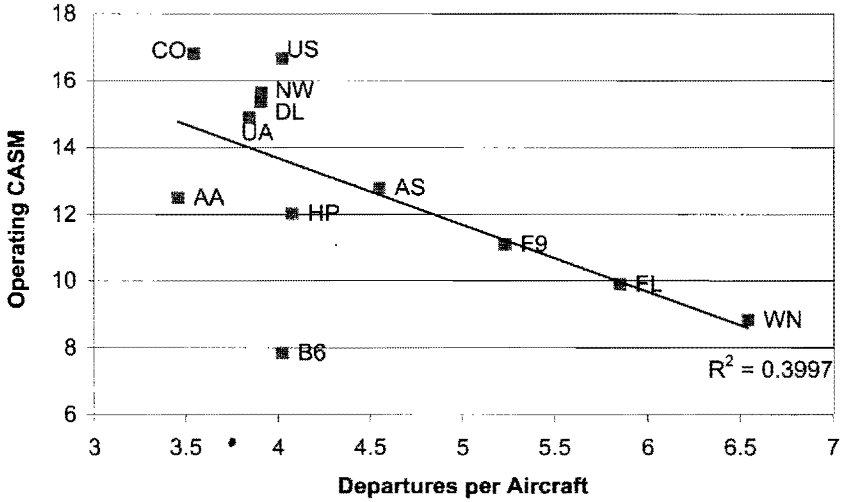

Figure 12.5 Correlation between operating CASM and departures per aircraft per day

Source: Compiled by the authors using Back Aviation Form41 data.

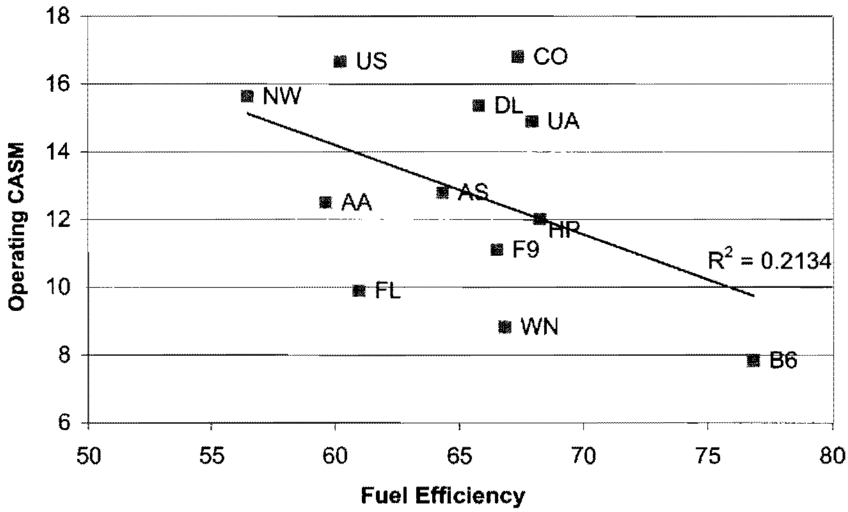

Figure 12.6 Correlation between operating CASM and fuel efficiency, 2006

Source: Compiled by the authors using Back Aviation Form41 data.

COST STRUCTURE COMPARISON

In the modern competitive aviation world, the most successful airlines are those with the lowest cost structure. Southwest, for example, has a long record of continually maintaining low operating costs, so that it can offer attractive airfares. Figures 12.7 through 12.10 provide a comparison of direct operating costs per ASM for four different aircraft that are commonly used throughout the United States by legacy and LCCs. The direct operating costs of the aircraft include fuel, labor costs of flying, all maintenance costs, and all ownership or leasing costs.[4]

Figure 12.7 provides a comparison of airlines using the older 737-300 aircraft. The figure shows clearly that not only is Southwest the largest operator of 737-300 aircraft, but it is also the most cost-efficient. Southwest has kept direct operating cost per seat mile at under seven cents per seat mile since 2000; this is especially significant because the aircraft are getting older, requiring more maintenance and increased fuel burn. While all the airlines exhibit a slight upward trend, (probably a result of increased fuel prices) both Delta and US Airways have had difficulty in controlling the escalating costs. Partly as a result of this, Delta retired its 737-300 fleet, with the last aircraft leaving the fleet in the second quarter of 2006. United has also used its 737-300 aircraft in its shuttle operation, but its costs per seat mile remain high. A possible explanation for the higher legacy carrier costs is that its 737-300s fly shorter routes in some parts of the country. This reduces seat miles in the face of relatively high fixed costs of operation and raises the cost per available seat mile (CASM). Nevertheless, Southwest Airlines also utilizes its 737-300 fleet on some shorter routes, but is still able to be a cost leader.

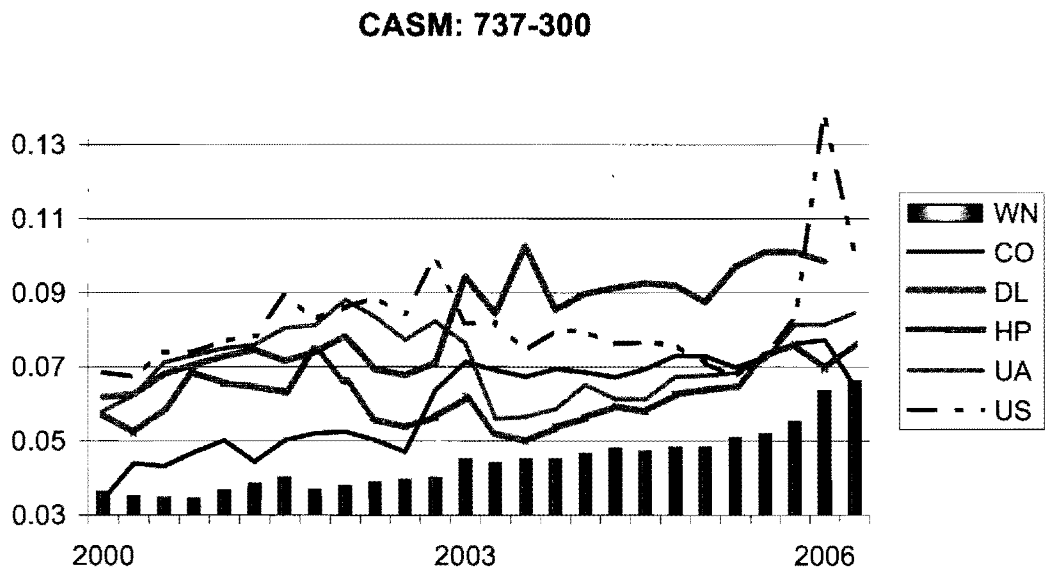

Figure 12.7 Cost per available seat mile for US 737-300 operators, 2000–2006

Source: Compiled by the authors using Back Aviation Form41 data.

4 All associated costs and ASM are for domestic US operations only.

Figure 12.8 provides a similar comparison of CASM for airlines using the newer Boeing 737-700 aircraft, and, again, Southwest Airlines is the cost leader. It is interesting to note that the CASM for the -700 series is at least two cents less than the CASM for the -300 series, which is a similar size aircraft. This could be the result of two factors: longer flights and/or increased efficiency. Since the -700 aircraft have a greater range, they typically fly longer flights, thereby increasing the ASMs while keeping flight costs relatively fixed. Also, the new generation of aircraft is much more fuel efficient so that fuel costs are decreased, resulting in an overall decrease in CASM.

One other interesting observation from Figure 12.8 is that, while the initial CASM for low-cost carrier AirTran is high, it eventually drops after the aircraft becomes fully operational. This reflects the initial start-up costs that must be incurred before the aircraft starts intensive scheduled flying. Although the gap between AirTran and Southwest is now significant, it most probably will be reduced as AirTran increases its fleet and receives benefits from economies of scale.

Figure 12.9 shows similar CASM cost comparisons for A319 operators. Frontier is the primary low-cost operator of the A319 in North America, but its cost structure is similar to United, America West, and Northwest. This is probably due to the fact that Frontier started as a short-haul legacy carrier and is yet to make the complete transition to an LCC. Only US Airways has a cost structure that is well above the rest of the industry.

The Airbus 319 also enables comparison to the Boeing 737-700, since they are both modern aircraft with similar seating capacities. Comparisons are complicated because the aircraft are operated by different airlines, but it appears that the general trend is that the Airbus 319 initially has a lower CASM, but the costs for the Boeing 737-700 have not increased as quickly as the Airbus 319. This could be a result of the introduction of blended winglets on the majority of Boeing 737-700 aircraft.

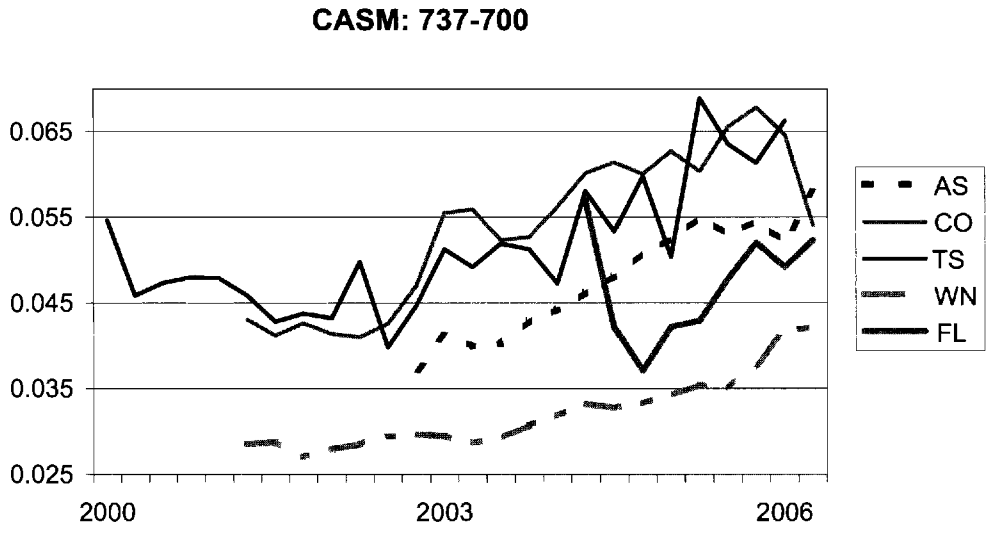

Figure 12.8 Cost per available seat mile for US 737-700 operators, 2000–2006

Source: Compiled by the authors using Back Aviation Form41.

Note: Some data for Southwest, Continental and Alaska was unobtainable.

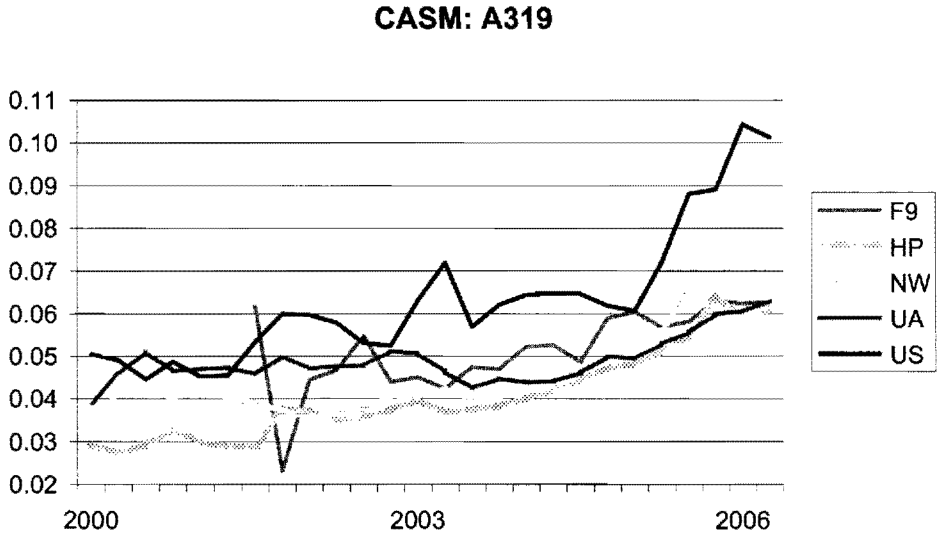

Figure 12.9 Cost per available seat mile for US Airbus 319 operators, 2000–2006

Source: Compiled by the authors using Back Aviation Form41.

