
Empirical Studies in Applied Economics

EMPIRICAL STUDIES IN APPLIED ECONOMICS

JEFFREY A. DUBIN
Division of the Humanities and Social Sciences
California Institute of Technology
Pasadena, California

Kluwer Academic Publishers
Boston/Dordrecht/London

*This book is dedicated to
Jackie, Andrew, and Ethan.*

Contents

List of Figures	xi
List of Tables	xiv
Preface	xv
Introduction	xvii
1. REVEALED MARKET POWER OF A NATURAL GAS PIPELINE	1
1.1 Introduction	1
1.2 Basis Differentials	2
1.3 Discounting	9
1.3.1 The Data Assembled	9
1.3.2 Analysis	12
1.3.3 Transaction Data Analysis	14
1.4 Econometric Results	21
1.5 Conclusion	30
2. THE DEMAND FOR NFL FOOTBALL	31
2.1 Introduction	31
2.2 The Demand for NFL Football Tickets	31
2.3 Bandwagons, Social Influences, and Group Demand	33
2.4 Becker's Model	34
2.5 The Demand for NFL Football	39
2.6 Conclusions	49
3. SHIFTS IN THE DEMAND FOR DIRECT MAIL	51
3.1 Introduction	51
3.2 Comparison of Original, Revised, and Updated Direct Mail Models	51
3.2.1 Comparison of Revised and Updated Models	52
3.2.2 Addressed Admail—Specific Results	53

3.3	The Shift in Demand Between 1995 and 1996	53
3.4	Comparing Forecast Accuracy	54
3.4.1	Forecasting Addressed Admail Volume	57
3.4.2	Measures of Forecast Accuracy	58
3.4.3	Forecast Results—Confidence Intervals	58
3.4.4	Forecast Results—Conclusions	59
4.	VALUATION OF A TECHNOLOGY PATENT	65
4.1	Introduction	65
4.2	Scope Limitations	66
4.2.1	The Variety of Computer Upgrade Technologies	67
4.2.2	Coprocessors and Empty Slots	68
4.2.3	Proprietary Upgrade Methods	69
4.2.4	Dual-Socket Technology Was Rare	70
4.3	Duration Limitations	71
4.3.1	A Product Lifecycle Model for CPU Sales	74
4.4	Limitations in the Royalty's Value	76
4.4.1	The Presence of Non-Infringing Substitutes Limits the Value of a Patent	76
4.4.2	The Value of a Consumer Upgrade Option	80
4.4.3	Trademark and Patent Valuation	82
4.5	Conclusions	83
5.	STATISTICAL ANALYSIS OF SYNERGY	87
5.1	Introduction	87
5.2	Case-Control Studies	90
5.2.1	Tests for Case-Control Studies—Multiplicative Case	91
5.2.2	Woolf's Method	95
5.2.3	Tests for Case-Control Studies—Additive Case	99
5.2.4	Maximum Likelihood	101
5.3	Bonferroni Intervals and Simulations	106
5.3.1	Bonferroni Interval	106
5.3.2	Simulation Methods	107
5.4	Synergy Indices	107
5.4.1	Rothman's <i>S</i> Index	108
5.4.2	Attributable Proportion Index	109
5.4.3	Additive-Multiplicative Measure	110
5.5	Cohort Studies	110
5.5.1	Determination of Death Rates	113
5.5.2	Variance of the Additive and Multiplicative Statistics	114
5.5.3	Variance of the Synergy Index	115
5.6	Conclusion	116
6.	TESTS OF SYNERGY	117

6.1	Introduction	117
6.2	Review of Studies	119
6.2.1	Hammond, Selikoff, and Seidman (1979)	119
6.2.2	Selikoff, Seidman, and Hammond (1980)	120
6.2.3	Berry, Newhouse, and Turok (1972)	121
6.2.4	Berry, Newhouse, and Antonis (1985)	122
6.2.5	McDonald, Liddell, <i>et al.</i> (1980)	122
6.2.6	Liddell, Thomas, <i>et al.</i> (1984)	123
6.2.7	Martischnig, Newell, <i>et al.</i> (1977)	124
6.2.8	Pastorino, Bearino, <i>et al.</i> (1984)	124
6.2.9	deKlerk, Musk, <i>et al.</i> (1991)	125
6.2.10	Bovenzi, Stanta, <i>et al.</i> (1993)	126
6.2.11	Blot, Harrington, <i>et al.</i> (1978)	126
6.2.12	Blot, Morris, <i>et al.</i> (1980)	127
6.2.13	Kjuus, Skjaerven, <i>et al.</i> (1986)	128
6.2.14	Vena, Byers, <i>et al.</i> (1985)	129
6.3	Tests of Hypothesis	129
6.3.1	Hammond, Selikoff, and Seidman (1979)	129
6.3.2	deKlerk, <i>et al.</i> (1991)	131
6.4	Meta-analysis of the Fourteen Synergy Studies	138
6.5	Conclusions	140
7.	CONCENTRATION IN THE CHEMOTHERAPY DRUG MARKET	141
7.1	Introduction	141
7.2	Market Structure	142
7.2.1	R&D Economies of Scale	142
7.2.2	Monopoly Power from Statutory Exclusivity	143
7.2.3	Brand Loyalty	143
7.2.4	Generic Drug Introduction	144
7.2.5	Creating and Maintaining a Monopoly	145
7.3	Statutory Exclusivity in the United States	146
7.3.1	Pure Food and Drug Act of 1906 and Food, Drug and Cosmetic Act of 1938	146
7.3.2	1983 Orphan Drug Act	146
7.3.3	Drug Price Competition and Patent Restoration Act of 1984	147
7.3.4	Exclusive Government Agency Licenses	148
7.4	History of Paclitaxel	148
7.5	Market Power Measures	150
7.5.1	Herfindahl-Hirschman Index (HHI)	150
7.5.2	Weighted HHI	152
7.5.3	Weighted Market Share Statistic	152
7.5.4	Market Share Weighted HHI	153
7.6	The Relevant Product Market for Paclitaxel Drugs	153

7.6.1	Antitrust Markets	154
7.6.2	Market Delineation—Demand Substitutability	154
7.6.3	Market Delineation—Supply Substitutability	155
7.6.4	The Relevant Product Market for Taxane Drugs	155
7.6.5	HHIs—Market Concentration Results	157
7.6.6	Market Share Weighted HHI	162
7.7	Conclusion	163
8.	ALASKAN POLICE SERVICES	167
8.1	Introduction	167
8.2	Historical Background	168
8.3	Theory	170
8.4	Allocation Analysis	172
8.4.1	The Dataset	182
8.4.2	Analysis of Variance	185
8.4.3	Econometric Model	187
8.5	Conclusions	189
9.	FAST FOOD HEALTH SCARE	193
9.1	Introduction	193
9.2	Data Sources	195
9.2.1	Restaurant Trends Data	195
9.2.2	Bureau of Economic Analysis	195
9.2.3	National Income and Product Accounts	195
9.2.4	Bureau of Labor Statistics	195
9.3	Econometric Models	196
9.3.1	Revenue per Capita	197
9.3.2	Average Unit Volume	197
9.3.3	Stores per Capita	197
9.4	Simulations and Conclusions	197
9.4.1	Market Share Logit Model	198
9.4.2	Conclusions	201
	Index	207

List of Figures

1.1.1	Henry Hub Deliveries to Venice	4
1.1.2	Henry Hub Deliveries to Kosciusko	5
1.1.3	Venice Deliveries to Henry Hub	6
1.1.4	Kiln Deliveries to Venice	7
1.1.5	Henry Hub Deliveries to Koch Zone 2	8
1.2	Delivered Volumes by Type of Service	15
1.3	Weighted Average Transportation Rate	16
1.4	Volume Weighted Average Discount for All Transactions	17
1.5	Weighted Average Monthly Receipt and Delivery Basis Value	18
1.6	Weighted Average Monthly Basis Differential	19
2.1	Aggregate Demand Depends on Audience Size	36
2.2	Multiple Equilibria and Demand Instability	38
2.3	Ticket Sales as a Percentage of Capacity (1995–1999)	40
2.4	Attendance as a Percentage of Ticket Sales (1995–1999)	41
2.5	Attendance as a Percentage of Capacity (1995–1999)	42
2.6	Ticket Sales as a Percentage of Capacity (1995–1999)	44
2.7	Attendance as a Percentage of Ticket Sales (1995–1999)	45
2.8	Attendance as a Percentage of Capacity (1995–1999)	46
3.1	Shift in Demand Between Fiscal Years 1995 and 1996	55
3.2	Forecasted Volume versus Actual Volume	56
3.3	Estimated Confidence Interval for Addressed Admail	60
4.1	Intel Processor Timeline	73
4.2	Speed Comparison of Intel Processors	75
4.3	A Series of Technological Changes	77
4.4	U.S. Microsystems Market, 1981–1991	78
4.5	CPU Market Share (1990–1995)	79
4.6	1991 Unit Sales for Top Manufacturers	85
6.1	Lung Cancer Relative Risks	133
6.2	Synergy <i>t</i> -Statistics	134

6.3	Synergy Index Values and Confidence Intervals	135
6.4	Attributable Proportion Statistic	136
6.5	Gamma Value	137

List of Tables

1.1	Alternative Pipelines by Power Plant	13
1.2	Variable Glossary	22
1.3.1	Transportation Rate Model 1	23
1.3.2	Transportation Rate Model 2	24
1.3.3	Transportation Rate Model 3	25
1.3.4	Transportation Rate Model 4	26
1.3.5	Transportation Rate Model 5	27
1.3.6	Transportation Rate Model 6	28
1.3.7	Transportation Rate Model 7	29
1.3.8	Transportation Rate Model 8	29
2.1	Data Sources and Variables	43
2.2	Variable Glossary	47
2.3	Variable Statistics	47
2.4	Demand Models	48
3.1	Variable Glossary	52
3.2	Addressed Admail (Original)	61
3.3	Addressed Admail (Revised)	62
3.4	Addressed Admail (Updated)	63
3.5	Addressed Admail 95% Confidence Intervals	64
4.1	Selected Processor History*	72
4.2	Top Twenty Computer Manufacturers	84
6.1	Hammond, <i>et al.</i> (1979)	120
6.2	Selikoff, <i>et al.</i> (1980)	120
6.3	Berry, <i>et al.</i> (1972)	121
6.4	Berry, <i>et al.</i> (1972)	122
6.5	McDonald, Liddell, <i>et al.</i> (1980)	123
6.6	Liddell, Thomas, <i>et al.</i> (1984)	123
6.7	Martischinig, Newell, <i>et al.</i> (1977)	124
6.8	Pastorino, Bearino, <i>et al.</i> (1984)	125

6.9	deKlerk, Musk, <i>et al.</i> (1991)	126
6.10	Bovenzi, Stanta, <i>et al.</i> (1993)	126
6.11	Blot, Harrington, <i>et al.</i> (1978)	127
6.12	Blot, Morris, <i>et al.</i> (1980)	128
6.13	Kjuus, Skjaerven, <i>et al.</i> (1986)	128
6.14	Vena, Byers, <i>et al.</i> (1985)	129
7.1.1	Alkylating Agents, Anthracyclines, Antibiotics	159
7.1.2	Anti-Metabolites, Antineoplastics	160
7.1.3	Hormones	161
7.2.1	Concentration—Bristol-Myers Oncology	163
7.2.2	Concentration—TAP Pharmaceuticals	163
7.2.3	Concentration—Zeneca Pharmaceuticals	164
7.2.4	Concentration—Barr Labs	164
7.2.5	Concentration—Pharmacia & Upjohn	165
7.2.6	Concentration—Glaxo Wellcome Oncology	165
8.1	Places Used in Statistical Analysis	173
8.2	Alaska Allocation Analysis	191
9.1	Variable Glossary	196
9.2	Revenue per Capita	198
9.3	Average Unit Volume	199
9.4	Stores per Capita	200

Preface

How odd it is that anyone should not see that all observation must be for or against some view if it is to be of any service! –Charles Darwin, September, 1861.

The chapters in this monograph collect some of my empirical research conducted during the last few years. This book required the behind the scene efforts of many people.

For assistance with the projects, I'd like to thank Charles Cicchetti, Colin Long, Mike McMahon, Randy Sugarman, Jim Baldwin, and Captain Chris Stockard. For able research assistance, I'd like to thank Elizabeth Knebel, Akhil Ahuja, Kim Harle, and Phil Schofield. For helpful advice and support, Dan McFadden and Cheng Hsaio. For friendship, I must mention Harry Kraus, Glenn Westreich, and Doug Rivers. For a wonderful family, I thank my wife Jackie and my great children Andrew and Ethan.

On the production side, I thank Victoria Mason, Abbey Dellman, Ginny Feliciano, and Bob Turring. At Kluwer, I was assisted by Tom Randall, Suzanne St. Clair, Marion Scott and the anonymous referees.

As always, the opinions provided in these chapters are solely mine. I retain the responsibility for any errors which remain.

JEFFREY DUBIN

Introduction

This monograph contains nine previously unpublished studies in applied economics. Each chapter is intended to be self-contained and each provides a separate empirical analysis. The exceptions to this are Chapters 5 and 6 which provide a theoretical and empirical analysis of synergy in respective parts.

A summary of each chapter follows.

The Revealed Market Power of a Natural Gas Pipeline

In this chapter I discuss the collapse of the basis differential between the eastern and western portions of the southern United States and demonstrate how the basis differential affects a natural gas pipeline's ability to raise or lower its prices up to the maximum tariff rate. "Basis" refers to the difference in gas prices between a cash market at a given location and the future price at a different location for a particular time period. A basis point establishes a futures price for delivery of natural gas at a specific geographic location at a specific time. Arbitrage limits the transportation rate to the differences in basis differential in a perfect market where all connections are feasible and there are no limitations on flow direction. I examine the basis differential issue as well as other market conditions that potentially affect discounting. My empirical analysis uses transaction data for a specific pipeline and reveals intense competition in the geographic market in which it operates. I find that market forces such as the prevailing basis differentials, elastic demand, and the number of connected alternatives all play significant roles in determining a pipeline's discount behavior.

The Demand for NFL Football

This chapter considers the demand for NFL football. Football demand is unusual as compared with many goods and services. Football demand is characterized by excess demand and generally tight market. A theory of Becker (1991) helps explain the demand for football. Becker's theory is that demand in some situations depends on social interaction and the size of the crowd. Us-

ing data from 1995 through 1999 for all NFL teams during their regular season, I construct an econometric model of the demand for NFL football. I use this model to test Becker's theory and conclude that the demand curve slopes upward in the relevant range as anticipated by the theoretical model. In this chapter, I discuss the option values inherent in NFL season tickets, the conditions under which football ticket sales are likely to be successful, the importance of initial conditions for establishing successful future ticket sales, and the importance of advertising and marketing. I also find that initial conditions, which help determine the future attendance in a bandwagon sport such as football, are very important in determining future outcomes.

Detecting and Measuring Shifts in the Demand for Direct Mail

This chapter evaluates the forecast accuracy of a structural econometric demand model for direct mail in Canada. Direct mail also known as advertised admail is used to provide advertising to consumers through the mail system. My original model was developed in March 1986 and was based on the period ending January 1996 using twelve years of historical monthly data. In this chapter, I update the regression results for the April 1989 through January 1996 period and provide new results from twenty additional monthly observations for the period of February 1996 through September 1997. I conclude that a shift the demand curve for direct mail occurred during this period. The shift in demand was due to three factors: retail sales improvements; decline in real costs of paper and printing; and general trend effects. I also find that the direct mail demand model does a better job at predicting levels rather than turning points.

Valuation of a Technology Patent—Scope, Duration, and Royalty

The purpose of this chapter is to establish the value of a particular patent and illustrate the role of scope, duration, and royalty rate in determining this value. The patent I examine was awarded to the computer manufacturer Acer in 1991. The value of this technology patent is demonstrated to be small due to the availability of substitutes and the limited applicable scope. I discuss the Bass product life-cycle model as it applies to this patent and develop a theoretical model of a consumer's decision to upgrade a personal computer. Finally, I calculate the royalties that Acer might have gained if they had attempted to enforce their patent.

Statistical Analysis of the Additive and Multiplicative Hypotheses of Multiple Exposure Synergy for Cohort and Case-Control Studies

In epidemiological studies, where there are multiple causes of a particular disease, the issue arises as to whether the multiple causes have a synergistic relationship so that their combined effect is both greater than that of either activity alone, and greater than what one would expect by the sum of their individual risk contributions. Two hypotheses are frequently tested. The first hypothesis states that when the sources of disease act independently, the relative risk of disease, given exposure, is an additive relationship. A second hypothesis states that the relationship between disease and the two causal factors is multiplicative. In this case, the combined risk is the product of the individual risks. This chapter considers several methods for determining the relative odds ratio, including the case-control method and the cohort method. I focus on the theory of testing the hypotheses of additivity and multiplicativity for the relative risk measures. I also derive various synergy indices and develop their distributions.

Tests of the Additive and Multiplicative Hypotheses of Multiple Exposure

In this chapter, I consider the possible synergistic relationship between tobacco smoking and asbestos exposure. Asbestos refers to a group of naturally occurring flexible fibers that may be separated and woven. There are two major types of asbestos. The possible link between smoking and asbestos, was observed by Selikoff, et al. (1968). In this early study, Selikoff and his co-authors noted that combined exposure to asbestos and smoking produced more lung cancer deaths than might be expected under the simple addition of the risks taken separately. Thus, a potentially synergistic relationship was identified between the two carcinogens while the form of that synergy was left to be determined. In this chapter, I review 14 studies published between 1979 and 1992, which have applied either the cohort or case-control method to analyze synergy effects between smoking and asbestos exposure. I review these studies and summarize their findings in a standardized fashion to allow the calculation of relative risks and variances. I then summarize my findings for the additive and multiplicative hypothesis based on the tests developed in Chapter 5. I also perform a meta-analysis using the data from the 14 studies to see whether the specific conclusions regarding synergy are characteristic of certain study attributes.