The final CASM cost comparison (Figure 12.10) is for the larger A320 aircraft; this shows LCC JetBlue as the cost leader. On average, the airline has operated at a whole cent lower than the competition, and this could easily be the difference between profitability and non-profitability in any given quarter. For this aircraft, US Airways is actually one of the better legacy carriers, while quasi-LCC America West has one of the highest CASMs for the A320 fleet. This highlights the fact that the scheduling and operating procedures for individual carriers can have a significant impact on the CASM for any given aircraft. United's statistics for this figure include numbers for both mainline A320 flights and Ted[5] flights, so it appears (from the figure) that the introduction of Ted has not had a dramatic impact on United's CASM figures.

With the exceptions noted above, the general trend of comparisons by individual aircraft type show that LCCs have been able to achieve lower operating costs. This trend is further reinforced in Figure 12.11. Here the average domestic CASM between LCCs and legacy carriers since 1996 is compared. For this figure, the LCCs include US airlines such as Southwest, JetBlue, America West, and AirTran, while the legacy carriers include American, United, Delta, Continental, US Airways, and Northwest.

As the figure clearly shows, in the late 1990s the cost-structure difference between LCCs and legacy carriers was sizeable. Fortunately for them, the legacy carriers were still able to obtain high revenues during a boom time in the economy. However, following the tragic events of 9/11 and the subsequent dramatic drop in air traffic, the LCCs responded more quickly to the new economic reality—that is, they adjusted their cost structures, while legacy carriers struggled to adjust. Since then the difference in CASMs has remained fairly constant, with LCCs on average having a 20 per cent cost advantage over legacy carriers. Although the airline industry's profitability as a whole is highly correlated to the strength of the economy, this comparison shows that legacy carriers definitely experience greater volatility in their profits and/or losses than LCC. Figure 12.11 also shows that rather small changes in average costs can produce dramatic returns in revenues.

5 Ted is United's low-cost carrier, the name 'Ted' being derived from the last part of United.

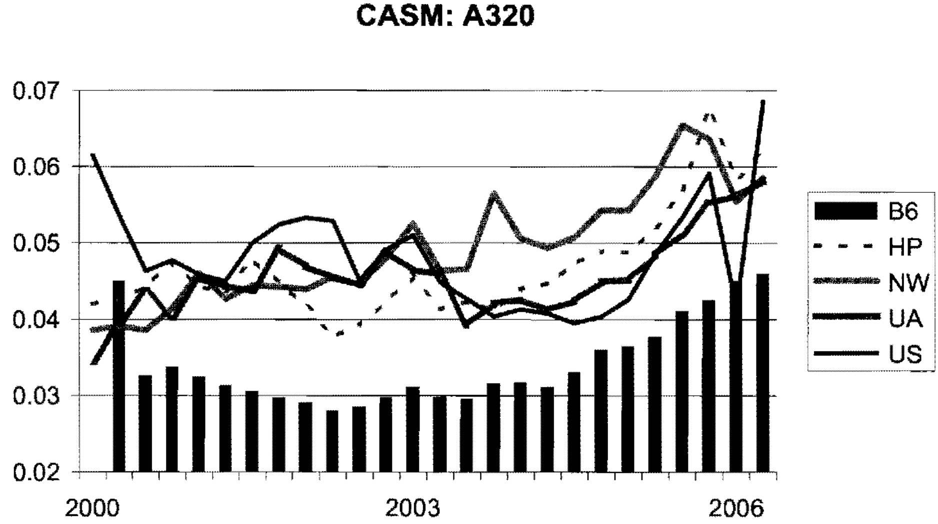

Figure 12.10 Cost per available seat mile for US Airbus 320 operators, 2000–2006

Source: Compiled by the authors using Back Aviation Form41.

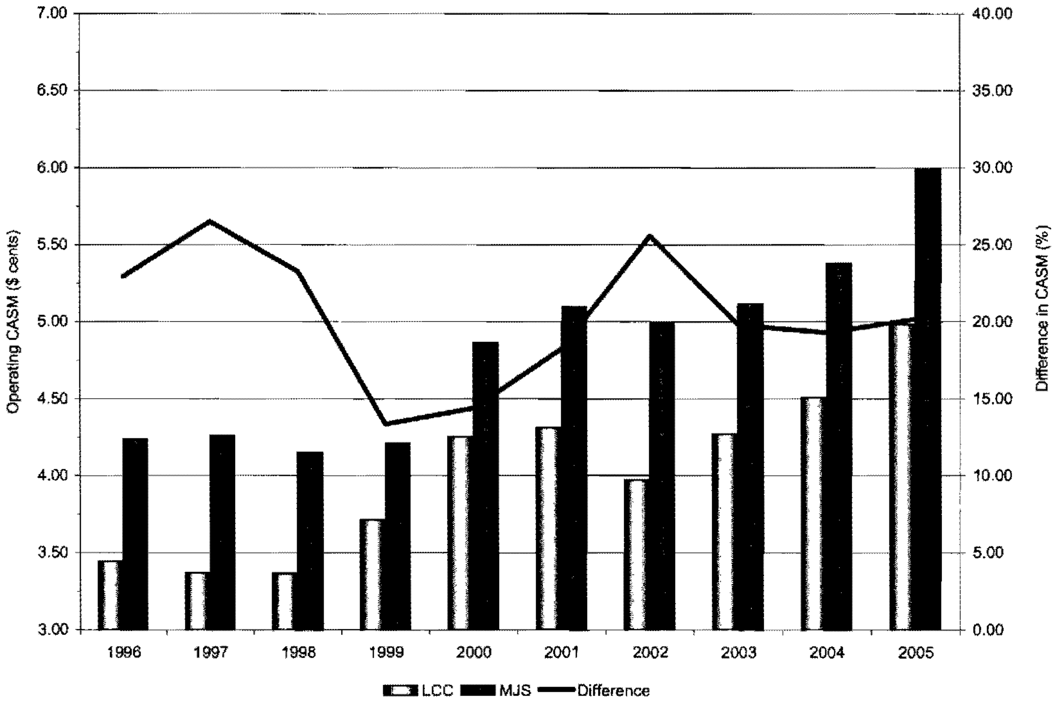

Figure 12.11 Comparison of LCCs' and legacy carriers' operating CASM, 1996–2005

Source: Compiled by the authors using Back Aviation Form41,

Table 12.5 provides a common-size income statement that compares the LCCs to the legacy carriers since 2000.[6] The table shows that LCCs have financially outperformed legacy carriers since 2000. In particular, LCCs have been much more successful in terms of operating profit and profit margin than the legacy carriers, and one of the main reasons for this has been their lower cost structure. Although the LCCs' total aircraft operating expenses consume a larger percentage of total operating expenses, they have had comparably lower direct operating costs.

Table 12.5 also shows that LCCs acquire a greater proportion of their revenues from passengers than from other sources. This is largely because, unlike legacy carriers, many LCCs carry no additional cargo on their flights (which could slow up aircraft turnarounds).

LCCs spend less of their total operating expenses on personnel than do legacy carriers. This could be caused by either of two factors: a quantity or a price effect. Under the quantity effect, legacy carriers would have the same wage levels as LCCs, but simply have (proportionately) more staff for their flights. The price effect would be the opposite—that is, where the proportionate staffing levels are the same, LCCs simply pay their employees less. In all likelihood, the difference in personnel expense proportions is probably caused by a mixture of the two effects.

Finally, LCCs spend a greater proportion of their operating expenses on fuel and maintenance in comparison to legacy carriers. A possible explanation for these higher cost proportions is that Southwest and AirTran have shorter average stage lengths. Since take-off consumes more fuel than any other stage of flight, and most maintenance programs revolve around the number of take-offs, the more departures per aircraft there are, the higher the expected proportion of fuel and maintenance costs. This theory is supported in Figure 12.12 which shows the average number of departures per aircraft in 2006 for major US airlines. Both Southwest and AirTran lead this list, indicating that they could be the driving factors behind the slight rise in the LCCs' percentage of fuel and maintenance costs compared to the legacy carriers.

INCUMBENT CARRIERS' RESPONSE TO LOW-COST CARRIERS

Legacy carriers have implemented two major strategies to combat low-cost carriers: the creation of their own LCCs and unilateral cost-cutting.

LCC Creation

Most major carriers in both North America and Europe have experimented in creating their own LCCs, but with little success. In the early 1990s Continental launched its own LCC, Continental Lite, which was configured with all economy seats, served no food, and operated flights of under 2.5 hours' duration (Bethune, 1998). Continental committed over 100 aircraft to the operation, but ceased operations shortly thereafter. "There was only

6 A common-size income statement essentially displays the income statement merely in percentage terms. All revenue categories are stated as a percentage of total revenue. Conversely, all expense categories are stated as a percentage of total operating expenses. The benefit of a common-size income statement in that it enables a better comparison of cost structures and helps highlight where money is being spent.

Table 12.5 Common-size income statement comparison between low-cost and legacy carriers, 2000–2005

		2000		2001		2002		2003		2004		2005	
		LCC	Legacy	LCC	Legacy	LCC	Legacy	LCC	Legacy	LCC	Legacy	LCC	Legacy
Revenues													
	Passenger Revenues	91.3%	88.4%	92.0%	87.1%	92.7%	88.5%	92.9%	82.2%	91.0%	76.9%	87.6%	75.2%
	Other Revenues	8.7%	11.6%	8.0%	12.9%	7.3%	11.5%	7.1%	17.8%	9.0%	23.1%	12.4%	24.8%
Total Revenues		100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Expenses													
	Personnel Expenses	8.2%	9.7%	8.6%	9.4%	9.5%	10.5%	10.1%	10.5%	9.6%	9.4%	8.9%	8.7%
	Fuel and Oils	18.9%	13.6%	16.8%	13.3%	16.1%	14.7%	17.7%	14.8%	19.8%	13.2%	23.7%	12.3%
	Other Direct Expenses	9.3%	8.4%	10.3%	7.7%	11.0%	5.1%	12.6%	5.6%	11.7%	9.6%	10.0%	14.2%
	Maintenance	13.2%	12.0%	12.6%	11.8%	12.0%	12.0%	12.1%	10.5%	11.0%	9.5%	9.6%	8.9%
	Depreciation and Amortization	5.8%	5.5%	6.7%	6.9%	5.9%	6.3%	5.1%	6.1%	4.9%	5.4%	4.4%	4.7%
Total Aircraft Operating Expenses		55.4%	49.2%	55.0%	49.2%	54.5%	48.6%	57.6%	47.4%	57.1%	47.0%	56.7%	48.8%
	Servicing and Administration Expenses	44.2%	47.0%	44.7%	46.5%	45.3%	46.4%	42.3%	42.7%	41.2%	38.3%	38.4%	34.9%
	Transport related Expenses	0.4%	3.8%	0.3%	4.3%	0.2%	5.0%	0.2%	9.9%	1.7%	14.8%	4.9%	16.3%
Total Operating Expenses		100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
	Operation Margin	10.2%	5.0%	1.3%	-13.5%	2.0%	-14.1%	6.8%	-5.9%	-0.9%	-4.5%	0.2%	-2.8%
	Profit Margin	5.6%	1.8%	1.5%	-10.1%	-2.6%	-16.3%	5.8%	-4.2%	-4.4%	-10.8%	-5.9%	-30.3%

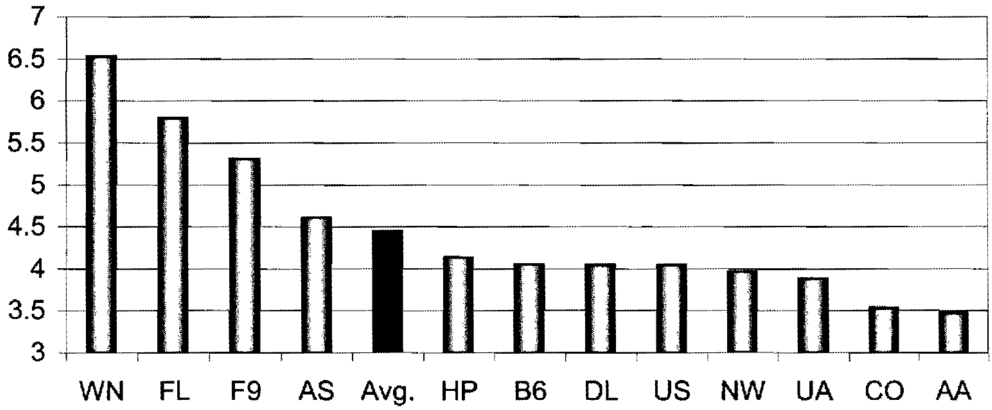

Figure 12.12 Average departures per aircraft per day, 2006

Source: Compiled by the authors using Back Aviation Form41 data.

one problem: People said, 'I don't want to buy that. That is not what I want'" (Bethune, 1998, p. 47). Since Continental Lite did not offer a competitive product, passengers did not choose the airline. Moreover, since Continental's name was closely associated with the new carrier, brand confusion occurred, and Continental mainline also lost passengers (Bethune, 1998). Continental's experiment with Lite was one of the first of a long list of failed attempts by legacy carriers to develop their own "low-cost carriers".