Concentration and Competition in the Chemotherapy Drug Market

The pharmaceutical industry ranks among the top industries in the United States in sales and research and development. This chapter considers competi-

tion in this industry and investigates the relationship between market structure, patent protection, and concentration. To illustrate several basic propositions in the pharmaceutical industry I narrow the focus of the analysis to the cancer drug market. One purpose of this chapter is to define the relevant economic market for a leading chemotherapy drug used to treat breast, lung, and ovarian cancer and to consider concentration in that market. I discuss the market structure of the pharmaceutical industry and consider research and development, brand loyalty, and factors which lead firms to create and maintain monopolies. I also discuss various market power measures including the Herfindahl-Hirschman statistic and modifications to this statistic used in a multi-market setting.

The Allocation of Police Services in Rural Alaska

In this chapter I examine whether the State of Alaska allocates scarce police resources in a race-based manner. My analysis explores this hypothesis by developing an econometric model to explain the provisions of police services in rural Alaska. I examine the provisions of police services in rural Alaska based on the number of hours Alaska State Troopers spend in 344 places in their territory or general area of responsibility in the State of Alaska. By examining Trooper activity I focus on decisions made at the State level for the numerous villages in Alaska and thus exclude from analysis the provision of police services in the few large municipalities of Alaska. To make a comparison of Alaskan native and non-native villages I examined several factors that might explain why Troopers expended the hours that they did. To accomplish this, I collected socio-demographic data for each place, geographic data, and information regarding year-round road accessibility. I also collected information on whether the place examined had some form of self-provision of law enforcement. My analysis reveals that no one factor can completely explain the variance in Trooper hours spent in any particular place. However, several factors are statistically significant. Race was not a significant factor in determining the Trooper's allocation. To the contrary, I find that Troopers are allocated rationally, in an efficient manner that attempts to balance many factors, especially scarce resource and wide disparities in access and distance to outlying places.

Financial Market Reaction to the Fast Food Hamburger Health Scare of 1993

In 1993 a bacteria epidemic killed several children in the Pacific Northwest. The bacteria was a virulent strain of *E. coli* and was traced to undercooked hamburgers at Jack-in-the-Box restaurants. In this chapter I use historical sales data to estimate the demand for hamburgers sold in the fast-food market. I

use this model to compare Jack-in-the-Box actual and forecasted sales for the two-year period following the E. coli outbreak. My conclusion is that Jack-in-the-Box sustained much greater losses than those reported in the press. I also find that the financial markets correctly anticipated that sales losses were worse than those reported. Additionally, the initial stock price decline reveals that investors expected that the sales losses would be permanent.

Chapter 1

THE REVEALED MARKET POWER OF A NATURAL GAS PIPELINE

1.1. Introduction

A natural gas pipeline operates in a regulated market environment in which transportation rates are set by regulators. Transportation rates are set using a fixed plus variable changing scheme in order to achieve a revenue path that covers the cost of providing service. Regulators have allowed pipelines to discount their rates from the maximum allowed by regulation in order to meet the competition in the market place. While isolated incidents exist, full deregulation of the natural gas transportation market has not been allowed by regulators. The discounting behavior of a pipeline in a regulated environment provides indirect evidence of the competitiveness in the market place.

In this chapter I discuss the collapse of the basis differential between the eastern and western portions of the southern United States and demonstrate how the basis differential affects a natural gas pipeline's ability to raise or lower its prices up to the maximum tariff rate. A basis point establishes a futures price for delivery of natural gas at a specific geographic location at a specific time. Arbitrage limits the transportation rate to the differences in basis differential in a perfect market where all connections are feasible and there are no limitations on flow direction. I examine the basis differential issue as well as other market conditions that potentially affect discounting. My empirical analysis uses transaction data from a specific pipeline, Koch Gateway (Koch Gateway Pipeline Company). The transactions reflect all movements of natural gas on the Koch Gateway system to power plants. The analysis therefore omits some transaction which are off system (i.e., movements to other pipelines) and deliveries to natural gas delivery companies. My discount analysis reveals that Koch Gateway faces intense competition in the geographic market in which it operates.

In section 1.2, I discuss the collapse of the basis differential between the eastern and western portion of the geographic market in which Koch Gateway operates. I show how this collapse affects Koch Gateway's ability to charge its maximum tariff rates. In Section 1.3, I discuss Koch Gateway's historic discounting for the pipeline transportation services it provides. In Section 1.4, I describe and present the results from my econometric analysis of Koch's historical discounting of its transportation services. In Section 1.5, I present my conclusions.

1.2. Basis Differentials

"Basis" traditionally refers to the difference in gas prices between a cash market at a given location and the future price at a different location for a particular time period. The future price is often measured at the Henry Hub. The Henry Hub is a basis point located in southern Louisiana. It is the official delivery point for natural gas futures trading on the New York Mercantile Exchange. "Basis differential" is sometimes used to refer to the difference in cash prices at two different points. For example, let's compare the price of South Texas gas to the price at the Henry Hub. Assume that South Texas gas is traded (bought and sold) at \$1.60 per MMBTU and Henry Hub gas is traded (bought and sold) at \$1.70 per MMBTU. The basis differential in this case would be negative 10 cents. These differentials are established by the market value of gas at different points to reflect location differences and market conditions.

The natural gas market is constantly evolving. This is reflected by the collapse of the basis differential between the western portion of Koch Gateway's system and the eastern portion of Koch Gateway's system. For example, for the period January 1991 through July 1994, on Koch Gateway's South Texas to North Louisiana segment, the basis differential averaged about 5 cents. As late as February 1994, the basis differential approached 15 cents on this portion of the Koch Gateway system. Similar basis differentials were present on other segments of the Koch Gateway system such as East Texas to Mississippi and Alabama. On the South Texas to Mississippi and Alabama portion of the system, the basis differential approached 25 cents in February 1994. These basis differentials no longer exist on Koch Gateway.

Five years ago, the geographic area in which Koch Gateway operated sat between the western market (characterized by low prices) and the tighter, consuming eastern market (where due to supply shortages, prices were higher). During periods when demand shifted to the east, prices would go up. Koch Gateway wanted to be able to charge market based prices so that they could try to capture some of this basis differential between the lower priced places on the western part of their system and the higher priced places on the eastern part of their system.

Today, new pipelines in the east and the new supplies of Gulf Coast gas have eliminated this price differential. Now, it is truly one big market with nearly equivalent prices across the entire market. This is largely due to the aforementioned increase in supply. Koch Gateway's ability to price its transportation services to capture the vanishing basis price differentials has essentially been eliminated in the market.

The basis differential is exogenous to Koch Gateway's behavior in the market place and reflects the broad market competition in the natural gas commodity market. When the basis differential is large, a non-negligible arbitrage opportunity exists where shippers can move gas to capture the basis differential. If the cost of the transportation move exceeds the differential, no profit could be earned. Consequently, the basis differential reflects one constraint on a pipeline's ability to charge its full transportation rate of the commodity.

Once the basis differential is less than the transportation cost, opportunities for arbitrage vanish. This is what has happened across the Koch Gateway system. As the basis differential between western and eastern points on Koch Gateway's system has virtually disappeared, so too have arbitrage opportunities.

The Federal Energy Regulatory Commission (FERC) has recognized this collapse of the basis differential. In Order 637,¹ the Commission recognized that it is now the commodity markets that determine the economic value of gas pipeline transportation services. The Commission stated that "...the implicit price for transportation represents the most any shipper purchasing delivered gas in a downstream market would pay to move gas from the lower priced market to the higher priced market. For instance, the implicit value of transportation between the Henry Hub and the Chicago City gate market was \$.07 in September 1999 (the difference between the \$2.67 price for gas in Chicago and the \$2.60 price as the Henry Hub)."

For several receipt and delivery locations, I display recent basis differentials at hub locations in Figure 1.1. I selected these locations because Koch Gateway transportation volumes to power plants demonstrate that these receipt and delivery pairs are relevant. Omitted from the data in Figure 1.1 through are the 32.3 percent of volume that are transported on the Koch Gateway system with short hauls and for which the basis differential is exactly zero. In selecting these basis pricing differentials, I used the volumes moved on the Koch Gateway system. For instance, as shown in Figure 1.1.1, 19 percent of delivery volume occurred between receipt and delivery points where the relevant liquid basis points were the "Henry Hub" point in South Louisiana and the "Venice" point in South-East Louisiana.

¹Regulation of Short-Term Natural Gas Transportation Services, and Regulation of Interstate Natural Gas Transportation Services 90 FERC 61, 109 (February 9, 2000) Order 637.

Figure 1.1.1: Henry Hub Deliveries to Venice (19% of Delivery V olumes)

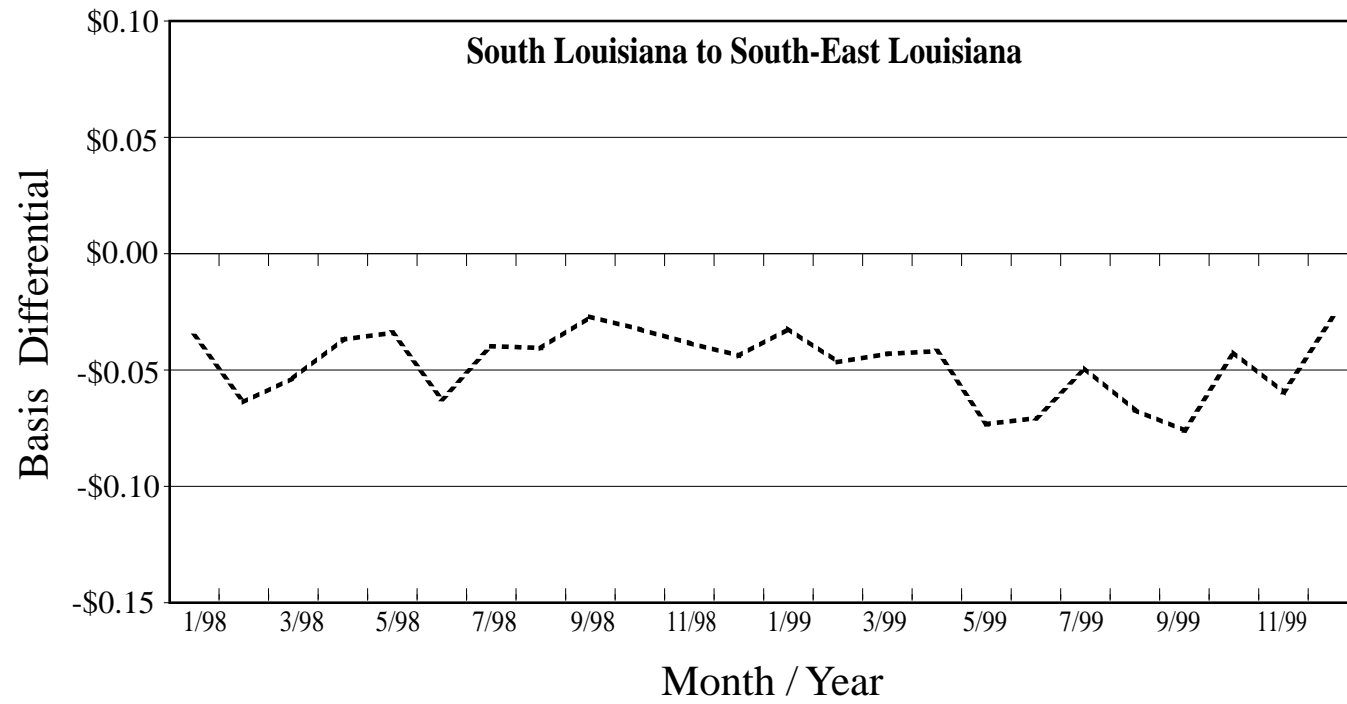


Figure 1.1.2: Henry Hub Deliveries to Kosciusko (12% of Delivery Volumes)

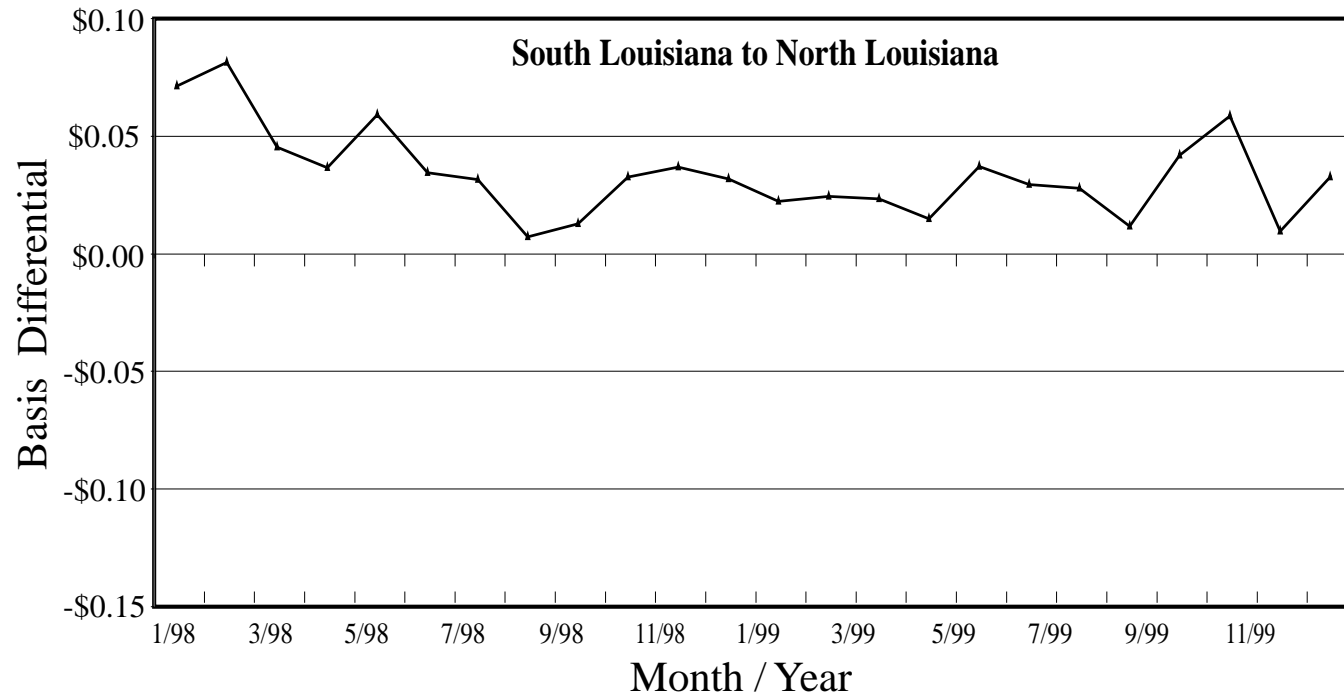


Figure 1.1.3: Venice Deliveries to Henry Hub (4% of Delivery Volumes)

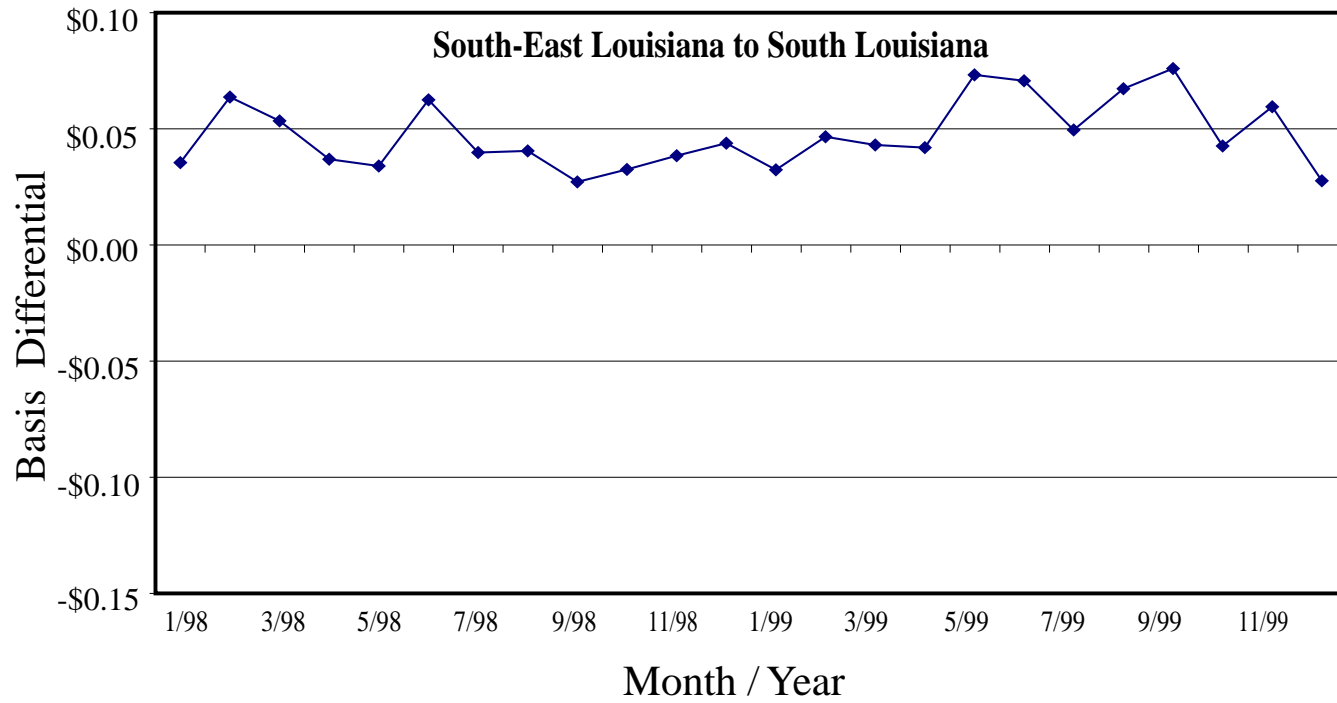


Figure 1.1.4: Kiln Deliveries to Venice (4% of Delivery Volumes)