Delta Air Lines has made two attempts at creating a low-cost carrier. Delta Express began operations in 1996 utilizing 737-200s in a high-density layout (O'Toole, 1999). Delta Express was based out of Orlando International and operated flights principally along the north-east corridor. It was created to compete with Southwest, Air Tran and eventually JetBlue (O'Toole, 1999). A 1999 study put Delta Express's CASM at 10.86 cents, which was considerably lower than Delta mainline, but still well above Southwest's CASM of 7.75 cents (O'Toole, 1999). After 9/11, Delta Express's operations were significantly reduced as leisure travel declined sharply (Johnston, 2001). It ceased operations in November 2003, shortly before Delta started its second LCC, Song. Song launched services in April 2003 amid much fanfare. Largely serving leisure routes from Florida and transcontinental flights from the north-east, the carrier attempted to create a hip, style-conscious brand by operating larger 757-200s in an all-economy configuration, with leather seats and an excellent in-flight entertainment system. Delta shut down the Song operation in April 2006, only three years after its start, as the carrier attempted to restructure under bankruptcy protection. Many of Song's characteristics are planned to be transferred into Delta mainline service.

US Airways launched MetroJet in 1998 to respond to low-cost competition from Southwest and Delta Express (Henry, 1998). The airline's base was Baltimore/Washington International (BWI) where Southwest also had a large operation. The airline operated 737-200s in an all-economy configuration, and the operation received labor concessions from the unions (Henry, 1998). MetroJet mainly focused on flying in the north-east and Florida, but faced fierce competition from Southwest and ended up mostly reducing its own mainline passengers. Much of the cost savings achieved by MetroJet were achieved by lower pay rates for employees and economies of density achieved through all-economy seating. However, while MetroJet did have lower costs than mainline US Airways, its cost structure was still high as a direct result of the fact that it was embedded within the mainline carrier. Following the 9/11 terrorist attacks, MetroJet's operations were shut down, and much of US Airways presence at BWI was never restored (Johnston, 2001).

United Airlines is another carrier that has operated two LCCs. Its first LCC, Shuttle by United, was an all-economy service utilizing 737-300/500s out of San Francisco International Airport (SFO) (Flint, 1996). During the late 1990s the Shuttle operation appeared to be quite a success for United, but actual statistics and data for Shuttle were never publicly released (Flint, 1996). Shuttle lasted for several years, and provided United with a focus on operational efficiencies, but like other first-generation legacy LCCs, Shuttle was wound up in 2001, with the aircraft being folded back into mainline service.

In 2004 United Airlines relaunched its low-cost model in the form of Ted (standing for the last three letters in United), which operates all-economy A320s from Denver. The airline operates to leisure markets such as Orlando and Phoenix ,replacing mainline service to such cities. All Ted flights are operated by United Airlines crew, as Ted does not have its own operating certificate. Of all the LCCs spun off by legacy carriers, Ted is the only airline still operating in 2006. It remains to be seen if it will last.

In Canada, Air Canada holds the dubious distinction of operating two LCCs at the same time, neither of which has survived. In 2001 Air Canada launched Tango, utilizing A320s configured in an all-economy layout. Although Tango was operated by Air Canada crews, the airline was totally autonomous from Air Canada mainline flights. This created a problem in that Tango's flights relied solely on O & D demand. Tango was created to respond to LCCs such as Canada 3000. However, shortly after Tango's launch, Canada 3000 fell into bankruptcy.

Air Canada's second discount carrier, called Zip, was launched in 2002 and operated as a totally separate airline with its own operating certificate, labor force, and management, but code-shared on all its flights with Air Canada. The carrier was based in Calgary to compete heavily against Calgary-based LCC WestJet. The airline operated 737-200s in an all-economy layout, but following Air Canada's entry into bankruptcy in 2004, both Zip and Tango disappeared.

In Europe a similar phenomenon occurred, with both British Airways and KLM setting up their own LCCs. In 1997 British Airways launched Go Fly using Boeing 737s based at London Stansted Airport (Goldsmith, 1998). The airline highlighted its ties with British Airways and posted a profit in 2000. When new management took over British Airways, Go became a liability as it was reducing the airline's core business (Goldsmith, 1998). In a move that is not in the business model for a low-cost carrier in the US, easyJet bought Go in May 2002 (Clark, 2002). Go's network was subsequently integrated into easyJet's.

In 2000 KLM launched Buzz to compete with other LCCs such as easyJet, Ryanair, and Go in the UK market (Dunn, 1999). Unlike most LCCs, Buzz operated two small fleets of Bae 146s and 737-300s. Since both fleets were small, no economies of scale could be realized, and operating costs were not "low-cost." Furthermore, from its base in London Stansted, Buzz flew into busy airports such as Amsterdam and Paris Charles de Gaulle (Dunn, 1999). In a deal similar to that done by easyJet, Ryanair bought Buzz in 2003 for £15.1million, but with Buzz having close to £11million cash on hand, the true cost of the purchase was very small (BBC News, 2003). Ryanair operated Buzz as a separate unit for a year, but eventually dissolved the operation and had Ryanair take over all operations. Therefore, KLM's experience with an LCC was short-lived and, like most other legacy carriers, unsuccessful.

In both Europe and North America, legacy carriers' experiments in creating their own LCCs have largely failed. Part of the problem lies in the operation never truly being low-cost, especially with regard to labor costs. The legacy carriers have also been very concerned about the new operator reducing their own core business. As the discussion above clearly shows, the legacy carrier strategy of creating LCCs of their own has been a complete failure (at least to date). Latin America is another growing market for low-cost model airlines. Since 2005 six low-cost airlines have started services in Mexico. In Brazil, GOL and other LCCs now account for more than 40 per cent of the domestic market (Boeing, 2007). And in the Middle East, Air Arabia and Jazeera Airways have already met with success.

Cost-Cutting

The other major response by legacy carriers has been unilateral cost-cutting of mainline services. In the United States, on-board food service has been reduced to the point where almost no food is served in economy class on any domestic flight. Legacy carriers have also begun to charge for such amenities as pillows, blankets, and in-flight entertainment.

While these measures enable legacy airlines to reduce costs, they also introduce the problem of lack of product differentiation. When legacy carriers reduce their service product to equal that of LCCs, they are largely competing just on cost. And, as we have seen above, competing solely on cost is risky, since LCCs have much lower cost structures than legacy carriers.

Legacy carriers have also attempted to reduce their cost structures by retiring older aircraft, receiving labor concessions, and reconfiguring aircraft seating layouts. In general, however, the legacy carriers' response to LCCs has largely been ineffective, and many carriers have attempted to avoid low-cost carriers by focusing on international flying (where they have definite competitive advantages due to legal restrictions). The future remains uncertain and is discussed in the following section.

THE FUTURE OF LOW-COST CARRIERS

While there have been many failures in the low-cost sector, there has also been tremendous success. And, it must be pointed out, many of the failures have been caused by the new airlines failing to retain the major characteristics of LCCs, so that they were not truly cost leaders. The message seems clear: the LCCs that were successful focused solely on reducing costs and being efficient.

The future of LCCs looks promising. On the one hand, the line between an LCC and a legacy carrier is blurring. LCCs such as JetBlue are providing free live television entertainment, while a legacy carrier may have no audio-visual entertainment. Some LCCs have adopted leather seats and increased legroom to provide additional amenities for their passengers. Also, airlines such Southwest have loyalty programs just like the legacy carriers. From a passenger standpoint, the differences are becoming very hard to distinguish.

On the other hand, LCCs will face tremendous competition from legacy carriers and LCCs alike. Since much of the LCCs' strategy is based on growth, the LCCs need to develop new markets to continue growing. In addition, with carriers such as Ryanair and easyJet having over 100 aircraft on order, they may be hard-pressed to find routes to fill their planes. It has been calculated that each of these aircraft will have to carry 250,000 passengers per year to break even (Turbulent Skies, 2004).

The last domain for legacy carriers has been international flying. The legacy carriers have dumped capacity into international markets, since this is where profits are being made. It is also the only area where legacy carriers do not face fierce competition from LCCs. This is principally a result of international air treaty regulations and the fact that international flying diverges from the low-cost model (since it generally requires different and larger aircraft types). While it is unlikely that Southwest will begin to fly transatlantic, there will probably be a few airlines in the future that will attempt to bring the low-cost model to the international market. The prospect of such a model succeeding is unknown (although the example of People Express shows that it is possible), but there is no doubt that the legacy carriers will protect their turf fiercely. If a carrier can achieve success using the low-cost model on international flights, then the world is truly the limit for low-cost air transportation. As this chapter's opening vignette stated, the twenty-first century could belong to the low-cost carrier.

SUMMARY

This chapter was devoted to the new paradigm of low-cost start-up airlines. It outlined the emergence of this phenomenon and followed this with a detailed list of the characteristics of low-cost carriers. Cost comparisons were made between the low-cost and legacy carriers, and the incumbent airlines' response to the low-cost carriers was discussed. Finally, the future of low-cost carriers was explored.

APPENDIX: SELECT AIRLINE TWO-LETTER CODES

Airline	Code
Air Canada	AC
Airtran Airways	FL
Air Transat A.T.	TS
Alaska Airlines	AS
American Airlines	AA
America West Airlines	HP
Continental Airlines	CO
Delta Airlines	DL
Easyjet	U2
Frontier Airlines	F9
JetBlue Airways	B6
Northwest Airlines	NW
Ryanair	FR
Southwest Airlines	WN
United Airlines	UA
US Airways	US
Westjet	WS

REFERENCES

BBC News. (2003). *Q & A: Ryanair swoops on Buzz.* Retrieved on 19 September 2006 from: http://news.bbc.co.uk.

Bethune, G. with Huler, S. (1998). *From Worst to First: Behind the Scenes of Continental's Remarkable Comeback.* New York: John Wiley & Sons.

Boeing (2007). *Current Market Outlook.* Available at: http://www.boeing.com/commercial/cmo/.

Buyck, C. (2005). Wooing Europe's New Breed. *Air Transport World,* 42(9), pp. 32–35.

Calder, S. (2002). *No Frills: The Truth behind the Low-cost Revolution in the Skies.* London: Virgin Books.

Clark, A. (2002). Easyjet Lines up Merger with Go: Shake-up of Budget Airlines Could Mean Higher Fares. *The Guardian,* 4 May, p. 2.

Doganis, R. (2001). *The Airline Business in the 21st Century.* London: Routledge.

Dunn, G. (1999). KLM Launches Low Cost Airline—named "Buzz." *Air Transport Intelligence News,* 22 September.

Field, D. and Pilling, M. (2005). The Last Legacy. *Airline Business,* 21(3), pp. 48–51.

Flint, P. (1996). The Leopard Changes its Spots. *Air Transport World,* 33(11), pp. 51, 54.

Goldsmith, C. (1998). British Airways Launches No-Frills Unit—Move May Risk Diluting Brand Name, Some Say. *Wall Street Journal,* 22 May, p. 5.

Henry, K. (1998). Aiming High with Lower Fares: US Airways' MetroJet Set to Debut, Battle Southwest, Boost BWI. *Sun,* 31 May, p.1.

Ionides, N. and O'Toole, K. (2005). Points of Sale. *Airline Business,* 21(3), pp. 42–45.

Johnston, C. (2001). Airlines Are Cutting Their Discount Services. *New York Times,* 18 November, p. 5.

O'Toole, K. (1999). Express Yourself. *Airline Business,* 28 January Retrieved on 28 November 2006 from Air Transport Intelligence at: http://www.rati.com/iata/oapd.

Smith, S. (2005). The Strategies and Effects of Low-Cost Carriers. *Steer Davies Gleave.* Retrieved on 20 September 2006 from: http://www.icea.co.uk/archive.htm.

Turbulent Skies (2004). *The Economist,* 10 July, pp. 59–63.

13

The Economics of Aviation Safety and Security

> Regulation has gone astray ... Either because they have become captives of regulated industries or captains of outmoded administrative agencies, regulators all too often encourage or approve unreasonably high prices, inadequate service, and anticompetitive behavior. The cost of this regulation is always passed on to the consumer. And that cost is astronomical.
>
> Senator Edward Kennedy, opening remarks to the Subcommittee on Administrative Practice and Procedure, 6 February 1975

In common usage, the term "safety" is often used incorrectly as an absolute value—that is, one is either safe or unsafe. However, safety is never absolute, since there is always some probability of an accident. Thus, safety depends on the given situation and the risks that are part of that situation. Nevertheless, many people claim that safety should be maximized, regardless of the cost. The reasoning goes something like this: if human life is deemed sacred, then it may seem reasonable to consider human safety to be sacred—shouldn't our goal be to achieve as much safety as possible? But, if this is true, then we should outlaw any activity where fatalities are even remotely possible. We need to ban swimming, skiing, fishing, flying, pregnancy, social gatherings (where disease may spread), driving at more than 15 miles per hour, and virtually everything else that people enjoy doing!

As the examples above point out, increases in safety are optimal only when the benefits of safety justify the costs; thus, minor increases in safety that impose major costs are never cost-efficient. This is the reason why people do not wear helmets all the time, or the country does not establish a national speed limit of 15 mph, although clearly both of these actions would reduce the risk of accidents. So, despite the fact that expressions like "safety must be preserved at any price" are commonly used, safety still needs to be judged within the economic context of a simple cost–benefit analysis.

The benefits of safety are undeniable—not only from a moral standpoint, but also from an economic standpoint. Some of the potential economic benefits of aviation safety include: strengthened consumer demand, strengthened labor supply, reduced insurance costs, lower cost of capital, lower liability risk, and reduced costs associated with government fines or penalties. While this chapter will generally analyze aviation safety from an economic standpoint, it will also provide some more specific facts and figures on aviation safety. The general outline for this chapter is as follows:

- The history of aviation safety

- Incentives for aviation safety, including:
  - Passengers' reaction
  - Labor reaction
  - Financial concerns
  - Insurance costs and liability risks
  - Government enforcement
- Causes of aviation accidents, including:
  - Flight crew error
  - Aircraft malfunction
  - Weather conditions
  - Airport/air traffic control
  - Maintenance
  - Miscellaneous/other
- Classification of accidents by phase of flight
- Classification of accidents by regions
- The basic economics of safety
- Is it possible to take the politics out of safety regulation?
- Safety prevention

Prior to discussing the topic of aviation safety, it is useful to understand some basic aviation safety terminology, since a few terms are sometimes used synonymously, but may contain different meanings. The National Transportation Safety Board (NTSB) defines an aviation accident as:

> ... an occurrence associated with the operation of an aircraft which takes place between the time any person boards the aircraft with the intention of flight until all such persons have disembarked, and in which any person suffers death or serious injury as a result of being in or upon the aircraft or by direct contact with the aircraft or anything attached thereto, or in which the aircraft receives substantial damage.
>
> (Vasigh and Helmkay, 2002, p. 502)

On the other hand, the NTSB defines an aviation incident as "an occurrence other than an accident, associated with the operation of an aircraft, which affects or could affect the safety of operations" (Vasigh and Helmkay, 2002, p. 502). In practice, the difference between an accident and an incident is largely based on the severity of the situation. If damage occurs to an aircraft, then the situation would probably be deemed an accident. However, an aircraft landing on a parallel taxiway or a runway incursion may incur damage (and therefore be classified as an incident), but may in fact represent a more serious threat to safety.

A fatal aviation accident is one which produces fatalities, or deaths. The fatalities could involve passengers, crew members, or people on the ground. A hull loss occurs when an aircraft is a complete write-off from an accident and is no longer flown. Typically, aviation safety is measured in terms of accidents, fatal accidents, fatalities, and hull losses.

Other terms that may appear in aviation safety literature include a near mid-air collision (NMAC), a pilot deviation and a runway incursion. Whereas a mid-air collision involves two aircraft making contact while in flight, a near mid-air collision is an incident associated with an aircraft flying within 500 feet of another airborne aircraft (Vasigh and Helmkay, 2002).

Pilot deviation refers to the actions of a pilot that result in the violation of a Federal Regulation or a North American Aerospace Defense Command (NORAD) Air Defense Identification Zone (ADIZ) Directive (Vasigh and Helmkay, 2002). Pilot deviation simply means that the aircraft enters airspace that is totally restricted or the aircraft enters such an airspace without taking the appropriate procedural steps; these deviations may be a result of equipment malfunctions, weather conditions, operational factors, and/or pilot experience (Vasigh and Helmkay, 2002.

Finally, a runway incursion refers to any occurrence on an airport runway involving an aircraft and any object or person on the ground that creates a collision hazard, or results in a loss of separation with an aircraft taking off, intending to take off, landing, or intending to land (FAA, 2002). Although aviation accidents have been diminishing over the past few years, runway incursions have not. Runway incursions are further classified into three operational categories and four severity classifications.

THE HISTORY OF AVIATION SAFETY

> If you were born on an airliner in the US in this decade and never got off you would encounter your first fatal accident when you were 2300 years of age and you would still have a 29 per cent chance of being one of the survivors.
>
> Les Lautman, Safety Manager, Boeing Commercial Airplane Company, 1989

In 1903 the Wright brothers made the first-ever heavier-than-air flight. Today, more than a century later, tens of thousands of airplanes are in the air at any one time, with those aircraft spanning all shapes and sizes. While aviation technology has developed at a tremendous rate, so too has aviation safety. Once a highly risky method of transportation, aviation has developed into the safest mode of transportation available to the public. In fact, in terms of fatalities per passenger miles in the United States from 1995 to 2004, air transportation was 40 times safer than passenger automobiles. Table 13.1 displays the fatality rate for various modes of transportation.

As Table 13.1 shows, air transportation is an incredibly safe mode of transportation, with 1998 and 2002 actually producing zero fatalities. These statistics support the fact that one is more likely to be involved in a fatal accident while driving than while flying. Although people may complain about the service provided by airlines, the one thing consumers have little to complain about is safety. Over the years airlines have proven that they successfully meet their foremost objective—safety.

The increase in aviation safety has been most dramatic over the past few decades. The Second World War was a significant event in aviation history as aircraft technology advanced rapidly to cope with wartime demand. Following the war, commercial aviation soared as the new technologies and aircraft developed during the war were transferred to commercial applications. An excellent example of this technology transfer was the German Messerschmitt 262, which was the world's first operational jet-powered fighter aircraft. This aircraft's technology was subsequently used as a basis for future jet aircraft.

Table 13.1 US fatality rates for various modes of transportation

US Fatalities per 100 Million Passenger Miles				
Year	Auto	Bus	Train	Airline
1989	1.12	0.04	0.06	0.09
1990	0.99	0.04	0.02	0.003
1991	0.91	0.04	0.06	0.03
1992	0.83	0.04	0.02	0.01
1993	0.86	0.02	0.45	0.01
1994	0.91	0.03	0.04	0.06
1995	0.97	0.03	0.00	0.04
1996	0.96	0.02	0.09	0.08
1997	0.92	0.01	0.05	0.01
1998	0.86	0.05	0.03	0.00
1999	0.83	0.07	0.10	0.003
2000	0.80	0.01	0.03	0.02
2001	0.78	0.02	0.02	0.06
2002	0.77	0.06	0.05	0.00
2003	0.74	0.05	0.02	0.005
2004	0.71	0.05	0.02	0.002
1989–2004	0.87	0.04	0.07	0.03
1995–2004	0.83	0.04	0.04	0.02

Source: Compiled by the authors using Air Transport Association (ATA) data.

And, although it took some time for it to power the majority of commercial aircraft, the development of the jet engine marked a significant event in commercial aviation safety. By allowing aircraft to fly further, faster, and higher, the jet engine proved to be more reliable than the piston engine, thereby increasing safety. This increase in safety is illustrated in Figure 13.1, which displays the accident rates for commercial aircraft in the United States. The evolution of the jet engine through the 1960s resulted in an accident rate that was decreasing exponentially, signaling a dramatic change in aviation safety.

Other new technologies, such as fly-by wire, have also made commercial aviation safer. Advanced computer simulation and modeling have made it easier to design redundant and fail-safe components on aircraft. While all of these factors have played a role in improving safety, they have also been accompanied by improvements in pilot training, specifically with advanced simulation training and research into crew resource management, and this has reduced the primary cause of aviation accidents—pilot error. Because of these advances and others, aviation has continually become safer and safer. However, in the

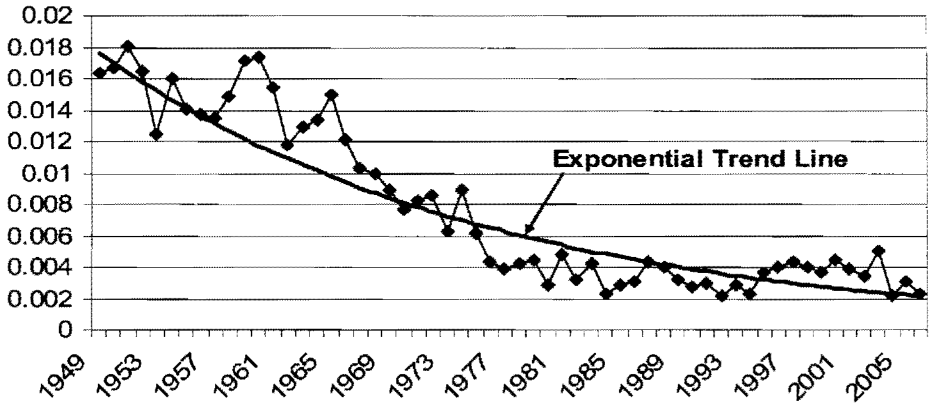

Figure 13.1 Commercial aircraft accidents per million departures in the United States

Source: Compiled by the authors using Air Transport Association (ATA) data.

past 20 years aviation safety has appeared to plateau: in general, the accident rate has hovered between 0.002 and 0.004 accidents per million aircraft departures since the mid-1980s. This is due in part to fewer advances in aviation technology, but it also may well be that aviation safety has reached an economic equilibrium. This implies that accidents could conceivably be further reduced, but the costs of doing so may be excessively high.[1] Therefore, aviation safety may be reaching a point where the benefits of safety are approximately equal to the costs (the cost–benefit analysis of commercial aviation will be discussed in more detail later in this chapter). The plateau effect discussed above is displayed in Figure 13.2.

As the figure shows (in comparison with Figure 13.1), fatalities are more random than accidents, with significant fatality years followed by years with no fatalities. Because of this variability in the data, a five-year moving average was constructed and is portrayed in the figure in order to more properly identify the trend in the data. The five-year moving average shows a significant downward trend in aviation fatalities, mirroring the downward exponential trend in aviation accidents. Today, the number of fatalities in commercial aviation is incredibly minute. In fact, based on 2006 data, it would take a passenger flying every single day, 43,720 years until they would experience a fatality in an aviation accident.[2] The probability of being killed in an aviation accident is less than a billionth of 1 per cent. In other words, the probability of being killed in an aviation accident is practically zero. In fact, many airlines such as JetBlue, Southwest Airlines, Virgin Atlantic, and Emirates have never experienced a fatal accident, and many defunct airlines like Go, Laker Airways, Song and MetroJet went through their whole operations without a fatal accident.

---

1 A good analogy here might be automobile safety. It would be possible to reduce automobile fatalities if every abutment on the interstate highway system was surrounded by crash-absorbing material. However, it is clear that this would be too costly for the few fatalities that it might prevent.