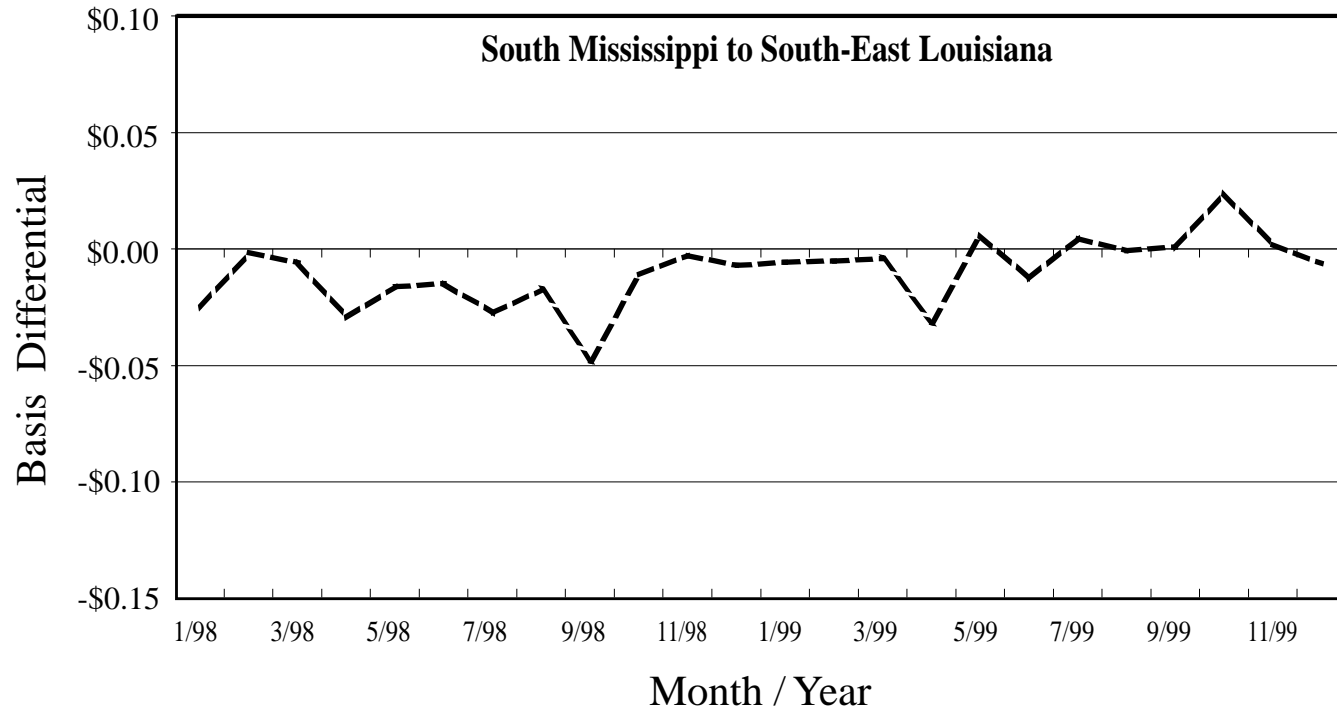
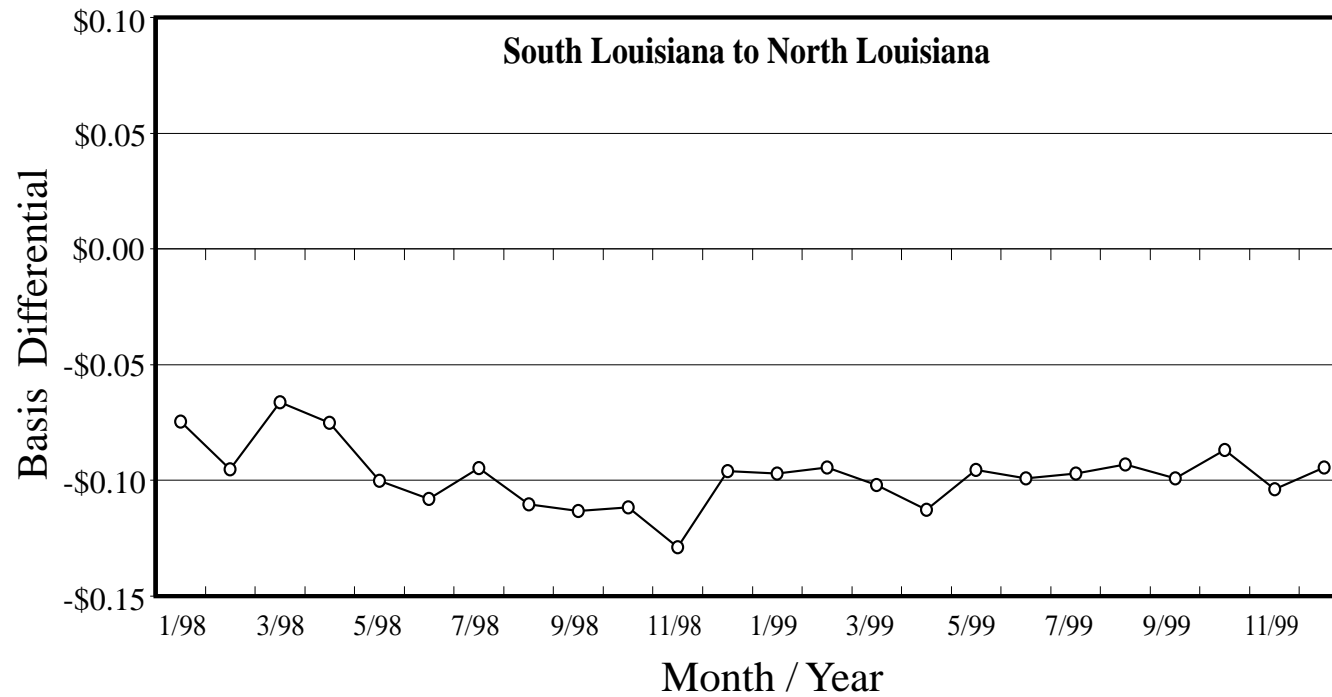


Figure 1.1.5: Henry Hub Deliveries to Koch Zone 2 (3% of Delivery Volumes)



I make two observations from these figures. First, the basis differentials have been declining over time. At present, the average differential is small and is often zero or negative. Additionally, the variance in the pricing differential is also declining to levels reflecting random difference rather than large systematic swings. Second, relative to the early 1990s, the basis differentials are much smaller. While differentials on the order of \$0.20 to \$0.30 were not uncommon in the early 1990s, the differentials are nearly zero at present.

My analysis, below, demonstrates that to the extent price differentials still exist on occasion, these price differentials have a statistically significant effect on Koch Gateway's discount setting behavior. However, as I will discuss later, as the basis differentials are nearly always negligible, there is a constraint on Koch Gateway's ability to charge maximum tariff rates for its transportation services. Below, I also discuss how this affects Koch Gateway's discounting decisions. Importantly, I conclude that the basis differential collapse is a significant factor in Koch Gateway's discounting behavior.

1.3. Discounting

Koch Gateway's decision to charge a rate for its transportation services that is lower than the full tariff rate is a decision that is governed by Koch Gateway and the market. I expect that Koch Gateway will discount its transportation rates because it is a price-taker rather than a price-setter. When Koch Gateway's shippers can effectively dictate prices, Koch Gateway lacks upstream market power. Conversely, when Koch Gateway is less likely to offer discounts or does not discount at all, whatever upstream market power Koch Gateway possesses is still constrained by its "just and reasonable" regulated tariff rate. Thus any potential market power is negated. Because discounting behavior will provide evidence as to whether Koch Gateway possesses upstream market power, I analyze Koch Gateway's upstream market power based on its historical discounting behavior.²

1.3.1. The Data Assembled

To investigate Koch Gateway's propensity to discount, I collected data on Koch Gateway's transactions to its power plant customers during 1998 and 1999. The data was summarized at the monthly level, and represents the simultaneous receipt and delivery of gas. The Koch Gateway transportation transaction data includes information on the delivery and receipt meters (Station Location Numbers or SLNs), including the meter's location, the volume shipped, the type of service, the shipper, and the actual customer costs. This

²The relationship between pricing (discounting) and the Herfindahl Hirschman Index (HHI) statistic is explored in Hannan (1997), Rhoades (1995), and in earlier studies such as Connor and Peterson (1992). A discussion of Koch's historical discounting behavior was provided in Dubin (1998).

dataset includes only volumes shipped through Koch Gateway connected to power plants for 1998 and 1999. The volumes transacted were aggregated at a monthly level. A total of 3,002 transactions are present in the transaction data. The dataset I analyzed has 144 separate receipt points, 16 delivery points (the power plants), and 24 months. Note that every delivery point did not receive gas from every receipt point.

The principal factors I considered in my analysis were the actual cost of the delivery represented in the monthly transaction, (the “contract rate”) and the undiscounted maximum tariff cost of the transaction. I added the transportation rate and the reservation rate together to determine the contract rate. I also added together the transportation and reservation rates allowed by the tariff to determine a tariff maximum rate. I derived the maximum tariff rate from Koch Gateway’s FERC tariff filings. This data includes the firm and interruptible service’s maximum allowable rates. I compared the contract rate to the tariff rate for the same time of the year, type of service and receipt and delivery zones. I calculated the discount by dividing the contract rate by the tariff rate and subtracting the result from one. This means that the discount value ranges from zero to one.

To explain the discount depth rather than the mere presence of a discount, I considered several factors. The first factor I consider is the basis differential between gas commodity prices at the index hubs nearest to the receipt point and delivery point. The basis point data provided a list of “liquid points” (i.e., basis locations) and the daily price at which gas traded at that point. The basis value at each point for the month is the average basis for that point. I chose the basis value for each receipt and delivery point based on the closest “liquid point” to the receipt-delivery point.

The next factor that I considered as a possible influence on discounting was prevailing weather at the time of the transportation pricing decision. Weather data was taken from the National Oceanic Atmospheric Agency (NOAA). This data consists of individual weather station daily heating and cooling degree-days, aggregated to a monthly heating and cooling degree-day value for each weather station. I selected the weather station for each power plant by locating the power plant on the Koch Gateway system map. I compared this location to a list of possible weather stations in the area and selected the station closest to the nearest large electrical load connected to the power plant. I measured both heating and cooling degree-days at the power plants as demand factors (i.e., larger heating degree-days imply colder season heating, thus increasing the demand for gas and the derived demand for gas transportation). Heating and cooling degree-days may be associated with above average demand, making it less likely that Koch Gateway would discount. I test this hypothesis in my econometric modeling reported below.

I also considered whether the transaction was firm or interruptible. Over 85.9 percent of Koch Gateway's transactions to power plants are on an interruptible basis (over 80.9 percent on a volume weighted basis). Given the contract provisions in long-term firm contracts, I expect interruptible transactions to be more heavily discounted than firm transactions; and, therefore, I test this hypothesis.

I also constructed a time trend variable to help identify any trends for Koch Gateway's discounting that are not discernible in the other factors. Next, I considered several measures of market competitiveness.

Connection and ownership data indicate the ownership of each plant and the number of competing pipelines that are connected or proximal to the power plant. The data also indicate the volume of gas shipped by Koch Gateway and others to the SLN for the period October 1998 to September 1999. I used this data to establish the number of connected alternatives and the number of close alternatives. I also used this data to create Koch Gateway's volume shares to each SLN. First, I considered whether the delivery point (i.e., a power plant) had an alternative to Koch Gateway within five miles and whether it was connected or not. Second, I considered the number of alternatives to Koch Gateway that were physically connected to a delivery point.

I also relied on Koch Gateway's 1996 survey of its delivery point SLNs and alternatives within five miles of each SLN to derive various concentration indices. Each record in the survey dataset corresponds to a single Koch Gateway SLN and contains information for as many as 15 alternative pipelines passing within five miles of the SLN. The survey dataset also contains Koch Gateway's diameter-squared flow capability at the SLN. Data for each alternative pipeline included: its company code; its diameter-squared flow capabilities; and its distance from the Koch Gateway SLN. There are 15 SLNs in the alternative pipeline data file. These represent all SLNs present in the transaction data set described earlier, less one power plant for which information was not available. The maximum number of alternative pipelines present at a single SLN is eight, the minimum is zero, and the average was just under four. I performed calculations to create variables measuring market competitiveness within each SLN's five-mile radius circle. These measures are based on market share and distance to nearby alternatives.

Using the distance data supplied by Koch Gateway in the alternative pipeline data file, I calculated variables measuring the minimum distance from each SLN to an alternative pipeline. The distance measurements represent the distance between the Koch Gateway SLN and each alternative pipeline within five miles.³ As an empirical matter, I test the hypothesis that the minimum distance

³The distance factor is relevant in the regression analysis only in those situations where Koch Gateway is not the only pipeline within five miles. In a situation where Koch Gateway is the only pipeline within

to alternative pipelines influences Koch Gateway's discounting, causing Koch Gateway to discount more when alternatives are closer.

Generally, I expect that markets where Koch Gateway has a large market share will be less competitive. In those situations, all other things equal, I expect that Koch Gateway will be less likely to discount. Hence, I hypothesized a negative relationship between Koch Gateway's market share and Koch Gateway's propensity to discount. In this case, an increase in Koch Gateway's market share at an SLN should decrease the likelihood that it will discount. This is because such a market type would be less competitive than others where Koch Gateway has a lower market share.

Other competitiveness measures are plausibly related to the propensity to discount. As noted above, I expected Koch Gateway's market share to be negatively related to its discounting probability. Similarly, the presence of more pipeline alternatives should equate to a higher probability that Koch Gateway will discount because Koch Gateway faces more competition in this particular market situation. I test whether connected pipelines or proximal pipelines which are not connected affect Koch Gateway's discount decision.

Finally, I constructed an indicator variable for situations where Koch Gateway controls 100 percent of an SLN level market. These are situations where there are no alternative pipelines within a five-mile radius of an operating Koch Gateway SLN. In these markets, I expect that Koch Gateway is less likely to discount its transportation services.

1.3.2. Analysis

There are two striking observations to be made initially about the Koch Gateway transactions. First, the average discount is 63.0 percent off maximum tariff rates in Koch Gateway power plant delivery transactions and 72.0 percent on a volume-weighted basis. Second, virtually all transactions receive some discount. Hence, discounting is a widespread and pervasive force in Koch Gateway's transactions with its power plant customers.

There are many reasons for this deep and pervasive discounting at power plants connected to Koch Gateway. These reasons include the many alternatives that the buyers have, multiple pipeline connections and buying points, the ability to negotiate deliveries globally for multiple plants, and a highly competitive upstream natural gas market. Below, I examine these and other factors that explain Koch Gateway's discounting behavior.

five miles, Koch Gateway has 100 percent of the capacity within a five-mile radius of the SLN. If, by a diameter-squared capacity market share measure, Koch Gateway's market share is not 100 percent, then another pipeline is within five miles and its distance to the SLN may influence Koch Gateway's discounting decision.

I summarize the transaction dataset for power plants in Table 1.1. In this table, I show the 16 delivery points (one for each power plant served by Koch Gateway). I also show Koch Gateway's volume, the total volume, Koch's market share and the number of physical and potential connections within five miles. As can be seen, during 1998/1999 Entergy was the primary power plant customer for Koch Gateway transportation services in this region.

Table 1.1. Alternative Pipelines by Power Plant

ENTERGY PLANTS					
Plant Name	Connected Alternatives	Close Alternatives	Koch Volume	Total Volume	Koch Share
Willow Glen	5	3	20,200,211	50,527,512	40.0%
Nine Mile Point	4	0	21,373,426	77,280,892	27.7%
Rex Brown	1	1	5,354,829	9,045,239	59.2%
Baxter Wilson	1	0	20,193,497	24,211,702	83.4%
Roy Nelson	5	2	31,968	28,896,960	0.1%
Little Gypsy	2	9	5,666,559	33,255,750	17.0%
Waterford	4	2	6,592,399	27,980,136	23.6%
Sterlington	2	7	7,556,036	12,225,171	61.8%
Michoud	4	1	18,809,138	34,016,211	55.3%

SOUTHERN COMPANY PLANTS					
Plant Name	Connected Alternatives	Close Alternatives	Koch Volume	Total Volume	Koch Share
Crist/Pensacola	0	0	3,795,160	3,795,160	100.0%
Eaton	0	2	2,680,700	2,680,700	100.0%
Jack Watson	0	0	8,349,519	8,384,776	99.6%

SWEPCO PLANTS					
Plant Name	Connected Alternatives	Close Alternatives	Koch Volume	Total Volume	Koch Share
Knox Lee	3	3	5,857,381	12,052,333	48.6%
Pirkey	2	2	154,482	154,482	100.0%

SMEPA PLANTS					
Plant Name	Connected Alternatives	Close Alternatives	Koch Volume	Total Volume	Koch Share
Benndale	0	0	50,013	50,013	100.0%
Moselle	1	0	4,159,428	8,130,180	51.2%

The number of connected alternatives ranges from one to five for Entergy plants, none for Southern plants, two to three for SWEPCO plants and zero to one for SMEPA plants. The number of additional close alternatives ranges from zero to nine for Entergy plants, zero to two for Southern plants, two to three for SWEPCO plants, and none for SMEPA plants. In total, there are between one and 11 alternatives to Koch Gateway at Entergy plants, zero to two alternatives to Koch Gateway at Southern plants, four to six alternatives

to Koch Gateway at SWEPCO plants and zero to one alternatives to Koch Gateway at SMEPA plants. Three plants (SMEPA–Benndale, Southern–Crist Pensacola, and Southern–Jack Watson) have no alternatives to Koch Gateway within five miles, whether connected or not. I compared these figures to the Koch Gateway 1996 survey data. This comparison reveals an increase in the number of alternatives at delivery SLNs (i.e., the number of alternatives has been growing.)