2 In a similar study, Barnett (2001) examined the mortality risk of air travel and found it to be extremely small. He estimated a death risk per flight of 1 in 13 million operated by countries that have a well-developed aviation industry. At this level of mortality risk, a passenger would have to take one flight per day for 36,000 years before having a fatal plane crash.

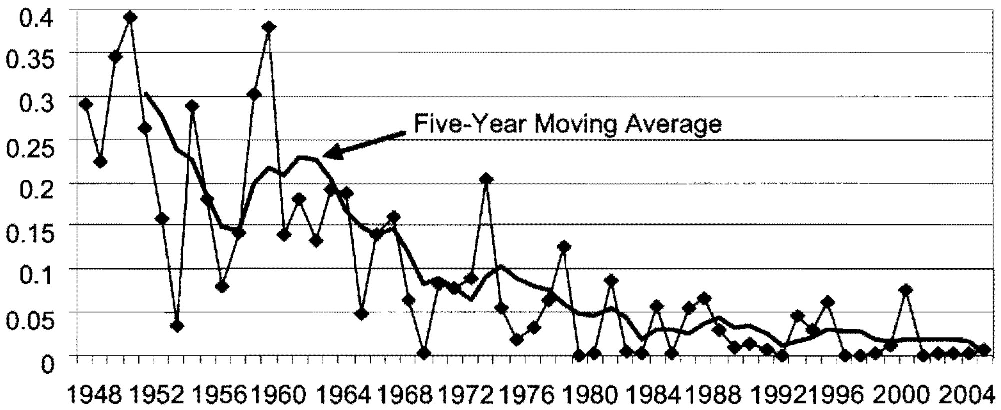

Figure 13.2 Commercial aviation fatalities per million aircraft miles in the United States with a five-year moving average

Source: Compiled by the authors using Air Transport Association (ATA) data.

INCENTIVES FOR AVIATION SAFETY

While some observers (usually in sensationalized media stories) will assert that airlines occasionally cut corners that compromise safety in the interests of greater profit, we will argue in the paragraphs below that just the opposite is probably true. There are strong incentives for airlines to avoid any accidents or incidents, so they are not likely to deliberately cut corners to compromise safety.

As the above data show, aviation is a tremendously safe mode of transportation. However, why do some people still believe that aviation is still unsafe? And why is the aviation industry continually focused on safety to the exclusion of many other considerations? The answers to these questions lie mainly (although there are other reasons covered below) in what might be termed asymmetrical media coverage of aviation accidents and incidents. While many more people in total are involved in automobile accidents and fatalities, aviation accidents typically involve more people in a single accident. Therefore, rightly or wrongly, the media sensationalize every aviation accident or incident. Such extensive media coverage creates a situation in which, if an airline has an accident, their logo may be emblazoned into the minds of consumers across the world for all the wrong reasons. While the media does not document the thousands of routine safe flights a day and the incredibly low probability of an aviation accident, any minor safety slip by an airline will draw extensive media coverage; this creates a climate where airlines have a very strong incentive to insure aviation safety.

However, media coverage is just one of many incentives that airlines have to provide safe and secure air travel. The other incentives for aviation safety can be grouped into five broad categories:

- passengers' reaction
- labor reaction
- financial concerns
- insurance cost and liability risks
- government enforcement.

Passengers' Reaction

Since consumers can choose amongst firms in the market, they can decide how important safety is to them. Therefore, an airline that is perceived to be less safe, as a result of an accident or investigation, is likely to see a decrease in demand compared to "safer" airlines. For passengers, the perception of the level of safety is ordinarily the key decision factor; this is opposed to the actual level of safety, since passengers usually know very little about the actual level of safety of their flight (Squalli and Saad, 2006). Perceptions of aviation safety are largely a result of accidents, particularly accidents that receive extensive media attention. In addition, perceptions are difficult to change and may persist for an extended period of time; this creates a situation where a single aviation accident may have a long-term impact on an airline's demand (Squalli and Saad, 2006).

While logical economic reasoning implies that demand for a particular airline should be reduced when it has had an accident, empirical studies have had difficulty in proving a decrease in demand. Borenstein and Zimmerman (1988) conducted an extensive study of 74 accidents in the United States from 1960 to 1985 and found that there was no statistically significant decrease in demand for the airline's services as a result of the accident.[3] Squalli and Saad (2006) did find a minor decrease in demand resulting from aviation accidents in the United States, while Wong and Yeh (2003) estimated a 22.11 per cent decline in monthly traffic lasting for 2.54 months resulting from an aviation accident in Taiwan. Part of the difficulty in statistically showing a decrease in demand is that, following an aviation accident, airlines are likely to take competitive action to help offset a shift in the demand curve, such as lowering ticket prices. While not statistically significant, Borenstein and Zimmerman (1988) did find that consumers responded more adversely to aviation accidents post-deregulation than they did during regulation.

Although it has been difficult to measure the decline in demand resulting from aviation accidents, economic logic suggests that, although the decline may not be large, it definitely exists. This is based on the simple fact that an accident on one carrier will undoubtedly cause some people (who are perhaps very risk-averse) to fly on another carrier. This may be particularly true of consumers who view aviation as unsafe in certain regions of the world, because there is a public perception that these regions have lower safety standards. While this may or may not be true, these so-called "unsafe" airlines will have difficulty in changing the public perception, since consumers typically remember bad things before good things.

The fact that the demand for an airline's flights could be severely affected provides an incentive for the airline to prevent accidents. Furthermore, an airline that is perceived to be safer than the competition may receive a modest increase in demand, if the consumers view safety as an important factor in their decision-making.

Labor Reaction

Similar in vein to passengers' reaction, the labor supply may also react adversely to an aviation accident. Employees do not want to work in an environment which they believe is unsafe. Therefore, an airline may experience two labor issues as a result of an accident: increased turnover and/or increased wage demands.

---

3 Borenstein and Zimmerman (1988) had results that displayed a 4.3 per cent reduction in consumer demand resulting from an accident during regulation and a 15.3 per cent reduction in demand post-deregulation. However, neither value was statistically significant.

As a result of the perception or the reality of reduced safety, employees, particularly flying crew, may feel that the airline is engaging in questionable safety practices. This could cause employees to leave the company or have the union enforce new safety measures. Moreover, an accident, or series of incidents, could make it more difficult for an airline to attract high-quality employees.

The other outcome affecting the labor supply resulting from decreased safety would be increased wages. Employees may demand better compensation for having to work in a less safe environment. In essence, this would be a form of "combat pay," where employees are compensated for working in an uncertain environment. While such demands may be difficult for employee groups to obtain in the short term, an airline's safety could become bargaining issues in the long term.

These impacts represent real costs to the airline, and therefore provide strong incentives to avoid accidents and incidents. Hence, the airline has an economic incentive to continue operating safely in order to avoid the costs imposed by the market forces resulting from an accident.

Financial Concerns

The stock market will always react negatively toward an airline that experiences an accident, particularly one that involves fatalities. The reason for this is the more or less obvious belief that such an accident will cause great uncertainty over the future of the airline in question. Borenstein and Zimmerman (1988) found that, on average, aviation accidents caused a 0.94 per cent equity loss for the firm on the first day of trading, which was statistically significant to the 1 per cent level. This value is slightly lower than found by two other studies that determined equity losses amounting to 1.18 per cent and 1.19 per cent on the first day of trading (Borenstein and Zimmerman, 1988). Mitchell and Maloney (1989) went one step further in analyzing the impact of an aviation accident on the firm's equity value in the long term. They found that, if the accident was proved to be the airline's fault, then the equity value dropped by 2.2 per cent. However, if the accident was not deemed to be the airline's fault, equity value dropped by only 1.2 per cent. Regardless of who is to blame for the accident, the airline's equity will decline, which is another incentive to avoid aviation accidents and promote safety.

A loss of equity value for an airline will have other, and possibly greater, negative financial effects—namely, a large increase in the cost of capital. Because of the greater risk and uncertainty associated with the decline in equity value, the, airline will find it more expensive to raise capital, and this can be quite serious since airlines are highly capital-intensive. While a loss of equity through a decline in the stock price is not an explicit cost against the airline, an increase in the cost of capital is a direct cost to the airline, providing yet another major incentive to avoid accidents and incidents.

Insurance Costs and Liability Risk

When an aviation accident occurs, airlines are usually fully indemnified from the losses through insurance. Insurance companies will pay out various liability and damage claims for the airlines, so that airlines suffer little direct financial loss from an accident. However, as a result of an airline accident, particularly if the airline is determined to be liable to any extent, the airline's insurance premiums are likely to increase dramatically in the future. Like automobile insurance premiums, airlines will see insurance hikes if they experience an accident or incident. Moreover, the insurance rate hike does not occur for just one year, but lasts for several years. Such increases can have a significant effect on an airline's profit margins. This is true since not only do airlines currently pay substantial insurance premiums, but it has also been estimated that increases in insurance rates explain about 34 per cent of equity loss (Wong and Yeh, 2003). Related to insurance premiums, airlines will find that their liability risks will increase substantially as a result of an accident. Therefore, the threat of increased liability risks provides airlines with one more economic incentive (from increased insurance premiums) to promote safety and avoid aviation accidents.

Government Enforcement

The final major incentive—the threat of government penalties—for aviation safety is not a true market incentive, but is nevertheless a very real one. Like traffic laws, where the threat of a speeding ticket helps deter many from speeding, aviation regulations are designed to deter airlines from violating safety procedures. However, unlike traffic penalties, safety fines levied by the FAA can be substantial. For example, in 2000 the FAA levied fines totaling $988,500 against Alaska Airlines for maintenance violations (FAA, 2000). Another example would be the $805,000 fine against United Airlines in 2002 for improper maintenance techniques (FAA, 2002). Although, clearly, fines levied by the FAA are substantial and provide an incentive for safety, the greatest threat posed to an airline would be that of a complete shutdown due to a severe violation in safety practices. The FAA has the authority to order an airline to cease operations: in such an event not only would the airline's revenue be effectively cut off, but it would also have to pay sizeable costs and penalties. The fear of shutdown is one of the greatest threats to an airline and, while the FAA has rarely used its authority to temporarily shut down an airline, the mere existence of this threat provides a tremendous incentive to adhere to FAA safety practices. Therefore, the presence of the FAA, or other aviation regulators, provide airlines with a final very real financial incentive to promote safe air travel. For example, ValuJet, Kiwi Airlines and Nation's Air all were shut down by the FAA for safety violations.

CAUSES OF AVIATION ACCIDENTS

> If you are looking for perfect safety, you will do well to sit on a fence and watch the birds; but if you really wish to learn, you must mount a machine and become acquainted with its tricks by actual trial.
>
> Wilbur Wright, from an address to the Western Society of Engineers in Chicago, 18 September 1901

Aviation accidents occur for a variety of reasons, and every accident is thoroughly investigated to help prevent future accidents. Accidents are rarely attributed to just one cause, as a variety of factors must go wrong for the accident to occur. Understanding the nature of the accidents and how they occur is important for the continual improvement of aviation safety and for helping to understand the economic principles of safety. Figure

13.3 displays the six major categories by which airline accidents are categorized, together with the percentages of each category:

- flight crew error
- aircraft malfunction
- weather conditions
- airport/air traffic control
- maintenance
- miscellaneous/other.

Flight Crew Error

As Figure 13.3 shows, flight crew error, or human error, is the number one cause of aviation accidents worldwide. While the period used in Figure 13.3 is only from 1996 to 2005, human error has always been the primary cause of aviation accidents. New technology has helped make aviation safer, but it still cannot compensate for errors made by humans. Much research has been conducted into reasons why flight crews make errors, and although areas such as crew resource management have helped reduce human error, the fact is that as long as humans are in control of the aircraft, flight crew error will probably occur.

Although aviation accidents are commonly a result of several contributory factors, most could have been avoided if the crew had done something differently. For example,

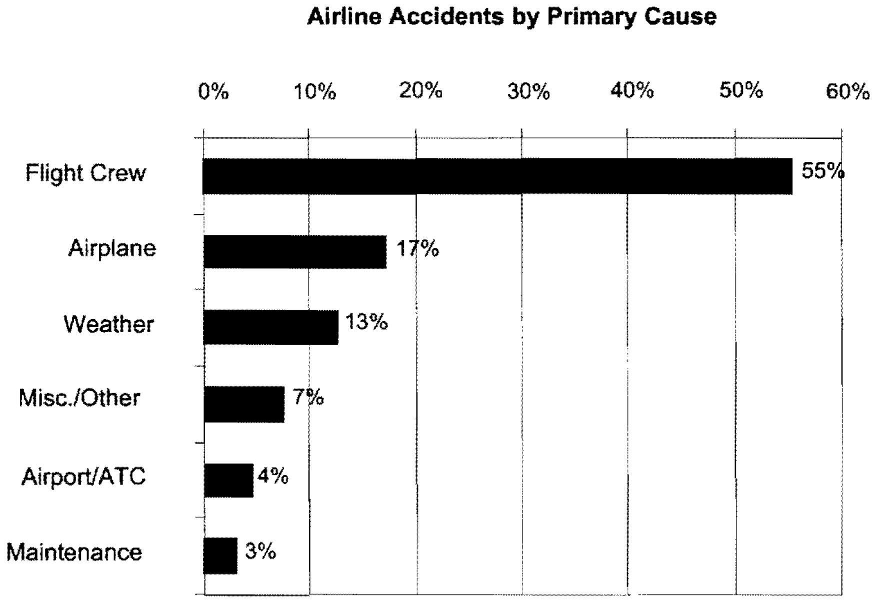

Figure 13.3 Commercial aviation fatalities per million aircraft miles in the United States, with a five-year moving average, 1996–2005

Source: Compiled by the authors using Air Transport Association (ATA) data.

one of the worst aviation accidents in history was a result of human error. The 1977 PanAm/KLM accident in Tenerife was a result of the KLM pilot entering the runway prior to receiving ATC clearance. The subsequent collision with a PanAm 747 killed 583 passengers in total. That said, many other fatal aviation accidents have been avoided due to exemplary efforts by the flight crew. For example, an Aloha Airlines 737-200 was able to land safely (with only one fatality) after part of the fuselage was torn apart by a sudden decompression. Therefore, while human errors have caused accidents, flight crews have also saved numerous lives.

The development of realistic flight simulators has made it possible for pilots to experience a variety of problems without ever taking to the sky. Thus, while technology has successfully made aviation safer, future effort towards better crew training and management may result in fewer flight crew errors.

Aircraft Malfunction

The second major determinant of an aviation accident is an aircraft-related malfunction, with 17 per cent of aviation accidents since 1996 attributed to this cause. Current aircraft are sophisticatedly designed with safety in mind; however, systems may still malfunction, and this can ultimately cause a serious aviation accident. Parts ranging from multi-million-dollar engines to trivial items have all been the cause of serious aviation accidents. The Aloha Airlines flight highlighted above is an example of an aircraft-related accident. In this case, metal fatigue caused part of the fuselage to deteriorate, and the aircraft experienced rapid decompression. Another example of an aircraft malfunction accident is United Airlines flight 232. Here, the aircraft crash-landed in Sioux City, Iowa, after one of the engines failed and thereby disabled the hydraulic systems on the aircraft. Most fatal aircraft malfunction accidents usually occur when the engines experience a problem, although other systems can also cause fatal accidents.

Figure 13.4 displays a ratio of aircraft hull losses for the major aircraft manufacturers between 1959 and 2005. While the number of aviation accidents per manufacturer includes accidents for a variety of reasons, a trend could possibly be extrapolated if it was assumed that the probability of an accident due to pilot error, weather, maintenance, or ATC is approximately the same for all manufacturers. (This may not be a good assumption for the smaller companies that do not have their aircraft spread throughout the world.) In terms of accident rates, Fokker was the highest with slightly over five accidents occurring per million departures. The next two highest were Boeing and McDonnell-Douglas, but this may be a result of the fact that Airbus is a newer company with fewer older aircraft.

To correct for this possibility, Figure 13.5 provides a similar comparison, but with Boeing 707s and DC-8s excluded. In this scenario, the accident rates between Boeing, Airbus, and McDonnell-Douglas are very similar. McDonnell-Douglas is slightly ahead due to a somewhat higher accident rate with the DC10/MD11 aircraft. In fact, when comparing Boeing's and Airbus's accident rate, there is no statistically significant difference, indicating that aircraft manufactured by both companies are the same in terms of safety (Vasigh and Helmkay, 2000). It seems reasonable to assume, therefore, that both Boeing (which now includes McDonnell-Douglas) and Airbus, the two largest and most successful manufacturers, have been producing safe and reliable aircraft.

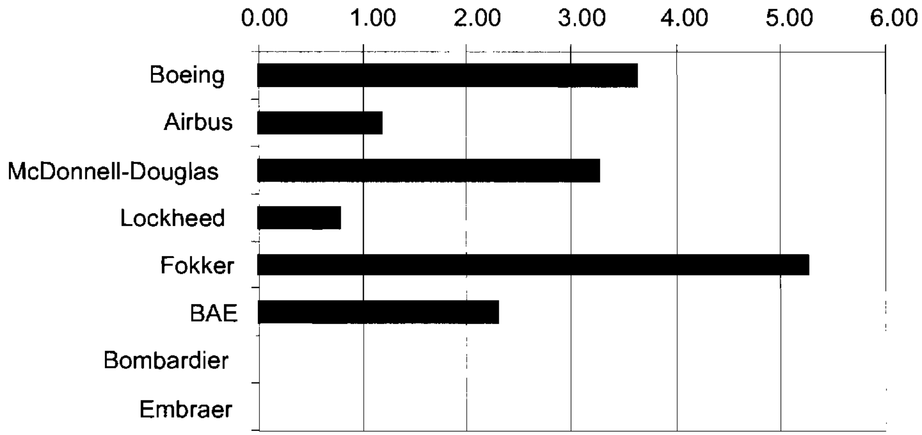

Figure 13.4 Commercial aircraft hull loss by manufacturer per million departures, 1959–2005

Source: Compiled by the authors using Boeing (2006) data.

Note: The sample includes all commercial aircraft during the time period, except for regional and commuter aircraft; therefore, Bombardier's CRJ-200 aircraft and Embraer's ERH aircraft are excluded.

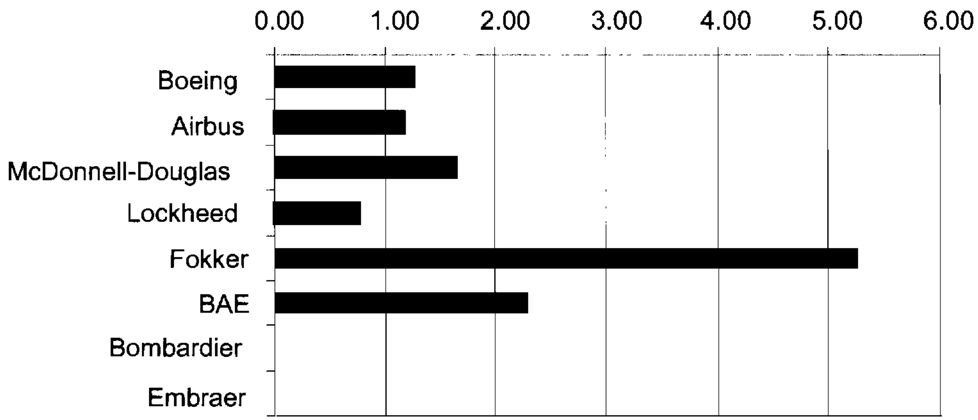

Figure 13.5 Commercial aircraft hull loss by manufacturer per million departures, 1959–2005, excluding Boeing 707s and DC-8s

Source: Compiled by the authors using Boeing (2006) data.

Weather Conditions

The third major cause of aviation accidents is the weather. Through the development of the jet engine, weather-related accidents have become less of a concern, since jet engines enable aircraft to fly higher and avoid any troublesome weather. In addition, the development of instrument landing systems allows aircraft to auto-land in adverse weather, reducing the chance of error during landing in bad weather conditions.

However, weather-related accidents still occur. For example, in 2005 an Air France A340 overran the runway at Toronto Pearson during adverse weather conditions—that is, poor visibility and strong winds. Icing can also be a major problem for commercial aircraft, as evidenced by the American Eagle ATR-72 that crashed in 1994 while waiting to land at Chicago O'Hare. While the primary cause of the accident was icing on the wing, it was also determined that the aircraft type had poor de-icing equipment. As a result of the accident, modifications were made to the aircraft to reduce the risk of another accident of this type.

Airport/Air Traffic Control

Air traffic control can also be prone to human error, resulting in accidents. Air traffic controller fatigue and stress, as well as poor communication, are important human factors that contribute to accidents. Also, the US air traffic control system generally dates from the mid-1960s and has had great difficulty in managing the recent huge increases in aviation traffic.

One of the more recent aviation accidents attributed to air traffic control was a 2002 mid-air collision between a Bashkirian Airlines Tupolpev 154 and a DHL 757 near the border of Germany and Switzerland. The air traffic controller ordered the TU-154 aircraft to disobey the aircraft's TCAS (Traffic Collision Avoidance System) warning, resulting in both aircraft descending and making contact. In this accident, a series of events led to the air traffic control system failing, resulting in the fatal crash.

Air traffic controller fatigue has also been blamed for several aviation accidents in the United States.[4] For example, although the 2007 Comair accident in Lexington, Kentucky was caused by several factors, the investigation found that the controller was working on just two hours of rest (Ahlers, 2007). Had the controller been better rested, it is possible that he may have noticed the CRJ aircraft beginning to take off on the wrong runway. Several other incidents may have resulted from air traffic controller fatigue, and regulations have been amended in order to help minimize human error by air traffic controllers (Ahlers, 2007).

Maintenance

The maintenance department of any airline is critical to ensuring that the airline operates safely. Aviation accidents sometimes occur as a result of maintenance being performed either incorrectly or not thoroughly. However, maintenance has accounted for only 4 per cent of hull loss accidents worldwide from 1996 to 2005; this in itself is a strong testament to the generally high quality of work maintenance personnel do on a worldwide basis. Through domestic and international regulations and inspections, maintenance is usually performed to strict standards. However, unfortunately, there have been exceptions to this rule.

In 1985 a Japan Airlines 747 crashed outside of Tokyo, killing 520 passengers and crew. The accident resulted from the aircraft losing its rear stabilizer and hydraulic systems due to an explosive decompression. The accident investigation determined that repairs performed by Boeing on an earlier tail strike of the aircraft were inadequate. Over time, the repairs began to fatigue, and finally the fuselage cracked causing a massive depressurization in the rear of the aircraft.

4 According to the National Transportation Safety Board (NTSB) four aviation mishaps, between 2001 and 2007, were contributed to by air traffic controller fatigue and lack of sleep.

Unfortunately, a similar accident occurred in 2002 when a China Airlines 747 flight from Taipei to Honk Kong crashed into the ocean, killing all on board. Once again the investigation uncovered the fact that the aircraft had experienced a tail strike over 20 years earlier, and the repairs were not made up to the appropriate standards. Eventually, metal fatigue caused rapid depressurization and the subsequent accident.

On 11 May 1996, a ValuJet aircraft crashed in the Florida Everglades and killed 110 passengers and crew. According to the NTSB, the ValuJet crash was the result of failures by the airline, its maintenance contractor and the FAA. Consequently, the Transportation Department's Inspector General required that the FAA should be more proactive in monitoring airline maintenance work performed by non-certified contractors. The ValuJet crash has led to changes at the FAA, including closer scrutiny of new carriers and more monitoring of their growth.

Miscellaneous/Other

The final category of aviation accidents is miscellaneous/other, which can include a variety of things, with hijackings representing the largest share. Unfortunately, commercial aircraft are still used for ulterior (usually political) motives, as the terrorist attacks of 11 September 2001 showed. Increased screening and security will assist in helping to prevent further terrorist attacks. However, increasingly restrictive and onerous security regulations can rapidly become more detrimental to the traveling public than any small (to non-existent) increase in safety that they generate. Finally, the cause of some aviation accidents is simply unknown due to a lack of evidence or unusual issues.