As shown in Table 1.1, Entergy owns nine power plants that are connected to Koch Gateway. All are physically connected to alternative pipelines and all but two have at least one additional pipeline close by. To the extent that it supplies gas transportation to these power plants, Koch Gateway discounts its transactions with Entergy aggressively to remain competitive.

Southern has three plants that are currently connected to Koch Gateway. None are connected to alternative pipelines and only one has nearby unconnected alternatives. Koch Gateway’s historical discounting behavior was not as aggressive as it was with Entergy, SWEPCO and SMEPA. Similarly, Koch Gateway provides transportation to two SWEPCO plants. Both have at least two other pipeline connections and several nearby unconnected alternatives. As with the Entergy plants, Koch discounts aggressively to SWEPCO. Finally, SMEPA has two power plants connected to Koch Gateway. Although one has no connected alternate pipeline and neither have any nearby alternatives, Koch Gateway discounts its transportation to both plants.

1.3.3. Transaction Data Analysis

The transaction data I analyze provides a compelling view how the Koch Gateway is used. For example, I examined Koch Gateway’s delivered and interruptible transportation volume. I present this graphically in Figure 1.2. As can be seen, the volumes shipped on Koch Gateway swing markedly. This demonstrates the seasonal and peaking nature of the demand for Koch Gateway’s services. This characterizes the way Koch Gateway is used as a swing pipeline by the power plants when demand is seasonally increased.

I also analyzed the weighted average transportation rate charged by Koch Gateway. These transportation rates are relatively small. Unlike the volumes transported, the weighted average transportation rates do not vary greatly. I show the weighted average transportation rates in Figure 1.3. Next, I considered and analyzed the volume weighted discounts that Koch Gateway gave on its interruptible transportation transactions over the same two-year period. On a volume-weighted basis, the average volume weighted discount provided by Koch Gateway was over 72.0 percent. In fact, as Figure 1.4 demonstrates, this volume weighted average discount for interruptible transportation never falls below 55 percent during the two-year period I analyzed. And, as we have seen, some plants receive considerably larger discounts than this average discount.

Figure 1.2: Delivered Volumes by Type of Service

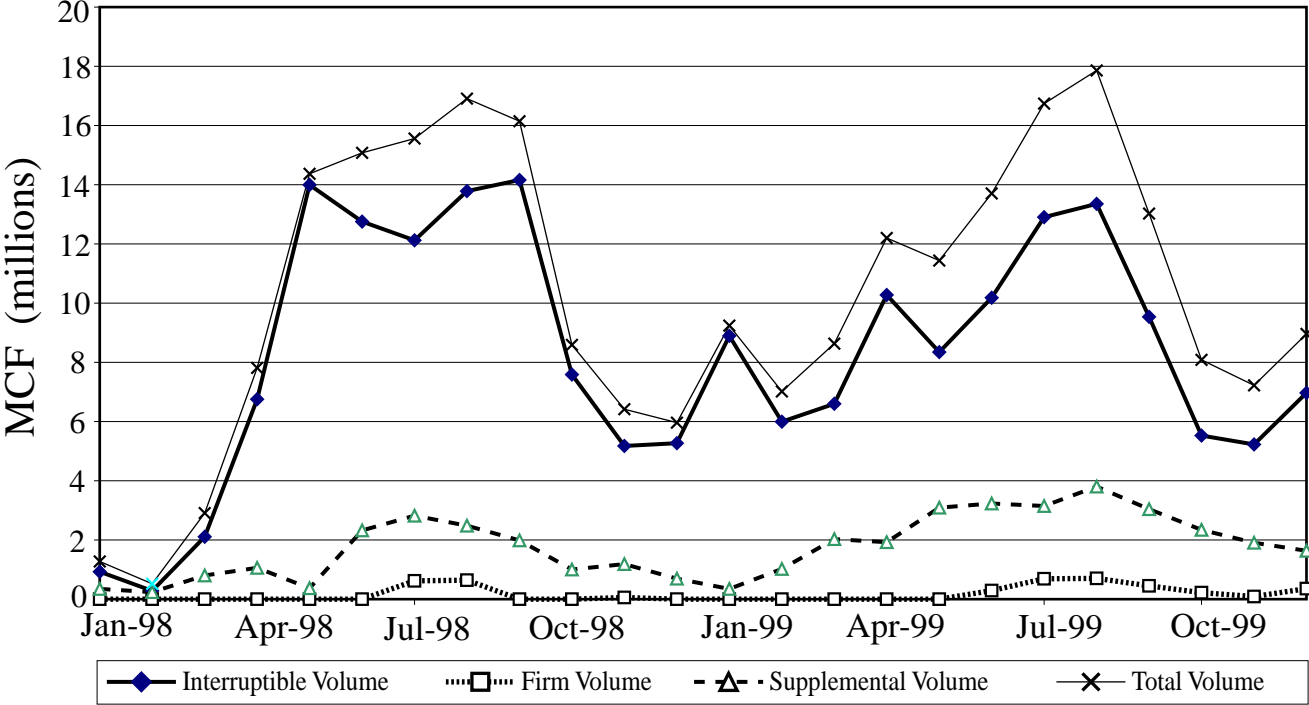


Figure 1.3: Weighted Average Transportation Rate

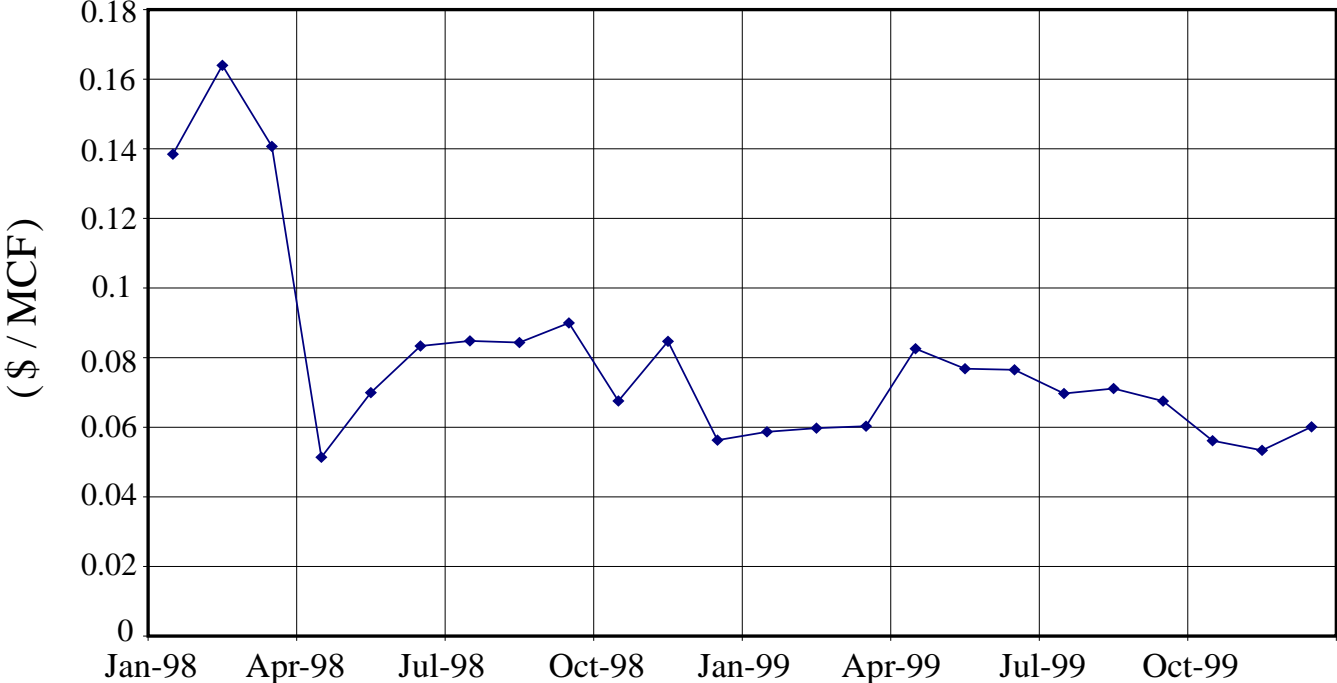


Figure 1.4: Volume Weighted Average Discount for All Transactions

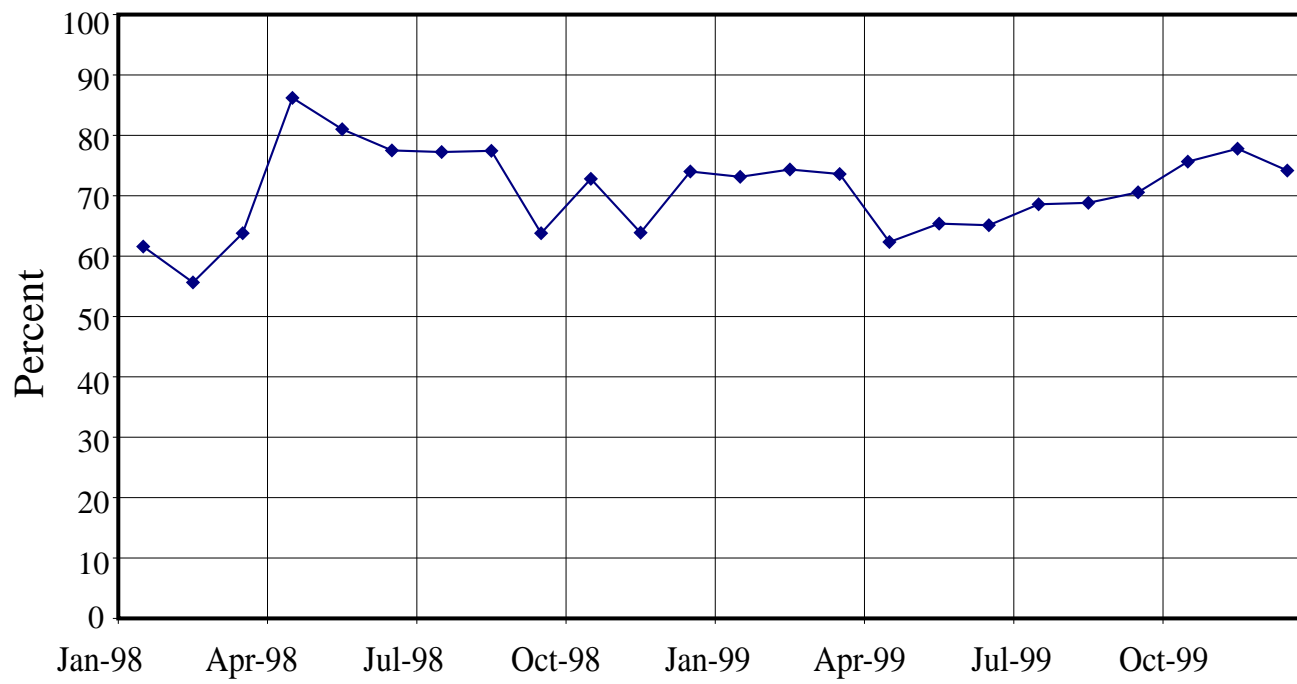


Figure 1.5: Weighted Average Monthly Receipt and Delivery Basis Value

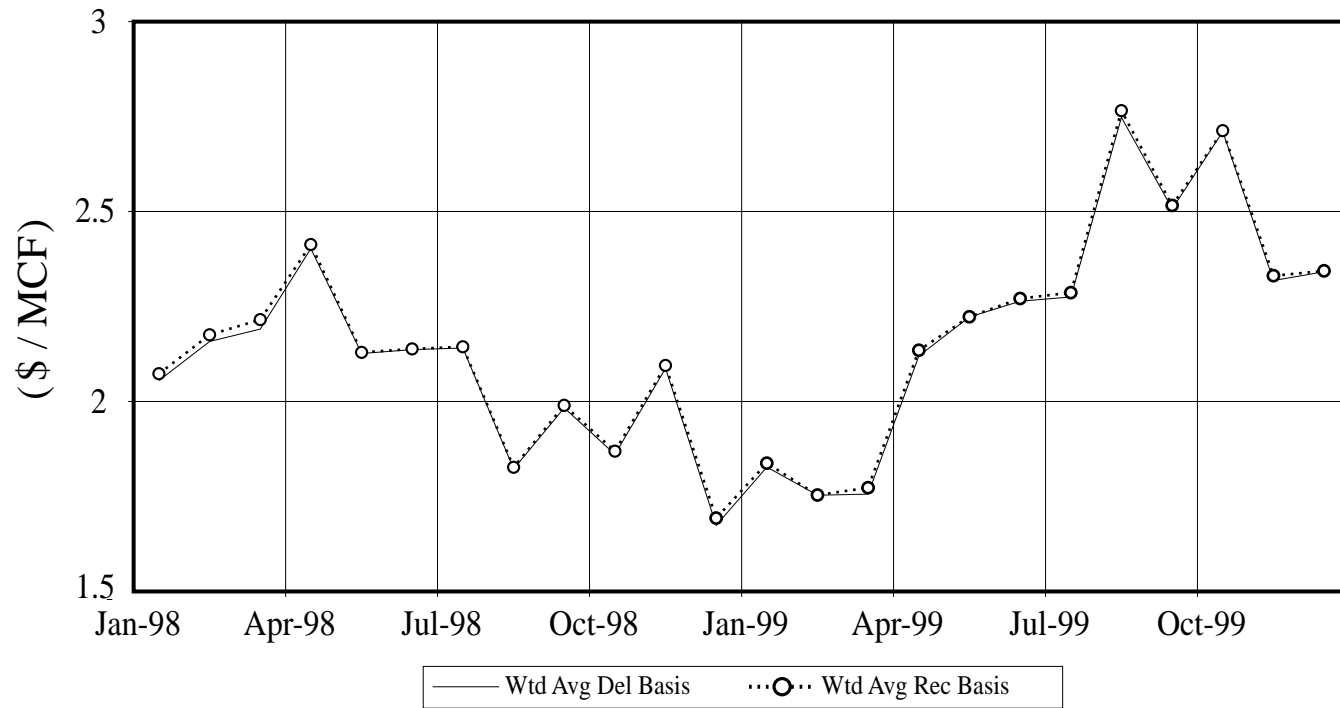
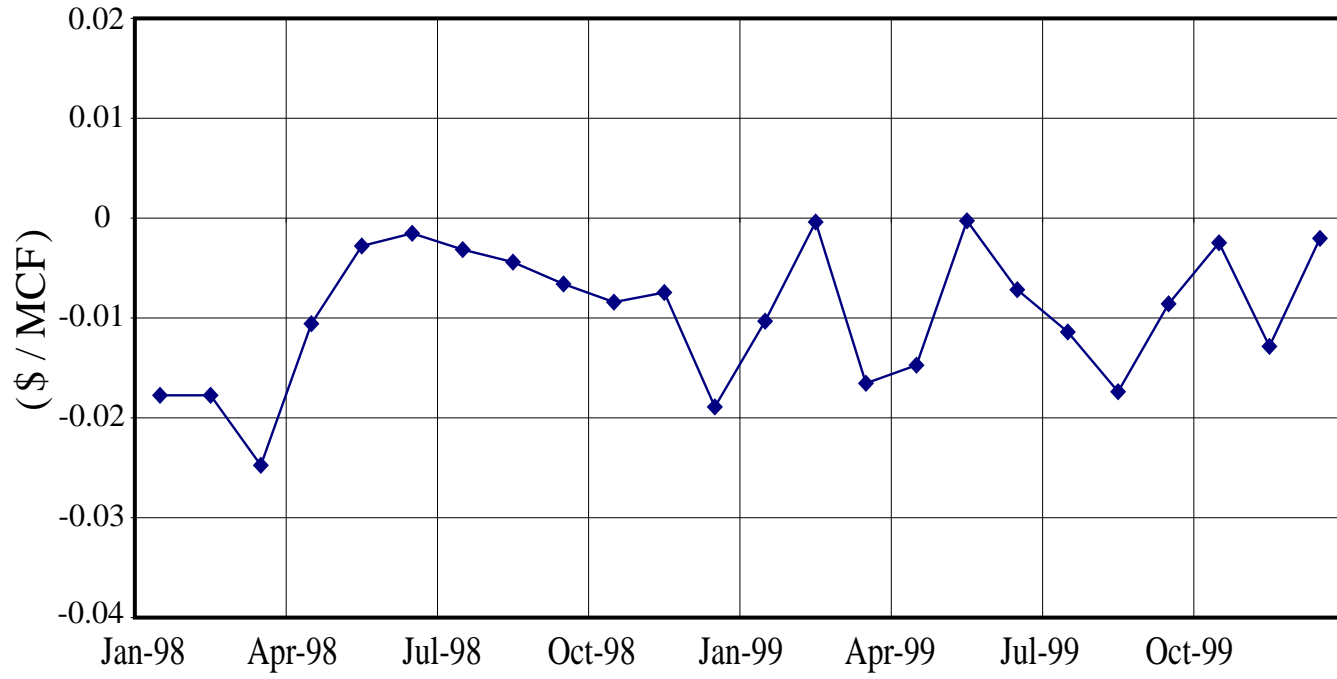


Figure 1.6: Weighted Average Monthly Basis Differential



Next, because the basis differential will, as I discussed earlier, often constrain Koch Gateway's ability to increase transportation prices, I examined monthly receipt point and delivery point prices. I did this on a weighted average monthly basis. As can be seen in Figure 1.5, the lines denoting prices at delivery points lie extremely close to the lines denoting prices at receipt points. In some cases, it appears as if there is only one line. This means that the prices at receipt points and delivery points are almost identical. Because there is virtually no market spread in the commodity price, Koch Gateway's ability to exercise market power in delivering natural gas is virtually non-existent. There are therefore few opportunities for arbitrage.

However, the basis differentials established at liquid points near the actual receipt and delivery point may not accurately reflect the commodity's actual price differential at receipt and delivery points, especially those that are somewhat distant from the liquid basis points. As an empirical matter, it is plausible that local basis price differentials will indicate both pricing constraints and the directions that shippers will generally be moving gas. While I don't expect a perfect correlation between basis differentials and Koch Gateway's discounting, I nevertheless test the hypothesis that these differentials matter empirically in determining Koch Gateway's discounting behavior.

In Figure 1.5, the index price at the nearest receipt liquid point receives the weight due to the volume transported in the transaction, while the nearest delivery liquid point receives the weight due to the volume in the same transaction. To take an extreme example, if ninety percent of Koch Gateway's transportation volume went from receipt point A to delivery point B then the graphs in Figure 1.5 would largely reflect the commodity prices at points A and B. It is precisely this fact that explains how the weighted average combination of negative and positive basis differentials shown in Figure 1.1 produce the negative differentials shown in Figure 1.5. A larger percentage of the volume flows in situations where the prevailing basis differentials are negative than in the situations where the basis differentials are positive. While aggregated data series illustrate average conditions, the regression analysis reported below utilizes the actual basis differential prevailing in each transaction at each time period. This procedure avoids any aggregation bias that might otherwise result by combining large and small transactions.

I also considered basis differentials weighted by volume. I show these results in Figure 1.6. While there are some variations in the basis differentials over time, these differentials are extremely small. As can be seen in Figure 1.6, the differentials are well under two cents in most periods. Examining the basis differential and index pricing shown in Figures 1.5 and 1.6, demonstrates that on a volume weighted basis, Koch Gateway transportation often flows against a negative weighted average basis price differential (i.e., the receipt basis price is higher than the delivery basis price). A negative spread indicates that ship-

pers will demand large discounts to move gas in largely uneconomic directions. Clearly, some of the distinction between negative basis differentials and Koch Gateway's actual behavior is explicable in that the liquid points are not perfect depictions of local conditions that prevail at specific receipt and delivery points.

The primary and most important point revealed here is that the natural gas market in which Koch Gateway operates is highly competitive. Koch Gateway does not have an opportunity for arbitrage and consequently cannot exercise market power. In these situations, Koch Gateway must negotiate discounts and settle for market-determined rates that are substantively below the "just and reasonable" regulatory cost of service rates.

1.4. Econometric Results

In this section, I consider econometric models of Koch Gateway's discounting based on several competitive measures, including the minimum distance to a competitor. I expect that as the distance to a competitor increases, Koch Gateway should be less likely to discount. Additionally, I consider models based solely on the number of alternatives physically or potentially connected to the power plant Koch Gateway serves. I expect that alternatives constrain Koch Gateway's pricing behavior. Below, I test the hypothesis that physical or close connections affect discounting. Hence, I test whether close connections (within five miles and unconnected) are sufficient to constrain Koch Gateway's discounting behavior.

I estimated regression models of Koch Gateway's discounting behavior with two dependent variables. In the first version, I use a regression model with the discount variable constructed from the transportation transaction without adjustment. In the second version, I treat all discounts of less than one percent as equivalent to zero discounts. This latter approach eliminates very small discounts as accounting irregularities. In the Koch Gateway transaction dataset, 193 transactions, or 6.43 percent of transactions, have discounts less than one percent. Thus, recoding these observations creates a mass point at the zero discount value. Even though these small discounts account for only 4.0 percent at the volume, it is appropriate in such a situation to use a Tobit model. My empirical investigation showed that Tobit models, regression models using adjusted discounts, and regression models with unadjusted discounts produced quantitatively identical results. The variables used in the regression analysis are defined in Table 1.2. The discount regression models are displayed in Tables 1.3.1–1.3.8.

In Table 1.3.1, I specify the discount to depend on: the basis differential; heating degree days; a trend term; whether the transaction was interruptible; the delivered quantity; Koch Gateway's market share at the delivery point; the number of close alternatives (other than Koch Gateway); and, an indicator for

Table 1.2. Variable Glossary

bas_diff	Basis differential between receipt and delivery
del_qty	Delivered quantity (MMBTU)
disc	Discount percentage
entergy	Entergy Power Plant
hdegday	Heating degree day
interup	Interruptible transportation
koch_shr	Koch Gateway's market share
koch_100	No alternative pipelines within 5-mile radius at delivery point
ksln_shr	Koch Gateway's market share at an SLN
min_dis	Distance to nearest alternative pipeline at delivery point
mswhhi	Market share weighted shipper Herfindahl
num_alt	Number of alternatives (connected or unconnected within 5 mile radius of SLN)
num_clos	Alternatives within 5 miles of power plant whether connected or not
num_cnxn	Number of physically connected alternatives
rkoch_100	No alternative pipelines within 5-mile radius of receipt SLN
rmin_dis	Distance to nearest alternative pipeline at receipt SLN
smepa	SMEPA Power Plant
southern	Southern Company Power Plant
supplem	Supplemental transactions type
swepco	SWEPCO Power Plant
trend	Time trend

Entergy plants. In Table 1.3.1 Koch Gateway's discounting behavior at Entergy power plants is compared to Southern, SWEPCO, and SMEPA power plants. The power plant indicators were nearly all statistically significant.

The regression model shows that as the basis differential increases, the magnitude of the discount falls. Hence, the regression model demonstrates that basis differentials constrain Koch Gateway's ability to set prices.⁴

⁴As competition increases in the upstream natural gas market and new pipelines enter the market, the basis differentials have been declining in this region over time, with the average at approximately -0.01 cents. The minimum and maximum basis differentials have reached levels of -14 to $+12$ cents in this period. These large differentials occur relatively infrequently and are uncharacteristic of the system on a volume-weighted basis.

Table 1.3.1. Transportation Rate Model 1

Dependent Variable: disc		
Variable	Coefficient	<i>t</i> -Stat
(1)	0.442	17.39
bas_diff	-0.154	-1.83
hdegday	0.000	1.93
trend	-0.003	-5.59
interup	0.148	14.59
del_qty	0.000	11.25
koch_shr	-0.231	-9.75
num_clos	0.002	1.16
num_cnxn	0.020	5.60
entergy	0.204	17.70
Observations	3002	
<i>R</i> -Squared	0.48	
Mean Discount	0.63	

Koch Gateway's discounts on its transportation transactions move opposite to the basis differentials. As the basis differential widens, Koch Gateway discounted less. At one extreme, where the basis differentials are nearly +12 cents, the discount will fall by approximately 5.1 percent.

Heating degree-days are positively associated with discounting. This reflects a volumetric discount in demand rather than a discount due to increased market power when demand is more inelastic. Similarly, delivered quantity is positively related to discounting. This reflects the fact that Koch Gateway is willing to offer quantity discounts and is willing to encourage larger transactions by lowering prices. This is consistent with a more elastic demand for larger shippers with more choices.

Interruptible transactions are more heavily discounted than firm transactions because most shippers taking this service have substitute fuel services. An interruptible transaction has approximately a 15 percent higher discount amount than does a non-interruptible transaction.

Turning to competition measures, I find that a greater number of connected alternatives (other than Koch Gateway) increases the discount amount. However, the number of close (less than five miles and currently unconnected) connections is not a statistically significant factor. Having tested the hypothesis that either the number of connected alternatives or the number of proximal

alternatives matters, I find that the number of connected alternatives acts as a constraint. However, the number of proximal unconnected alternatives does not affect discounting in a statistically significant manner. Finally, as Koch Gateway's market share increases, its propensity to discount falls. Koch Gateway's market power, measured by its actual discounting, is constrained by its competitors. As the market share (and market power) of Koch Gateway's competitors increases, Koch Gateway must meet the competition by lowering its price.

In Table 1.3.2, I consider Koch Gateway's discounting behavior but distinguish the discounts among the plants owned by Southern Company, SMEPA and SWEPCO. In Table 1.3.2, the basis differential effect is even larger than in Table 1.3.1. Heating degree-days are no longer statistically significant but are still positively associated with discounting. The estimated coefficients show that, in comparison to Entergy, Southern plants receive lower discounts (by approximately 35 percent), SMEPA plants receive higher discounts (by approximately 5.2 percent), and SWEPCO plants receive 11.8 percent lower discounts. The regression uses Entergy as the comparison plant.

Table 1.3.2. Transportation Rate Model 2

Dependent Variable: disc		
Variable	Coefficient	<i>t</i> -Stat
(1)	0.570	23.50
bas_diff	-0.430	-5.31
hdegday	0.000	0.40
trend	-0.004	-7.65
interup	0.155	16.05
del_qty	0.000	11.39
koch_shr	-0.078	-3.30
num_clos	0.003	1.80
num_cnxn	0.023	6.53
southern	-0.349	-26.02
sempa	0.052	2.66
swepco	-0.119	-5.98
Observations	3002	
<i>R</i> -Squared	0.55	
Mean Discount	0.63	

In contrast to the situation at SMEPA or SWEPCO, Koch Gateway is much more often the only supplier at Southern's plants. (The SMEPA Benndale plant is the only other case where there is not another physical connection).

Koch Gateway is never the sole supplier to Entergy plants. Regardless of the apparent lack of competition and even though Koch Gateway is more often the sole supplier at Southern plants, it was historically only able to reduce its discount by 35 percent (relative to its discounts at Entergy power plants). Historically Southern received an average 23 percent (volume weighted) discount in its transactions with Koch Gateway. This is considerably lower than the Koch Gateway system average discount to power plants, but well below the maximum tariff.

In Table 1.3.3, I again consider Southern, SMEPA and SWEPCO power plants relative to Entergy plants. In this model, I use the minimum distance variable in place of the number of connections (connected or unconnected) to measure competition. I find that as the minimum distance to an alternative increases, Koch Gateway is less likely to discount. The potential for physical connection is thus important and measurable. This effect was not discernible through the number of alternative potential connections but is revealed by the distance factor. (the distance to connect and the size of the pipe are the two main components of the connection cost.) A smaller distance to a potential competitor within five miles has a statistically significant effect on Koch Gateway’s discounting behavior. Here, the mere presence of unconnected pipeline alternatives affects Koch Gateway’s pricing for transportation service. In other words, potential competitors are in the same market with Koch Gateway because the barriers to entry are not high if the competition is physically close.

Table 1.3.3. Transportation Rate Model 3

Dependent Variable: disc		
Variable	Coefficient	t-Stat
(1)	0.660	52.27
bas_diff	-0.541	-7.26
hdegday	0.000	0.11
trend	-0.005	-9.17
interup	0.144	14.01
del_qty	0.000	10.10
(min_dis&!koch_100)	-0.000	-9.54
southern	-0.495	-53.22
smepa	-0.047	-2.54
swepco	-0.120	-6.12
Observations		2666
R-Squared		0.56
Mean Discount		0.62

In Table 1.3.4, I rely on information from Koch Gateway's survey of alternatives and use two specific competition factors available in this dataset. The first factor measures the number of alternatives at the delivery SLN. The second factor represents Koch Gateway's capacity market share at the delivery SLN. This model is based on somewhat fewer transactions since one power plant is not represented in the alternative pipeline dataset. Table 1.3.4 parallels my analysis in Table 1.3.2, except that it relies on the SLN alternative data. Here I find that as Koch Gateway's market share increases, or as the number of alternatives decreases, Koch Gateway is less likely to discount its transactions.

Table 1.3.4. Transportation Rate Model 4

Dependent Variable: disc		
Variable	Coefficient	<i>t</i> -Stat
(1)	0.635	40.61
bas_diff	-0.360	-4.34
hdegday	-0.000	-0.29
trend	-0.006	-10.00
interup	0.139	13.47
del_qty	0.000	10.55
ksln_shr	-0.001	-4.45
num_alt	0.006	3.68
southern	-0.415	-37.06
smepa	0.049	2.41
swepco	-0.137	-6.97
Observations	2666	
<i>R</i> -Squared	0.55	
Mean Discount	0.62	

The models considered, thus far, have focused on potential competition at the delivery point since the power plant is the originator of the transaction. In providing gas to a power plant, shippers have many alternatives. Presumably, shippers choose the path for which their profitability is largest. Since the transportation component of the delivered product is only a small part of its cost, a shipper may select a receipt point where, in theory, Koch Gateway could exercise some market power. To capture the potential market power in the transaction due to differentials in competition at the receipt point, I relied again on Koch Gateway's survey information. Specifically, I paralleled my analysis for the delivery point SLN by calculating the minimum distance to alternative pipelines within the five mile radius of the receipt point SLN. Specifying a regression model (Table 1.3.5) for Koch Gateway's discount decision shows that

increases in the minimum distance to a competitor pipeline within five miles of the receipt point have a negative partial correlation with discounting.

Table 1.3.5. Transportation Rate Model 5

Dependent Variable: disc		
Variable	Coefficient	<i>t</i> -Stat
(1)	0.873	41.99
bas_diff	0.055	0.41
hdegday	-0.000	-4.13
trend	-0.009	-10.01
interup	-0.009	-0.60
del_qty	0.000	3.08
(min_dis&!koch_100)	-0.000	-3.25
(rmin_dis&!rkoch100)	-0.000	-12.99
southern	-0.470	-30.03
smepa	-0.076	-2.54
swepco	-0.063	-2.58
Observations		748
<i>R</i> -Squared		0.66
Mean Discount		0.58

Thus as potential competitors to Koch Gateway became more distant at either the receipt point or the delivery point, Koch Gateway will be more likely to discount, all other factors equal. Conversely, as competitors' distance to Koch Gateway receipt and delivery SLNs decreased, Koch Gateway was more likely to discount.⁵

Shippers on Koch Gateway's system have their own market power. Potentially, the market power of shippers will affect Koch Gateway's discounting decisions. For instance, if Koch Gateway faces the same shipper in multiple transactions on its system, then the pricing decision it makes at one point in time with respect to a shipper could influence the shipper's decisions at other points on Koch Gateway's system or at other times. Thus, the profit maximization problem faced by Koch Gateway has intertemporal and spatial dimensions.

To examine the hypothesis of shipper market power, I calculated the Herfindahl Hirschman Index (HHI) for all shippers faced by Koch Gateway on a monthly basis, based on their transportation volumes. The HHI is a widely

⁵It should be noted that the survey data could only be matched in approximately 25 percent of Koch Gateway's receipt SLNs. Therefore, while these results are suggestive, they are not definitive. Importantly, none of my previous conclusions were altered in this alternative regression.

accepted method to measure market concentration. It involves first determining the relevant geographic and product markets, market participants, and each participant's market share. The HHI is calculated by squaring the market share possessed by each participant firm selling a particular good in a specific, well-defined geographic market and then summing the squares across all firms in the market.⁶

The shipper HHI is 1403 for those shippers using the Koch Gateway system to provide gas deliveries to power plants. To examine the influence of the shipper concentration on Koch Gateway's transaction decision at a single SLN, I considered a market share weighted HHI statistic for the shipper, defined as the product of the shippers market share of gas transported at the delivery SLN times the overall shipper HHI. Including this factor in the regression analysis (Table 1.3.6) indicated a positive, yet statistically insignificant, influence on Koch Gateway's discount.

Table 1.3.6. Transportation Rate Model 6

Dependent Variable: disc		
Variable	Coefficient	<i>t</i> -Stat
(1)	0.562	22.40
bas_diff	-0.419	-5.09
hdegday	0.000	-0.01
trend	-0.004	-6.59
interup	0.154	15.76
del_qty	0.000	11.23
koch_shr	-0.077	-3.24
num_clos	0.003	1.83
num_cnxn	0.023	6.55
southern	-0.354	-26.08
smepa	0.038	1.80
swepco	-0.117	-5.89
mshwhi	0.000	0.80
Observations	2952	
<i>R</i> -Squared	0.55	
Mean Discount	0.63	

After eliminating all supplemental firm transactions from the transaction dataset, I find that the market share weighted shipper HHI is statistically significant (Table 1.3.7). More generally, I find that Koch Gateway discounts its

⁶For further discussion see Dubin (1998).

supplemental transactions less aggressively than other firm transactions (Table 1.3.8).

Table 1.3.7. Transportation Rate Model 7

Dependent Variable: disc		
Variable	Coefficient	t-Stat
(1)	0.712	17.85
bas_diff	-0.453	-5.67
hdegday	0.000	1.35
trend	-0.006	-11.58
interup	0.049	1.50
del_qty	0.000	8.70
koch_shr	-0.078	-3.35
num_clos	0.002	1.06
num_cnxn	0.021	6.03
southern	-0.399	-30.77
smepa	0.115	4.80
swepco	-0.123	-6.80
mshwhi	0.000	1.92
Observations		3558
R-Squared		0.62
Mean Discount		0.65

Table 1.3.8. Transportation Rate Model 8

Dependent Variable: disc		
Variable	Coefficient	<i>t</i> -Stat
(1)	0.650	15.26
bas_diff	-0.433	-5.35
hdegday	0.000	0.38
trend	-0.004	-7.77
interup	0.075	2.08
del_qty	0.000	11.31
koch_shr	-0.078	-3.33
num_clos	0.003	1.84
num_cnxn	0.023	6.60
southern	-0.350	-26.10
sempa	0.055	2.79
swepco	-0.119	-5.98
supplem	-0.085	-2.29
Observations	3002	
<i>R</i> -Squared	0.55	
Mean Discount	0.63	

Hence, shippers exert their market power primarily in interruptible transactions. This alternative regression model reveals that shippers have market power on the Koch Gateway system at specific points and use that market power to make Koch Gateway take lower prices for its transportation services.