Figure 13.6 displays the number of worldwide fatalities resulting from aviation accidents between 1987 and 2005. The fatalities are classified according to an ICAO accident taxonomy based on the cause of the accident. Because of varying definitions, the cause of an aviation accident could be embedded amongst the different categories. For example, flight crew error could be categorized as CFIT (controlled flight into terrain) or LOC (loss of control), depending on the accident. Regardless of such cross-classifications, the ICAO taxonomy provides a standardized worldwide definition of aviation accidents that can be of use in safety research. Since controlled flight into terrain and loss of control (the number one and number two causes) of accidents are generally caused by flight crew errors, Figure 13.6 underscores the fact that the vast majority of aviation accidents are still caused by human error.

As Figure 13.6 displays, CFIT is the number one category of fatalities in aviation. A CFIT accident can result from numerous issues. although pilot error is usually a central cause. LOC is the second major category, with LOC accidents being a result of numerous issues. Although the ICAO classification is different, the major cause of aviation accidents remains roughly the same—human error.

CLASSIFICATION OF ACCIDENTS BY PHASE OF FLIGHT

Another classification of aviation accidents is the phase of flight when the accident occurred. Figure 13.7 displays both the number of accidents and fatalities for the worldwide commercial jet fleet between 1996 and 2005 as categorized by the phase of flight. The figure shows the large disparities between accidents and fatalities for various phases of flight. While nearly half of worldwide aviation accidents occur during landing, only 16 per cent of fatalities occur during this stage. Conversely, 42 per cent of aviation fatalities

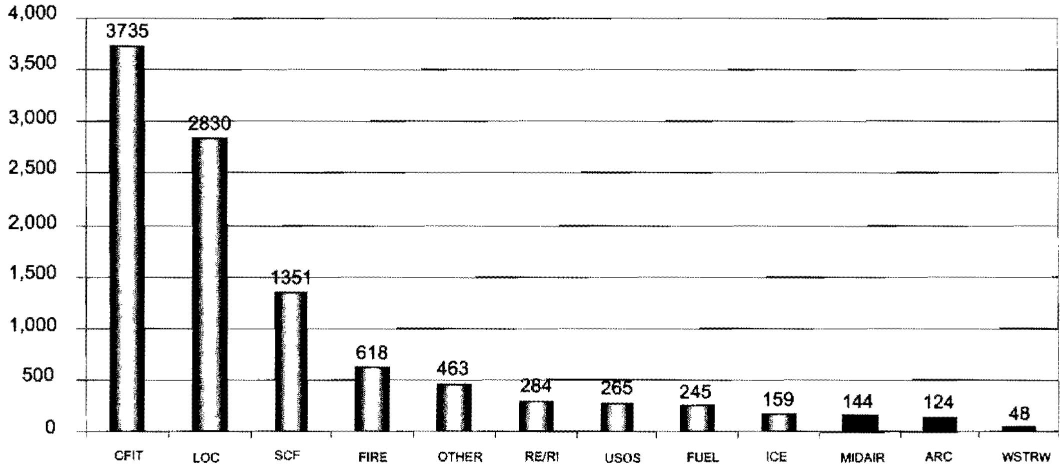

Figure 13.6 Classification of aviation fatalities, 1987–2005, according to ICAO accident taxonomy

Source: Compiled by the authors using Boeing (2006) data.

occur during climb, yet only 12 per cent of accidents occur during this phase of flight. The statistics indicate that climb is the most dangerous phase of commercial flight (as far as fatalities are concerned), since engine failures in this particular phase of flight can easily result in a fatal accident. Cruise is the safest phase of flight, even though it takes up roughly 57 per cent of a flight's duration. Although accidents and fatalities have occurred at cruise, pilots generally have more time to react and avoid more serious consequences. An example of a cruise incident that ended safely was the Air Transat A330 that ran out of fuel and glided to safety in the Azores in 2001. Had the aircraft run out of fuel at a lower altitude or during climb or approach, the outcome could have been catastrophic.

CLASSIFICATION OF ACCIDENTS BY REGION

Another important classification of aviation safety is by region. Figure 13.9 displays hull loss rates per million departures for various regions worldwide from 1996 to 2006.

As Figure 13.8 shows, aviation safety varies widely by region, with Africa being the most unsafe and North America being the safest. While these differences in aviation safety can be attributed to a variety of reasons, overall economic prosperity appears to be correlated with aviation safety. Poor regions, such as Africa, do not have the same level of safety oversight, nor infrastructure, as do developed nations. Air traffic control coverage can be sporadic across Africa' and lack of instrument landing systems can make rough-weather landings even more dangerous. Moreover, since the primary cause of aviation accidents is human error, training standards are extremely important, and it is difficult to gauge the overall training standards in the less developed countries. Finally, many airlines in developing nations use older aircraft which may be more prone to accidents. Because of these factors, one would expect the number of accidents in developing nations to be higher than in developed nations—although probably not to the extent that currently exists.

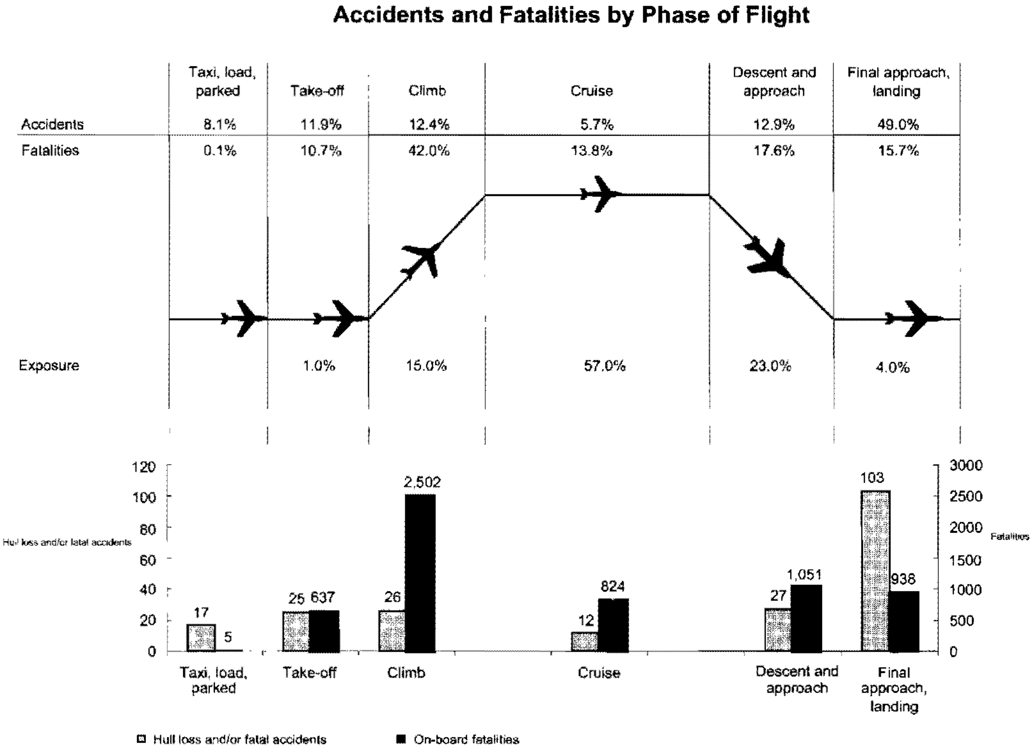

Figure 13.7 Worldwide commercial accidents and fatalities, 1996–2005, categorized by phase of flight

Source: Compiled by the authors using Boeing (2006) data.

Notes: CFIT = controlled flight into or toward terrain.
LOC = loss of control, either in-flight or on the ground.
SCF = system/component failure or malfunction, either powerplant or non-powerplant.
FIRE = fire/smoke (non-impact).
OTHER = other or unknown.
RE/RI = runway excursion or incursion, either by vehicles, aircraft, people, or animals.

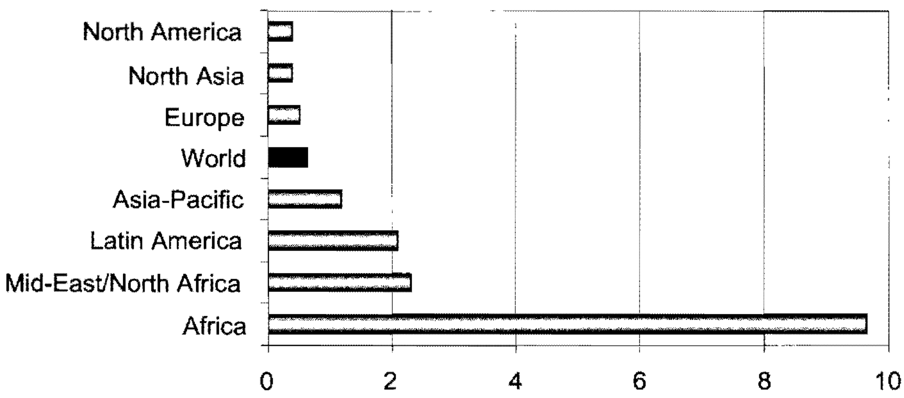

Figure 13.8 Hull losses per million departures by regions, 1996–2006

Source: Compiled by the authors using IATA (2006) data.

Figure 13.8 also enables a comparison of regions to the worldwide hull loss rate per million departures. Three regions, North America, Europe, and North Asia, are below the average, yet these three regions are also the three largest aviation markets. Three other regions, Africa, Middle East/ North Africa, and Latin America/Caribbean, were all substantially above the worldwide average. While not specifically broken down, another region of particular concern is Russia and the CIS states. In 2006 alone their hull loss rate per million departures was 19.6, which was substantially above any other region (O'Brien, 2006). As a comparison, the second-highest region, Africa, had a hull loss rate per million departures of only 2.37 in 2006 (O'Brien, 2006). Clearly Russia/CIS need to improve on their safety standards, as a hull loss rate of 19.6 is unacceptable. Substantial change will need to occur in this region to bring it in line with the rest of the northern hemisphere.

THE BASIC ECONOMICS OF SAFETY

To understand the economics of aviation safety, one needs to look at the industry from the macro-level, where the benefits of safety regulations to consumers and companies are weighed against the costs of imposing the regulations. Since the costs and benefits of safety are both explicit and implicit, it is sometimes difficult to fairly evaluate a regulation; this difficulty is one of the main reasons why many aviation safety regulations put into place by the government do not always make economic sense.

Many safety regulations enacted by governments are blanket responses to potential threats, or media-generated reactions that merely alleviate passenger concerns while not increasing safety in any substantial way. An example of this would be the requirement for all passengers to always take off their shoes while going through security screening. This regulation was created immediately following a potential terrorist threat, but the increased level of security that this extra check provides is probably negligible.

This example highlights the fact that there is a strong probability that a number of aviation safety regulations generate more costs than benefits. Indeed, it is quite likely that some costly airline regulations, even if they provide some positive safety benefits to aviation, will actually decrease net safety in society! This follows because airline regulation will drive up prices for consumers, as well as sometimes making flying more cumbersome and time-consuming, thus causing some who would have otherwise flown to instead travel by rail or, worse, drive. Since, as explained earlier, these other modes of transportation are far more dangerous than flying, lives are lost whenever costly regulations convert air travelers to ground travelers. Likewise, eliminating costly airline regulations that do little or nothing to improve safety would reduce ticket prices, draw people away from cars and rail and to aircraft, thus improving total travel safety. To some extent, we can increase total safety by decreasing airline safety.

The "Southwest rule" is a good example of a regulation that, if abolished, would probably increase total safety. In fact, the benefit of this regulation is probably about zero, since in all the years prior to the rule there were no injuries caused by the aircraft slowly taxiing toward take-off position as passengers continued to settle into their seats. Ben-Yosef (2005, Chapter 6) argues that the FAA's eventual 1996 decision to ground ValuJet Airlines, the low-cost leader in the eastern United States at the time, also cost far more in increased highway fatalities than any conceivable benefit. Ben-Yosef also documents the general public's misperception that aircraft maintenance failures cause most accidents, and, more broadly, the huge disconnect between public perception regarding the causes of airline crashes versus the reports of safety regulators and other industry experts. Breyer (1993) maintains that the interaction of public misperceptions and the political process produces an essentially random agenda. It is, of course, difficult for politicians to rise above the politics of safety and make decisions based on costs and benefits. The notion that airlines may well be too safe, forcing travelers into riskier transportation modes, appears to be too sophisticated to be effectively dealt with through the political process. Perhaps this will change in time with continued efforts to educate the public in this regard; however, such efforts have so far been remarkably ineffective and probably will not succeed in the future.