1.5. Conclusion

My analysis finds that market forces such as the prevailing basis differentials, quantity discounts reflecting elastic demand, and the number of connected alternatives all play significant roles in determining Koch Gateway's discount behavior. These factors all indicate competitive or workably competitive market conditions that influence Koch Gateway's discounting decisions.

The trend in basis differentials in the region in which Koch Gateway operates and the number of alternatives to Koch Gateway at the SMEPA, Southern and SWEPCO power plants are the key factors that govern Koch Gateway's discounting decision. Additional factors, that cause Koch Gateway to discount its transportation services, include competition, the collapse of the basis differential and the concomitant loss of arbitrage opportunities, buyer market power, and elastic demand.

Chapter 2

THE DEMAND FOR NFL FOOTBALL

2.1. Introduction

This chapter considers the demand for NFL football. Football demand is unusual as compared with many goods and services. Rather than market clearing at competitive price levels or limited demand induced by monopoly pricing, the equilibrium position for many football teams is one of excess demand. Football attendance is characterized by excess demand and generally tight market. A theory of Becker (1991) helps explain the demand for football. Becker's theory is that demand in some situations depends on social interaction and the size of the crowd. DeSerpa (1994) has adapted this theory for NFL football and other situations where the crowd composition is as important as the crowd size. Using data from 1995 through 1999 for all NFL teams during their regular season, I construct an econometric model of the demand for NFL football. I use this model to test the Becker/DeSerpa theory and conclude that the demand curve slopes upward in the relevant range as anticipated by the theoretical model. The next section of this chapter discuss NFL football demand. In Section 2.3, I discuss bandwagons, social influences, and group demand behavior. In Section 2.4, I discuss the Becker model and DeSerpa's extensions. Section 2.5 presents the econometric model while Section 2.6 provides conclusions.

2.2. The Demand for NFL Football Tickets

For many products, sellouts and excess demand are the prevailing phenomena. Evidence of excess demand includes season ticket programs, personal seat license contracts (PSL), near capacity crowds, and "sellouts." In the context of NFL ticket demand, these factors indicate tight markets and are consistent with

the theoretical notion of excess demand.¹ For instance, for many NFL teams, fans cannot obtain a season ticket directly from the team. Similarly, it is often not possible to directly obtain a playoff ticket to see these teams play in the post-season. In such situations, aftermarkets develop for providing tickets to eager fans at prices well in excess of the ticket's face value.

This pattern of excess demand is not limited to football. It is also seen in rock concerts, popular restaurants, and Broadway plays. In situations with excess demand (i.e., where demand exceeds supply) the price of the good in question should rise until the excess demand is eliminated. However, prices rarely rise to the point where only those most willing to pay gain admission. This anomaly can be explained.

First, the outcomes of sporting events are uncertain. Many factors affect a team's success. For example, when an NFL team has an unsuccessful season, it is awarded a higher draft choice, raising its chances to improve its personnel, and perhaps improve its winning percentage. Conversely, when an NFL team has a successful season, it is "rewarded" with a more difficult schedule in the next season. These choices by the league are clear attempts to balance competition and create parity among the teams. Additionally, injuries or plain luck can affect a team's performance in any given year. These qualities contribute to a situation that is characterized by considerable uncertainty in how consumers value individual games. Some games will be relatively low-demand games while others are likely to attract more fans, and thus have high demand. In the football market, fans are given the opportunity to buy a season worth of tickets at one time. In principal, a season ticket package sells for the number of games in a season times the face value of an individual game ticket. Consumers can then choose which games to attend based on their reservation valuation. In other words, when fans perceive that their reservation value exceeds the price already paid for the ticket, they will attend the game. Conversely, when fans perceive that their reservation value is less than the amount they paid for the ticket, they may choose to skip a particular game. Given the transaction costs associated with reselling a ticket; the restrictions on reselling tickets that exist in some markets; and the possibility that a consumer will not attend every game in the season, consumers generally place less value on the full set of tickets than they do for each game purchased at face value (as evaluated on a game-by-game basis for only those games they attend).

On the other hand, purchasing season tickets provides an individual with some option value. First, in markets where demand exceeds supply, the only way to guarantee admission to a game is by owning a season ticket. Second, the season ticket provides an option to enjoy games that may rise in consumer

¹The PSL is an upfront one-time charge placed on top of the season ticket price to guarantee the right to purchase the same or better season ticket for some period of time.

value during the season. Third, the season ticket provides an option to purchase post-season tickets for playoff games. The fee that management charges for post-season tickets may price a consumer out of the market, but the option value remains. On average, a season ticket is valued by the consumer and priced by management to reflect the individual game ticket price, the likelihood of post-season play, the price differential between the market and face value of a post-season ticket, and the transaction costs associated with reselling or purchasing tickets in the market during the regular season. With considerable uncertainty about any given team's success, sellers face a complex market in setting season ticket pricing policies.

Management can adopt an exploitative position and charge what the market would bear in any given season. However, if a team performs poorly, the season ticket price would have to be lowered the following year to ensure demand. The cost to management of exploitation is the loss of stable demand (i.e., large fluctuation in attendance and fan loyalty). Consumers, on the other hand, are willing to pay a price to ensure access to seasons when the team plays well, and suffer through the seasons when the team does not meet performance expectations.

The value of stable demand to management reduces its incentive to change season ticket prices often (i.e., season to season). Rather, in order to ensure stable demand, management might charge a price that reflects the consumers' option value of guaranteed access to regular and post-season play. Furthermore, management might offer a discount to consumers to ensure stability in demand. These factors help explain the fact that season ticket prices do not rise to clear excess demand in the short term. Specifically, excess demand in the short term may be followed by excess supply in a subsequent period. Consumers will pay a price for season tickets that reflects both the good times and the bad times.

However, the foregoing explanation of ticket pricing does not demonstrate why persistent excess demand is often the rule rather than the exception in professional football. It does explain why management prefers stable pricing and favors uniform pricing during the season, with only moderately increased prices for post-season play. Nonetheless, it cannot explain why markets in the NFL can sustain excess demand year after year. To explain this, I rely on a theory of group behavior.

2.3. Bandwagons, Social Influences, and Group Demand

Many products that consumers purchase have the property that the enjoyment or utility provided to the consumer depends on the crowd. Audience reaction and participation are consumed with the commodity itself. For instance, witnessing a live performance by a popular music group, eating in a popular restaurant, attending a Broadway play, and watching a football game all pro-

vide opportunities for the consumer to interact in a social setting with other consumers who are similarly enthused by the event and are able to signal this enjoyment. The audience's performance can also be an important contributing factor to the performance of the attraction (the team may react positively to fan support, or the rock star might provide a better show with a lively audience).

Gary Becker (1991) helped explain why consumers reveal excess demand for popular restaurants and leave nearby restaurants empty. Becker's theory is similar to Leibenstein's (1950) bandwagon model where an individual's demand depends on the aggregate consumption. Similar models have been put forth by Bass (1969) and others who explain that the likelihood of purchase may increase with the behavior in aggregate. However, such models are different from Becker's model in that they do not require social interaction. DeSerpa (1994) and DeSerpa and Faith (1996) have adapted Becker's model to explain that the composition of those participating in a jointly consumed venue may sometimes be as important as the number of participants.

I review Becker's explanations below. Three observations are important. First, higher demand is rewarded with higher valuations by consumers (i.e., the demand curve slopes upward); a crowded restaurant entices more patrons than an empty one. Second, multiple equilibrium are simultaneously possible with similar market prices (empty restaurants and full restaurants) with the equilibrium resulting in excess demand being desirable to owners. Third, the equilibrium with excess demand is potentially unstable. A corollary to Becker's theory is that initial conditions are often important in establishing the observed equilibrium. Failure to initially achieve an excess demand equilibrium may lead to stagnant outcomes in which it is impossible to achieve the more desirable excess demand equilibrium. Analogously, the inability of a team to sell out its stadium may produce a downward cycle of future poor attendance. I discuss this in greater detail below.

2.4. Becker's Model

Becker's model assumes that an individual's demand depends on the aggregate demand level where:

$$D = \sum d'(P, D) = F(P, D)$$

with $d'(P, D)$ denoting the demand of the i th consumer and D denoting the aggregate or market demand. The summation is taken over all purchasers $i = 1, 2, \dots, N$. I expect $f_p = \partial F / \partial P < 0$ since all individuals have downward sloping demands.² Similarly, I expect $F_d = \partial F / \partial D > 0$ because individual demands increase with group demand by assumption. For each value of market

²I use the notation F_p to denote the partial derivative i.e., $F_p = \partial F(P, D) / \partial P$.

demand, D , the equilibrium price solves $D = F(P, D)$. Let $P = G(P, D)$ denote the inverse demand function such that $D = F(G(D), D)$. Denote the function $G_d = \partial G / \partial D$. The sign of the G_d is indeterminate. G_d may be either positive or negative depending on the size of the group effect on individual demand. Using the implicit function theorem, we have

$$\frac{dP}{dD} = G_d = \frac{1 - F_d}{F_p}$$

If $F_d > 1$, an increase in aggregate demand increases the demand price (recall $F_p < 0$). On the other hand, if $F_d < 1$, an increase in aggregate demand decreases the demand price.

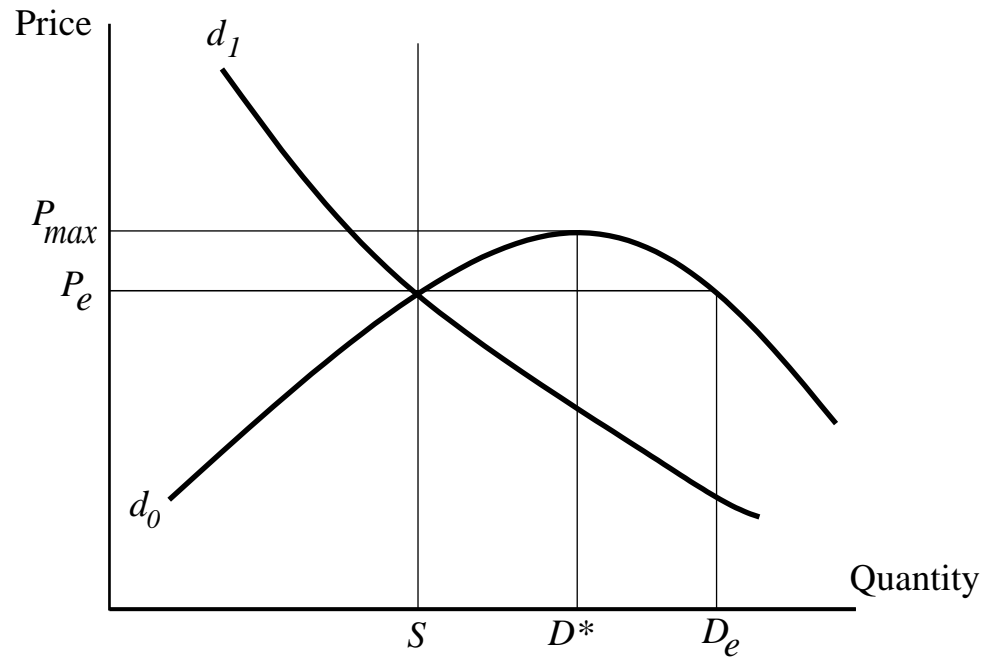
In Figure 2.1, I show the function for the aggregate demand $F(P, D)$ peaking at some value D^* ; this function is labeled d_0 . For purposes of illustration aggregate demand is shown increasing up to level D^* and decreasing after D^* .

An alternative function d_1 denotes aggregate demand without the presence of group effects, or where group effects are not dominant. The vertical line at quantity S denotes fixed available capacity. At price P_e , d_0 demand equals S and there is no excess demand. However, for the aggregate demand function, larger audiences are more desirable to consumers (at least over some range). At demand level D^* , consumers are willing to pay a higher price P_{\max} . Since demand exceeds supply at price P_{\max} , the available seats must be allocated by lottery or on a first-come first-serve basis. Given the restriction on capacity, excess demand is necessary to achieve the best price from consumers.

There is a second equilibrium in Figure 2.1 where quantity is D_e . At this equilibrium, consumers will also pay the market clearing price P_e . However, market demand exceeds S . DeSerpa (1994) observes that demand increases up to a point with increasing prices because of the optimal audience size. However, if capacity limits the audience size, increases in excess demand cannot fuel the excitement unless the rationing of excess demand affects audience composition.

Implicitly, for Becker's theory to apply to football, we must assume that demand depends on the aggregate quantity desired rather than upon those who actually succeed in the lottery and are able to partake of the experience. DeSerpa extends Becker's theory and argues that as long as the "buzz" or noise level (audience excitement) is inversely related to reservation prices, then demand will be increasing in price (i.e., the price consumers are willing to pay will increase with audience size). He argues that those who have the largest reservation values (the highest demand buyers) generate a relatively low noise level. By assigning the available seats at random to a larger audience, individuals with potentially lower reservation values are allowed to participate and generate a greater buzz for all who attend. One should contrast the football fan wearing a suit with the fan who paints his face and strips to the waist in subzero

Figure 2.1: Aggregate Demand Depends on Audience Size



weather. Andrew Zimbalist, as quoted in Swift (2000), supports the notion that the buzz contributes to fans' and teams' overall experience at a sporting event. "Corporate customers," he claims, "tend to be more sedate, which lessens the home field advantage." DeSerpa's assumptions seem reasonable for football, but are not as applicable to the restaurant setting considered by Becker.

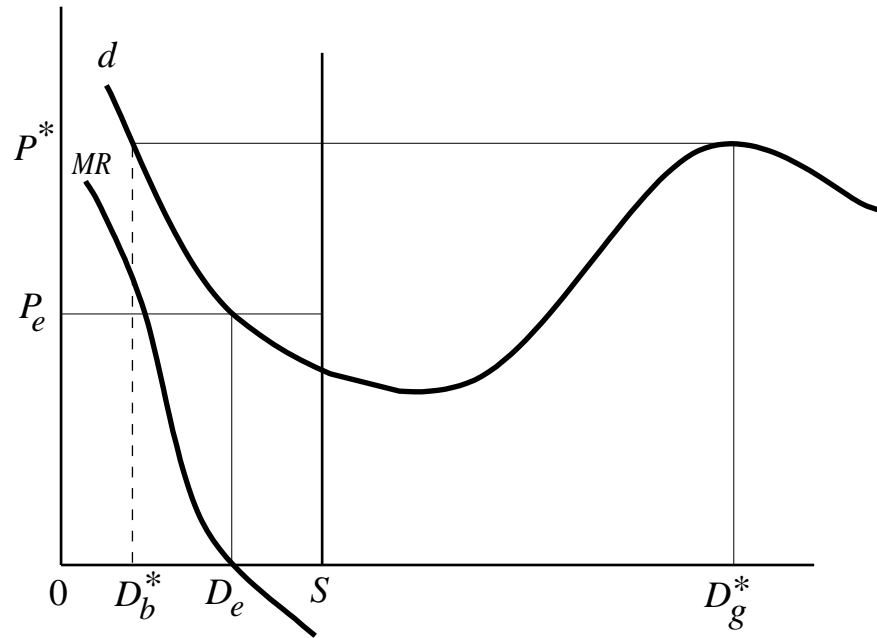
To illustrate the concepts of multiple equilibrium and instability of equilibrium, consider Figure 2.2. Here I assume that demand is downward sloping for low levels of demand, and is possibly elastic. For low levels of demand, declines in price cause increased demand, but the social interaction effects are not strong enough to overcome the decrease in demand price. Then, as audience size gets larger, the demand price increases. Eventually, the demand price turns down when the potential audience gets too large (i.e., in DeSerpa's case, when any individual's chance of getting in gets too low to put a large value on the outcome).

In one equilibrium, management can charge P_e where marginal revenue is zero. In this equilibrium, demand levels are below available capacity S . Alternatively, a price level P^* could be charged. At P^* , two equilibrium are possible. The first has low demand, more excess capacity and positive marginal revenue. This is not a desirable outcome for owners and corresponds to the "empty" restaurant Becker discussed or an undersold football stadium. Alternatively, at the same price, an equilibrium at quantity D_g^* with excess demand and large expected noise occurs. When maximizing revenues, owners would clearly prefer the second equilibrium to the other two. But can it be achieved?

One method to achieve the excess demand equilibrium is to create the impression that the stadium is smaller than it actually is. For example, consider the Miami Heat, an NBA franchise team in Florida. The Heat understands the value fans place on participating in a sold-out event. Although the Heat has one of the league's best records, won its regional division and advanced to the league playoffs during 1999-2000, "the team has closed the balconies seven times this year and obscured. . . the seats up there with black curtains." Hiding the empty balcony seats essentially fosters the impression of a sold-out arena; there are fewer vacant seats between people.

A second method to create the excess demand equilibrium is through advertising. For instance, Becker discusses how advertising and publicity can help get consumers to such a point. But failing to advertise or market properly could be a disaster. Similarly, a false start in ticket sales or a mismanaged season ticket program could be equally disastrous in seeking to achieve the excess demand equilibrium. Suppose that management sets too high a price (i.e., a price higher than P^*) in the first season. In this situation, demand will be much too low and far short of the desirable state of excess demand. Additionally, the equilibrium at P^* can be fickle. Consumers who lose confidence in the

Figure 2.2: Multiple Equilibria and Demand Instability



team's management will lower their demands. Only the high demand loyalists will remain to generate revenue. Thus, first impressions may be crucial.

Management does not want to expand capacity in a fragile equilibrium situation. When consumers are fickle, it may be more sensible to avoid a possible costly expansion that leaves the stadium undersold and underwhelmed. An ill-timed expansion can be equally disastrous.

In the next section, I examine empirical evidence on NFL football demand and use this model in part to confirm the Becker/DeSerpa theory.

2.5. The Demand for NFL Football

I follow the general approach of Noll (1974) and Welki and Zlatoper (1991, 1999) in specifying an empirical demand model for NFL football. I specify and estimate equations for both attendance and ticket sales. The data I compiled is for all teams during their regular seasons for the period 1995-1999. Each game is separately considered, allowing the specifics of the home team and visiting team to be included in the demand model. This study generalizes the previous published studies by using multiple seasons, all regular season games, and all teams collectively. In total, the regression is based on observations of over 1,200 data points ($5 \text{ years} \times 8 \text{ home games} \times 30 \text{ teams} = 1,200 \text{ observations}$). Table 2.1 reveals the source of the data and the factors derived from this data. Table 2.2 provides a glossary of factors assembled for the econometric analysis.

I consider econometric models that measure demand either as ticket sales as a proportion of stadium capacity, or attendance as a proportion ticket sales. These two variables normalize demand by measures of scale (i.e., differences across teams due to stadium size differentials or tickets sold). Ticket sales as a percentage of capacity are shown in Figure 2.3. These percentages are based on regular season contests for the period 1995–1999 and reflect the frequency of occurrence. For instance, in the period 1995–1999, 40.8 percent of all contests recorded ticket sales that were between 98 and 99 percent of capacity. Similarly, attendance as a percentage of ticket sales is shown in Figure 2.4.

A related concept is the ratio of attendance to capacity. This factor is shown in Figure 2.5. Attendance in proportion to capacity is the product of the measures for ticket sales as a proportion of capacity and attendance as a proportion of ticket sales. While a separate regression model was not developed for the ratio of attendance to capacity, this ratio is interesting because it demonstrates how full the stadiums are and therefore is related to the notions of sellouts and noise levels.³

³Welki and Zlatoper (1991) use attendance relative to capacity measures in their demand analysis.

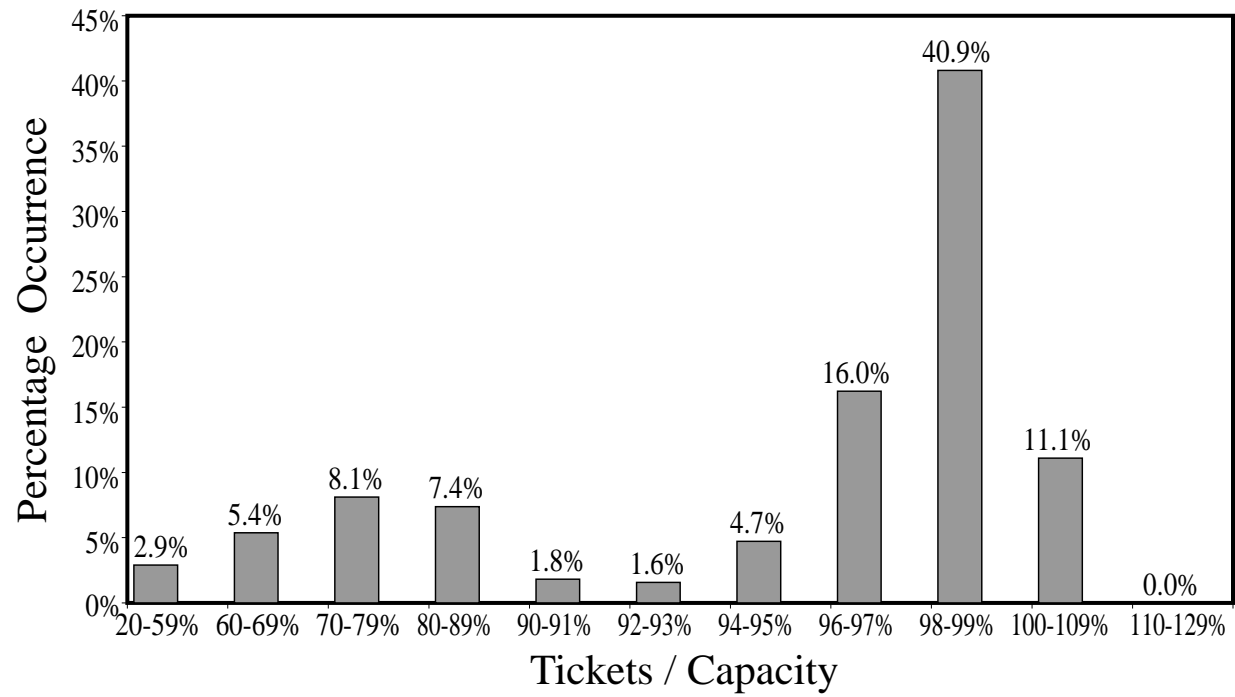
Figure 2.3: Ticket Sales as a Percentage of Capacity (1995-1999)

Figure 2.4: Attendance as a Percentage of Ticket Sales (1995-1999)

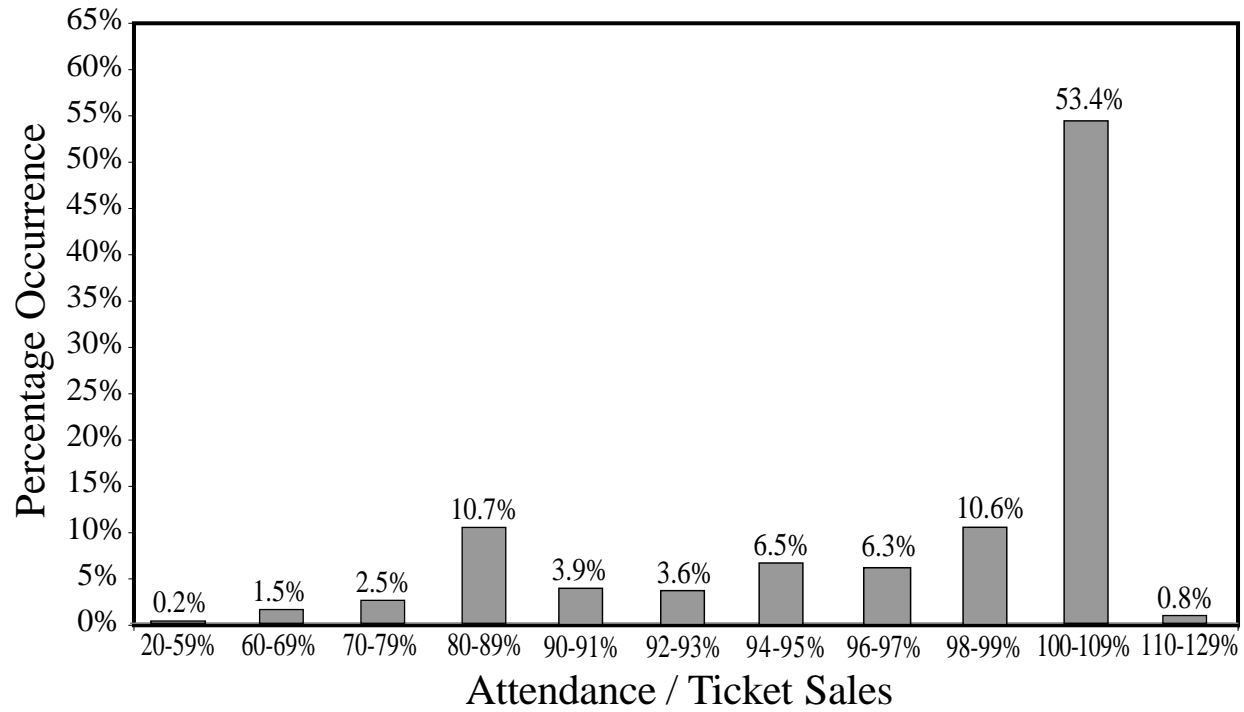


Figure 2.5: Attendance as a Percentage of Capacity (1995-1999)

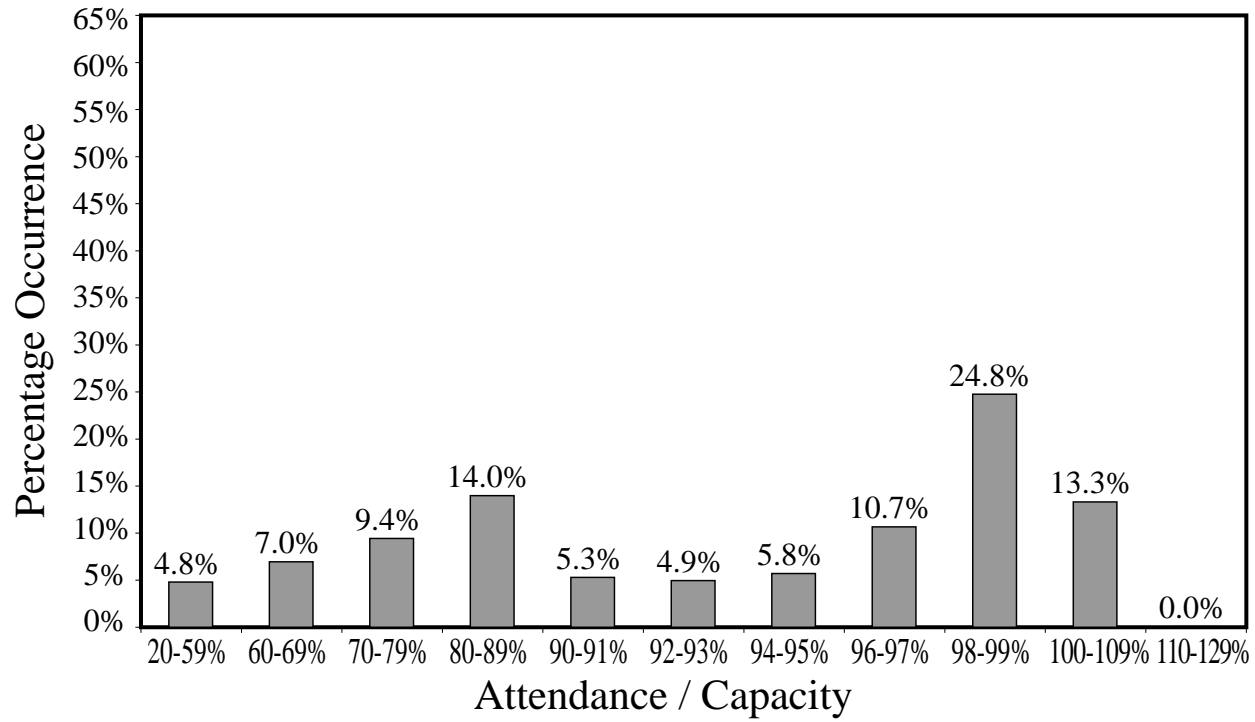


Table 2.1. Data Sources and Variables

I	National Football League Summary of Attendance Statistics (1995–1999): Average ticket price, including seat premiums, per season (by location) Capacity per manifest Date of contest Home team Indications of weather Location of contest Tickets sold per manifest Visiting team
II	Official National Football League Record and Fact Book (1995–1998): Attendance per manifest Indications of which team won each contest
III	National Football league Website http://www.nfl.com (1999): Attendance per manifest Indications of which team won each contest
IV	Socio-demographic Data from http://www.oseda.missouri.edu for the 1990 census data at specific metro locations Percent of population, Black Percent of population, Hispanic Percent of population, White Percent of population below poverty 1990 per capita income
V	Street and Smith’s Sports Business Journal: Average December temperature
VI	Factors derived from source material: Day of the week the contest was held Home game winning percentage (per team) (calculated by considering all regular-season home games) Overall winning percentage (per team) (calculated by considering all regular-season games) 1999 population (calculated using the average of the rates of change in population for 95–96, 96–97, 97–98, and applying the average to 1998 population figures by city)
VII	Quick Stats Website from http://www.quickstats.com/index.htm : Start time per manifest (1995-1999)

In Figures 2.6, 2.7, and 2.8, I display the same three ratios by time period. The average value calculated in each season is the weighted average across all teams playing in that season.

Following the standard economic demand theory and the sports demand literature, demand is specified to depend on ticket price (average ticket prices are used, including premiums for higher quality seats and season ticket sales) and the performance of the contestant teams (both the performance of the home team and that of the visitor). I also considered potential explanatory factors for the racial composition, income levels, and population of the local metro area along with variables to measure the weather on game day and whether

Figure 2.6: Ticket Sales as a Percentage of Capacity (1995-1999)

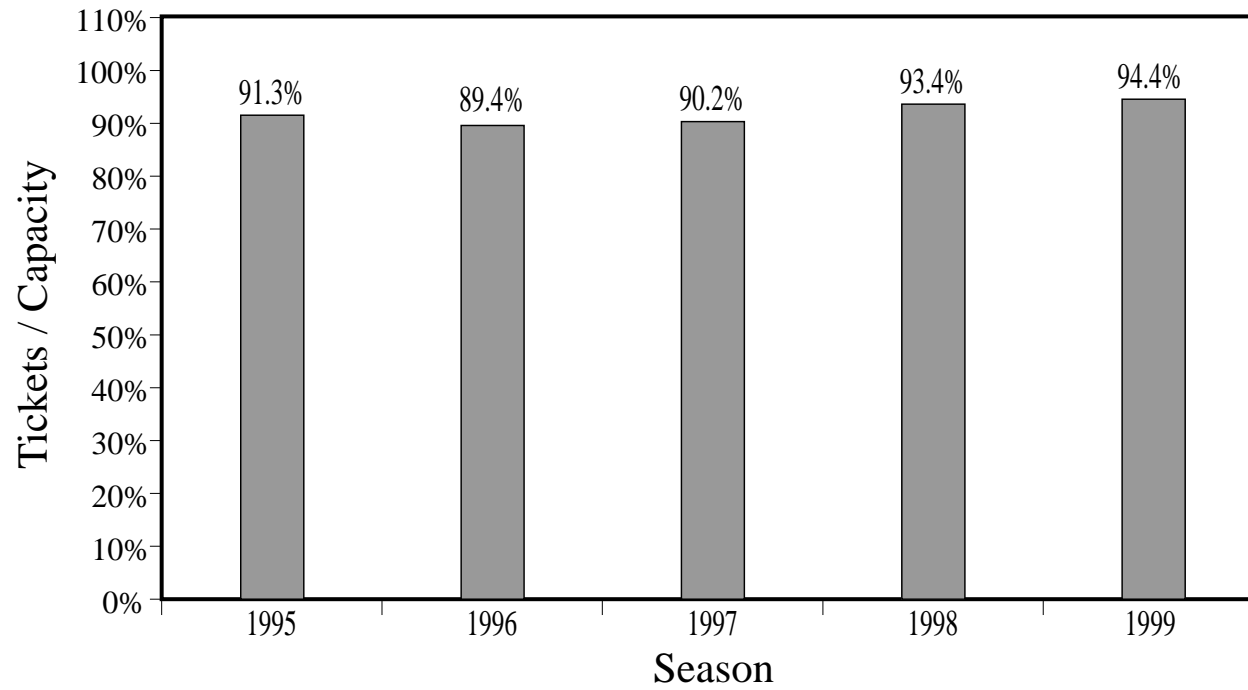


Figure 2.7: Attendance as a Percentage of Ticket Sales (1995-1999)

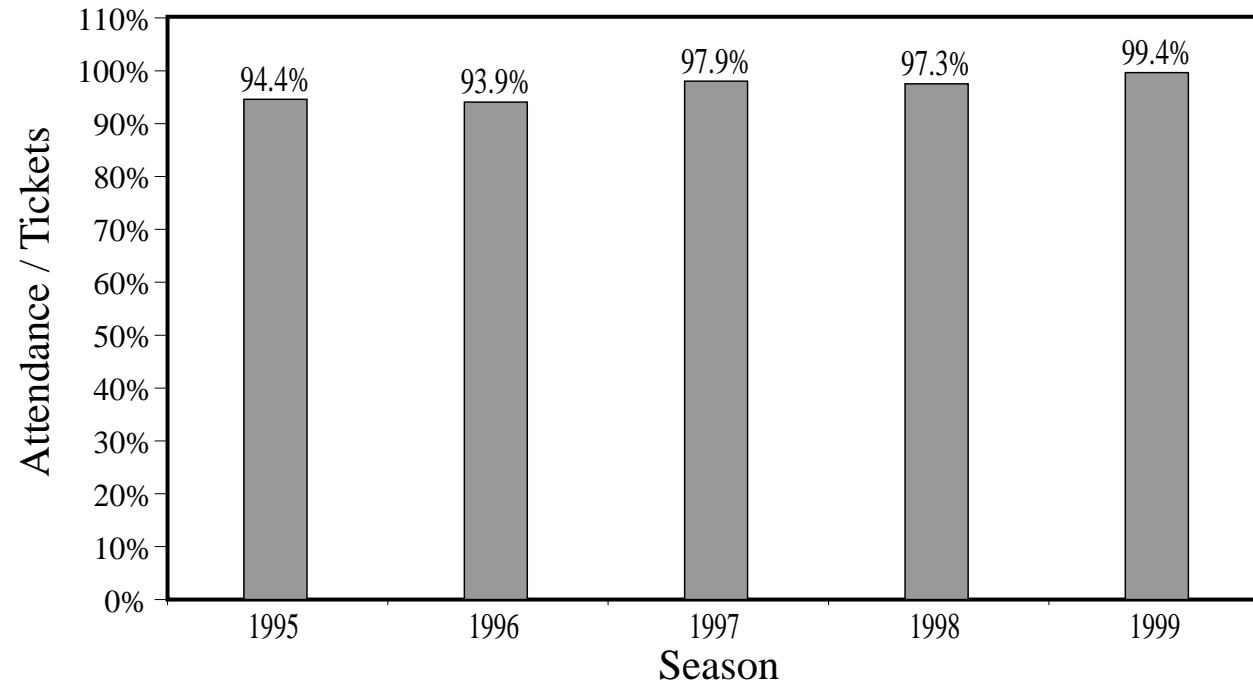


Figure 2.8: Attendance as a Percentage of Capacity (1995-1999)