The very basic analysis provided above is an example of a simple economic cost–benefit analysis for aviation safety. Such an analysis could be conducted for almost all safety regulations to determine whether the regulation is economically efficient. However, since it may be difficult to quantify all the benefits from improved safety, such analysis is rarely undertaken. Therefore, and as pointed out above, the political and bureaucratic process usually assumes, especially where aviation safety is concerned, that the benefits of almost any safety or security regulation outweigh the costs. Moreover, as mentioned earlier in the chapter, the reasons for this are easy to identify; although commercial aviation is a critical sector of the economy, the vast majority of people who fly do so only a few times a year. Therefore, the actual number of people who fly on a regular basis is only a small proportion of the population.[5] Add to this the media tendency to sensationalize all aviation-related accidents or incidents, and the natural inclination of regulators to avoid even the remote appearance of not being vigilant on safety, and one can readily see that even elementary cost–benefit analysis would be difficult to implement in this culture.

Furthermore, it is also probably true that, if one were to perform the economic cost–benefit analysis for all safety regulations, many would pass due to the substantial benefits from improved safety. However, many regulations would also fail, largely because they were enacted in response to political pressure. Some safety decisions do not receive economic scrutiny because there are other competitive factors in play. For example, the regulation banning aircraft push-back until everyone is seated was created in part as a response to other airlines lobbying against Southwest's practices. Therefore, for the reasons discussed earlier, it is probably true that, contrary to popular belief, aviation safety and security exceed the levels that might be considered economically efficient.[6]

IS IT POSSIBLE TO TAKE THE POLITICS OUT OF SAFETY REGULATION?

Robert Poole (1981) has argued that airline safety regulation can be effectively privatized, thus driven by economic analysis rather than shallow politics. Poole envisions a system where, in essence, private insurers replace politicians as the ultimate safety authorities. Insurers have a vested interest in assuring that airlines do not take imprudent risks, since they must pay for any damages caused by an accident. However, insurers do not have to explain their decisions to uninformed voters. Thus, insurance companies are unlikely to have any interest in continuing the "Southwest rule" or any other regulation that doesn't

5 Again, this can be contrasted to the automobile where the ill-fated safety regulation tying the ignition of the car to a fastened seatbelt was quickly abandoned when a significant proportion of the population (automobile drivers and voters) discovered what a nuisance this particular regulation would be in practice.

6 The opposite is probably true of automobile safety.

truly improve safety. Likewise, insurers would allow airlines to cancel safety programs that produce more costs than benefits—that is, an airline would be allowed to slightly increase risk as long as a higher insurance premium was paid to cover it. This sort of behavior is already observable, in that consumers can buy personal injury insurance, at a higher price, for a motorcycle or subcompact car even though such vehicles are substantially less safe than standard size automobiles. Insurers know better than to try to eliminate all risk—any movement in that direction would result in customers leaving them for a more reasonable insurer. The fact is that insurers ban only imprudent risks, and insist that customers pay more for any increase in prudent risks. In this setting, government could merely require that airlines purchase legitimate insurance, and then let insurers handle the details.

The FAA and comparable regulators in other countries could continue their same basic mission but be converted into a private organization, paid by insurers rather than by taxpayers. The head of the FAA would, of course, no longer be a political appointee but a private manager, appointed by a board of stockholders, comparable to any other corporate CEO. Ideally, this privately reborn FAA would be driven by economic analysis, able to maintain a more long-term focus on true accident risks, rather than being driven by the latest headlines and the whims of politics.

Although there is currently no private regulator of airline safety, Poole points to a number of examples of private safety regulation in other areas. Underwriters Laboratories, for example, sets safety standards for a number of electronic components. (It seems that many people have assumed that the company is some sort of public agency, since their function is so commonly associated with government.) US fire departments, though usually government bureaucracies themselves, are, in effect, regulated by a private insurance organization, the National Board of Fire Underwriters (NBFU). The NBFU inspects fire departments to rate their response time, quality of equipment, staffing, and so on, but does not, of course have the authority to demand corrections where there are problems. However, if a local fire department is poorly rated, then fire insurance premiums in the area are raised, immediately exposing inadequacies. Consequently, fire departments generally work with the NBFU to correct problems and improve themselves as needed. The NBFU began this function in 1890, stepping in to deal with the problem of widely varying quality in firefighting, and, after more than a century of experience, continues to operate without incident. The standards set by the NBFU do not appear to have anything comparable to the Southwest rule or other inappropriate regulation.

More broadly, any independent private agency that provides product information and ratings is performing a function similar to the private airline regulation envisioned by Poole and other supporters. Just as consumers have difficulty judging whether an airline is appropriately safe, they may also have trouble judging the safety and general quality of many products. Government agencies sometimes provide such judgments to some extent, and do not charge consumers for this service. However, government does not do this extensively enough to satisfy consumers; therefore, we have private companies like Consumer Digest and Consumer Reports that inspect products and make recommendations to consumers. Likewise, when investors want to know more about a company, they turn not to the Securities and Exchange Commission, but to Moody's or Standard and Poor's.

Naturally, many people would question whether private regulation can really perform better than government. It is instructive to compare the incentives and operational nature of private and public entities in this area. Poole and other privatization proponents argue that private regulators have superior incentives and better flexibility to deal with problems.

If the FAA were private, working under contract for insurers, it could be replaced, in part or in full, by another organization if it performed poorly. Basically, any entrepreneur, perhaps a former FAA employee, could approach an insurer and make a case that she could provide inspection in a given area more efficiently. Knowing it could be fired, the private FAA would seem to have a strong incentive to operate efficiently, appropriately monitor the competence and integrity of its workers, and so on. Likewise, being private, the FAA would be able to more freely adjust policies or fire employees who weren't performing well and more rapidly promote those who were. In contrast, the current FAA, critics argue, knows that any failure in its mission is likely to be greeted with an increased budget. Likewise, bureaucratic red tape limits its ability to be efficient and effectively deal with employees.

SAFETY PREVENTION

The main reasons for the rapidly decreasing aviation accident rate since the 1950s are the various safety programs/inventions adopted by safety regulators, airlines, and aircraft manufacturers. All three groups have combined resources to make aviation the safest mode of transportation. This increase in safety has ultimately helped stabilize the industry, and make it a more attractive transportation option for consumers. Based on the incentives described previously, methods to increase safety have sizeable economic benefits. Some examples of these are: aging aircraft regulations, collision avoidance systems, wind-shear detection, de-icing, and human factors.

As mentioned previously, aging aircraft can compromise aviation safety in some specific cases. In order to help minimize this problem, the FAA and JAA (Joint Aviation Authority)[7] require specific component overhauls at specified intervals (well ahead of the time the components would be expected to fail). Some countries take aging aircraft regulations to an extreme by not allowing airlines to operate commercial aircraft over a certain age. As pointed out in the previous section, this type of rather arbitrary safety regulation might improve safety somewhat, but it will also impose significant extra costs on the industry which will, of course, ultimately be passed down to the passengers. From an economic standpoint, passengers may not be better off as a result of such stringent aircraft age regulations (with no decrease in overall safety levels).

Although mid-air collisions have never been the number one cause of aviation accidents, joint research by governments and industry resulted in the development and deployment of the Traffic Alert and Collision Avoidance System (TCAS). TCAS alerts pilots of a possible mid-air collision and provides instructions to help avoid a serious accident. The invention of the TCAS has reduced the number of mid-air collisions, although it has not eliminated them altogether, as evidenced by the fatal accident involving the DHL and Bashkirian Airlines aircraft over Europe in 2002. Furthermore, the TCAS is not immune to human error, as pilots and air traffic controllers can still make mistakes and disobey TCAS warnings, resulting in tragedy.

Wind-shear represents another significant threat to aircraft, since it can cause an aircraft to become uncontrollable. Previously wind-shear was undetectable; however, through government and industry research, warning devices have been created to alert pilots of possible wind-shear conditions. Based on the wind-shear warnings, regulations

7 The European equivalent of the FAA.

have been developed to help insure that aircraft do not fly during dangerous wind-shear conditions. The wind-shear alerts enable a pilot to take appropriate action to avoid dangerous situations.

Although the American Eagle ATR72 de-icing accident highlights the fact that fatal accidents still occur due to ice forming on the wings, advances in anti-icing have significantly reduced this type of accident. Aircraft manufacturers have designed aircraft with anti-icing boots, while chemical compositions have enabled de-icing to be carried out on the ground.

SUMMARY

This chapter has provided an overview of the state of aviation safety and security. In general terms, safety and security in aviation have been highly effective from an economic point of view, although there are probably numerous rules and regulations that could be relaxed with no decrease in overall safety. In fact, stringent regulations for the introduction of new technologies and procedures probably act to decrease, rather than increase, overall safety in the industry. Future developments in this field will have to center on replacing human judgment with automated technologies. However, these developments have been, and continue to be, extremely difficult to implement due to bureaucratic and political inertia.

REFERENCES

Ahlers, M.M. (2007). NTSB: Air Controller Fatigue Contributed to 4 Mishaps. *CNN*, 10 April. Retrieved on 23 April 2007 from: http://www.cnn.com/2007/US/04/10/controller.fatigue/index.html.

Barnett, A. (2001). Air Safety: End of the Golden Age? *Journal of the Operational Research* Society, August. Based on the Year 2000 Blackett Memorial Lecture, Royal Aeronautical Society, November 2000.

Ben-Yosef, E. (2005). *The Evolution of the US Airline Industry: Theory, Strategy and Policy*. Studies in Industrial Organization, Volume 25. Dordercht: Springer.

Boeing (2006). *Statistical Summary of Commercial Jet Airplane Accidents*, May. Seattle: The Boeing Company. Retrieved May 2007 from: http://www.boeing.com/news/techissues/pdf/statsum.pdf.

Borenstein, S. and Zimmerman, M.B. (1988). Market Incentives for Safe Commercial Airline Operation. *The American Economic Review*, 78, pp. 913–35. Retrieved on 9 April 2007 from Proquest at: http://www/proquest.com.

Breyer, S. (1993). *Breaking the Vicious Cycle: Toward Effective Risk Regulation*. Cambridge: Cambridge University Press.

Federal Aviation Administration (FAA) (2000). FAA Proposes Maintenance Fines Totaling $988,500 against Alaska Airlines, 4 December. Retrieved on 12 April 2007 from: http://www.faa.gov/news/press_releases/news_story.cfm?newsId=5296.

Federal Aviation Administration (FAA) (2002). FAA Proposes $805,000 Fine Against United Airlines, 3 December. Retrieved on 12 April 2007 from: http://www.faa.gov/news/press_releases/news_story.cfm?newsId=6372.

Mitchell, G. and Maloney, M.T. (1989). Crisis in the Cockpit? The Role of Market Forces in Promoting Air Travel Safety. *The Journal of Law and Economics*, 32, pp. 329–56.

O'Brien, B. (2006). IOSA: IATA Operational Safety Audit. IATA. Retrieved on 12 April 2007 from: http://www.iata.org/NR/rdonlyres/AAA225BF-0CA4-456F-8F9B-F2AD8639DF80/0/safety_iosa.pdf.

Poole, R. (1981). *Alternatives to Federal Regulatory Agencies.* Lanham, MD: Lexington Books.

Squalli, J. and Saad, M. (2006). Accidents, Airline Safety Perceptions and Consumer Demand. *Journal of Economics and Finance*, 30, pp. 297–305. Retrieved on 9 April 2007 from Proquest at: http://www.proquest.com.

Vasigh, B. and Helmkay, S. (2000). An Empirical Examination of Airframe Manufacturers' Safety Performance: Boeing Versus Airbus. *Public Works Management & Policy*, 5, pp. 147–59.

Vasigh, B. and Helmkay, S. (2002). Airline Safety: An Application of Empirical Methods to Determine Fatality Risk. In D. Jenkins (ed.), *Handbook of Airline Economics*. New York: McGraw-Hill, pp. 501–11.

Wong, J.T. and Yeh, W.C. (2003). Impact of Flight Accident on Passenger Traffic Volume of the Airlines in Taiwan. *Journal of the Eastern Asia Society for Transportation Studies*, 5, pp. 471–83.

Index