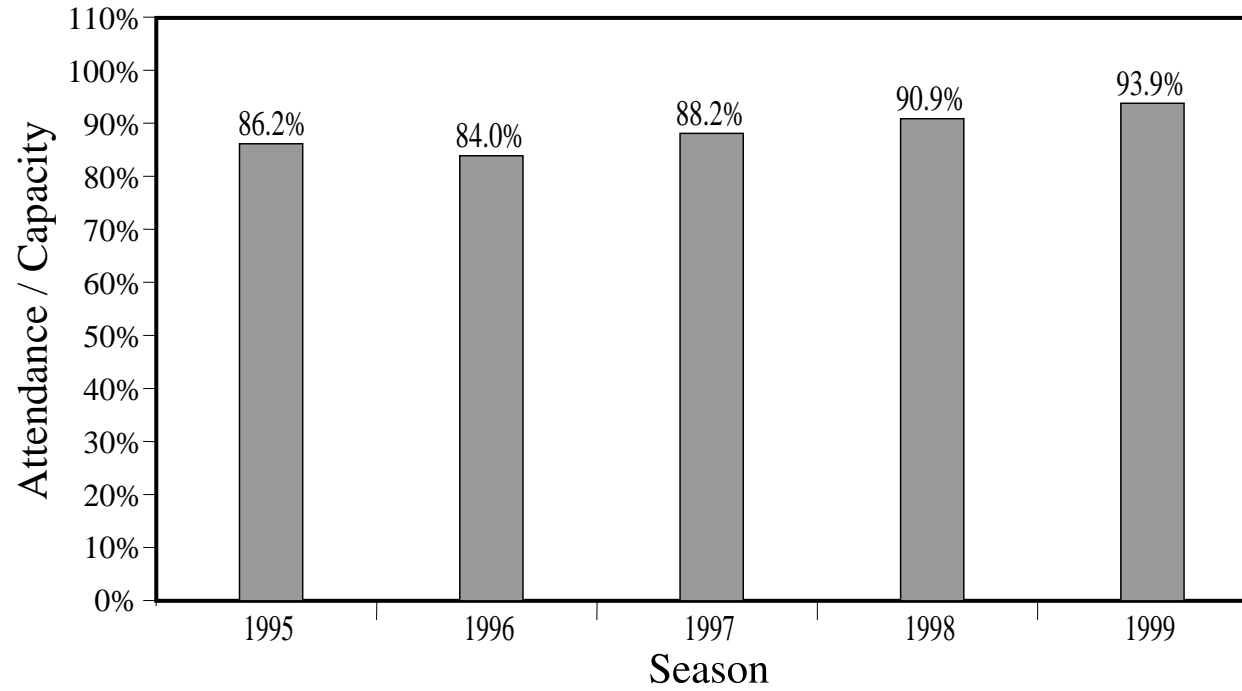


Table 2.2. Variable Glossary

attndnce	game attendance
cap	stadium capacity by season
lpricesp	log of average ticket price per season, including seat premium
p_black	percent of metro area residents, Black
p_hisp	percent of metro area residents, Hispanic
p_white	percent of metro area residents, White
pwin_h	home team's winning percentage (as of last game played)
pwin_v	visiting team's winning percentage (as of last game played)
rain	rainy day
seats	stadium seats per 100,000 metro residents
ticket	tickets sold per game
trend	season + (game-1)/16
monday	game played on a Monday
night	game played at night (after 4 p.m. local time)

the game was played on a Monday or at night. I controlled for the size of the stadium relative to the population using a variable specified as the ratio of stadium seats to population. All size variables were transformed using logarithms. Table 2.3 provides basic summary statistics for the variables listed in Table 2.2.

Table 2.3. Variable Statistics

Variable	Number of Observations	Minimum	Maximum	Mean	Standard Deviation
log(attndnce/ticket)	1,208	-1.25	0.23	-0.04	0.10
log(seats)	1,208	6.96	11.03	9.32	0.90
log(ticket/cap)	1,208	-1.38	0.04	-0.10	0.18
lpricesp	1,208	3.25	4.40	3.68	0.22
p_black	1,208	0.006	0.76	0.32	0.19
p_hisp	1,208	0.006	0.62	0.11	0.13
p_white	1,208	0.22	0.94	0.59	0.17
pwin_h	1,133	0.00	1.00	0.49	0.27
pwin_v	1,132	0.00	1.00	0.51	0.27
rain	1,207	0.00	1.00	0.05	0.23
trend	1,208	1.00	5.94	3.48	1.45
monday	1,208	0.00	1.00	0.07	0.25
night	1,208	0.00	1.00	0.21	0.41

The regression models for ticket sales and attendance are presented in Table 2.4. Demand for NFL tickets depends significantly on the percentage of games won in the season to date for the home team. The winning percentage of the visiting team also shows a significant positive effect. The magnitude of

the coefficients shows that the winning percentage of the home team is more important than the winning percentage of the visiting team as they affect ticket sales and attendance.

Table 2.4. Demand Models

Explanatory Variable	Model 1	Model 2
	Log (Tickets Sold/Capacity)	Log (Attendance/Tickets Sold)
Constant	-0.447 (-3.87)	-0.567 (-8.47)
Home Team's % Wins	0.168 (9.16)	0.077 (7.24)
Visiting Team's % Wins	0.054 (3.04)	0.028 (2.76)
Log of Average Ticket Price Per Season, Including Seat Premium	0.147 (5.74)	0.0475 (3.20)
Trend	0.002 (0.53)	0.0097 (4.38)
% Black	0.223 (3.27)	0.0745 (1.89)
% Hispanic	-0.037 (-0.85)	-0.049 (-1.94)
% White	0.394 (5.12)	0.061 (1.36)
Rain	0.0173 (0.83)	-0.0615 (-5.12)
Log (Capacity/Population)	-0.0670 (-10.45)	0.0232 (6.26)
Night	0.0419 (3.03)	-0.0095 (-1.18)
Monday	0.0335 (1.50)	0.0337 (2.57)
Observations	1131	1131
R-Squared	0.192	0.195

t-statistics in parenthesis

There are no discernible trend effects in the model for ticket demand. Larger stadiums in relationship to population have lower ticket sales (normalized by capacity of the stadium). Controlling for race shows some significant race effects with whites and blacks leading to higher demand relative to Hispanics and others. The ticket price effect is statistically significant and positive; this effect confirms the Becker/DeSerpa model and supports a positive association between price and sales. Finally, night games are positively related to ticket sales whereas Monday night games apparently do not increase ticket sales. The

results of the model are very robust. Alternative specification using additional variables for weather and other factors produced similar results.⁴

The attendance model is similar to the ticket sales model in most respects. For instance, the attendance model also shows positive price effects. However, there are some key differences. In the attendance model, there are positive trend effects demonstrated. The results further indicate that attendance is positively affected by Monday night play but is not affected by night games more generally.

As a matter of economic theory, the more relevant demand model is based on ticket sales. The attendance model is useful because it demonstrates results that are both similar to and different from the models estimated in previous research. For instance, the variable indicating rain on the game day shows that rainy days are associated with lower attendance - a result found in the literature. But rainy days have no empirical effect on ticket sales. This might follow from the fact that very few sales are likely to be made on the day of game for professional football games. Most tickets are sold well in advance of game day. Similarly, night games are a bigger draw for sales (i.e. anticipated attendance) than for actual attendance.

2.6. Conclusions

This chapter has discussed the option values inherent in NFL season tickets, the conditions under which football ticket sales are likely to be successful, the importance of initial conditions for establishing successful future ticket sales, and the importance of advertising and marketing in establishing proper initial conditions. Establishing conditions of sellouts and excess demand are found to be crucial to the performance of NFL football at the box office. The Becker/DeSerpa model was tested by empirical analysis. I found support for the bandwagon theory of football demand. I also found that a team's performance, while helpful in generating ticket sales, is not by any means the only contributing factor.

The demand for football depends in part on the fans' experience. Demand for many teams is characterized by excess demand, and fans may wait several years before getting a chance to own season tickets. While excess demand is not uncommon in football, there are, from time to time, unfortunate outliers. Thus the initial conditions that characterize the future attendance in a bandwagon sport such as football are important in determining future outcomes. This is especially true as there is a very high correlation of sellouts over time.

⁴Using a Tobit analysis to account for the truncation of demand at capacity produced very similar results. This occurs because few teams sell out all of their capacity even if games are defined as "sell outs" by league standards.

Chapter 3

DETECTING AND MEASURING SHIFTS IN THE DEMAND FOR DIRECT MAIL

3.1. Introduction

This chapter evaluates the forecast accuracy of a structural econometric demand model for direct mail in Canada. Direct mail also known as advertised admail is used to provide advertising to consumers through the mail system. My original model was developed in March 1986 and was based on the period ending January 1996 using twelve years of historical monthly data. A complete discussion of the original regression model “original” is provided in Dubin (1998). In this chapter, I update the regression results for the April 1989 through January 1996 period “revised” and provide new results from twenty additional monthly observations for the period of February 1996 through September 1997 “updated”.

In Section 3.2, I present the original, revised, and updated regression models for direct mail. In Section 3.3, I discuss the shift in demand for direct mail in Canada and the accuracy of the structural economic model. Section 3.4 provides my estimates of forecast accuracy and conclusions.

3.2. Comparison of Original, Revised, and Updated Direct Mail Models

To review, my econometric model for addressed admail used seasonal dummy variables (seasonal indicators), real paper and printing prices, real product price, real retail sales, and trend to explain historical demand. A variable glossary is presented in Table 3.1.

Table 3.2 presents the results of the direct mail regression models (as originally estimated) for short-long (standard business), oversized (non-standard business size), and combined admail demand for the period ending in January

Table 3.1. Variable Glossary

(1)	constant term
s1-s11	month dummy variables
pi_mnppr/cpi	real paper and printing cost index
all_sl_v	short-long advertised admail volume
all_us_v	oversized advertised admail volume
all_ad_v	combined advertised admail volume
all_sl_p	short-long advertised admail price
all_os_p	oversized advertised admail price
all_ad_p	combined advertised admail price
ret_m_al/cpi	real retail sales index
trend	time trend

1996.¹ Generally the fit of the model was quite good and the explanatory factors were significant and in accord with economic theory.

In Table 3.3, I re-estimate the regression models for the time period ending in January 1996. As there were revisions to the Canadian historical time-series direct mail data and revisions to the Statistics Canada macroeconomic data, these models differ from my original reported model. I use the revised data regressions in the forecasts of admail demand for the 20-month period that occurred between my 1996 study (Dubin 1998)) and this study. Generally, these models are very similar to the models as originally estimated and, therefore, I do not discuss them in significant detail. A complete discussion is provided in Dubin (1998). For instance, the original price elasticity of the demand for admail was estimated to be -0.42 and was re-estimated using the revised data to be -0.40 . I note that the new estimate using the additional 20 months of data is -0.43 .

3.2.1. Comparison of Revised and Updated Models

I present the updated regression models for Addressed Admail in Table 3.4.

These models for short-long, oversized, and total admail are very similar to the models as originally estimated. The regression fit is approximately the same as in the original model, with R^2 values at roughly 87 percent. The estimates for the elasticity of total admail demand with respect to paper and printing prices fell from -0.32 to -0.28 , the retail sales elasticity fell from 0.7 to

¹For data sources and variable construction see Dubin (1998).

0.6, and the estimated trend effect was approximately the same. The price elasticity for total admail rose slightly from my previous estimate of -0.42 to the current estimate of -0.43 . The short-long price elasticity fell from -0.26 to -0.16 and the oversize price elasticity estimate rose from -0.51 to -0.65 . In sum, the changes in the estimated regression models are small and the general conclusions I previously reached remain unaltered. Specifically, the demand for addressed admail is still growing, still responsive to paper and printing costs, still moves pro-cyclically with the Canadian economy, and still reveals inelastic price elasticity.

3.2.2. Addressed Admail—Specific Results

Paper and printing costs remained steady in real terms until the rapid increase and subsequent fall in 1995. The estimated price elasticity of real paper and printing costs is estimated to be -0.28 (t -stat = -3.3) indicating that an increase in the real cost of paper and printing causes the demand for addressed admail to decline. Moreover, customers are captive to this cost since the elasticity is less than one. Retail sales have a positive and statistically significant effect on admail demand. A percentage increase in retail sales leads to a 0.61 percentage increase in the demand for addressed admail.

As noted above, the estimated price elasticity of -0.43 (t -stat = -1.94) shows that an increase in the real price of admail causes admail volumes to decline. The inelasticity estimated by the model indicates that an increase in the price of Canadian addressed admail products will result in lower volume, but higher overall revenues at a higher price level. Profitability is also predicted to increase as costs decline with lower total volumes.

In order to examine the separate price elasticities for specific types of admail, I specified and estimated two additional models. I present these models using the updated data in Table 3.4. The first model considers short-long admail. The own-price elasticity is estimated at -0.16 (t -stat = -0.66). Thus, it is less elastic than both oversized or combined addressed admail with elasticities of -0.65 and -0.43 , respectively. Short-long admail is also revealed to have grown more rapidly than oversized admail after controlling for other factors. Paper and printing costs have a negative effect on both short-long and oversized volumes. Given the estimated elasticities, it is likely that a uniform increase in the price of all addressed admail types will create a greater volume decline in oversized pieces than in short-long pieces. These conclusions, therefore, remain unchanged relative to my 1996 study (Dubin (1998)).

3.3. The Shift in Demand Between 1995 and 1996

The average price of addressed admail in fiscal year 1995 was \$0.261 per piece. In fiscal year 1996, this rose to \$0.289 per piece. Volumes on an annual

basis rose from 1,458,510 to 1,510,016 pieces. Additionally, between fiscal year 1995 and fiscal year 1996, several key explanatory factors showed important changes. First, the Canadian economy improved. Second, the real cost of paper and printing fell. Third, other trend effects caused demand to rise. These shifts would have combined to increase demand even if addressed admail prices had remained constant. Therefore, the price increase between fiscal years 1995 and 1996 caused a movement along the demand curve while other factors caused a shift in the demand curve.

I illustrate this using the estimated demand curve for fiscal years 1995 and 1996 as depicted in Figure 3.1. The figure shows the consequences of the price increase as a movement along the demand curve from Point A to B. The movement from Point B to C represents the shift in the demand curve. Hence, the move between fiscal years 1995 and 1996 (the movement between Points A and C) has two components—the price component (A to B) and the shift component (B to C).

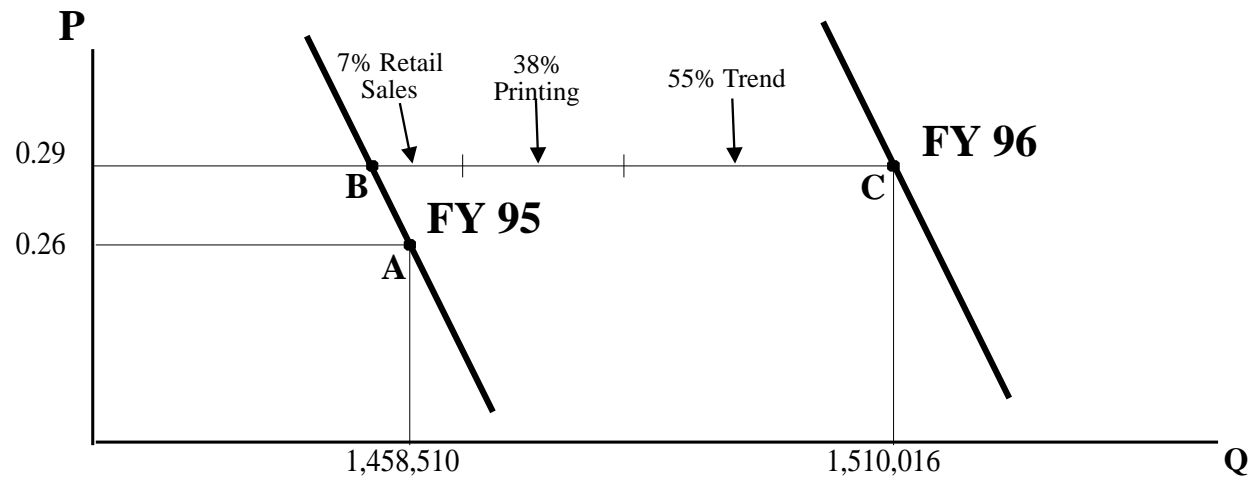
Between fiscal years 1995 and 1996, average addressed admail prices rose from \$0.261 to \$0.289 per piece. In nominal terms, this was an 10.7 percent price increase. In real terms, the price increase was lower, at 8.8 percent as average inflation in Canada between the two fiscal years was approximately 1.8 percent. During this same period, volumes increased about 3.5 percent. As a consequence of the price increase, volumes fell (the movement along the demand curve). At the same time, general trend effects increased demand for addressed admail while improvements in retail sales caused a further increase in demand. The substantial decline in real paper and printing costs of 14 percent caused demand to increase substantially. Thus, the change in volume due to increased admail price per piece was almost fully offset by lower paper and printing costs.

Therefore, the shift in demand from Point B to Point C was due to three factors: retail sales improvements; decline in real costs of paper and printing; and general trend effects. Using the estimated regression model, I find that approximately 7 percent of the shift was due to the improvement in the Canadian economy, 38 percent of the shift was a result of lower paper and printing costs, and the remaining 55 percent of the shift was due to general trends causing admail demand to increase each year.

3.4. Comparing Forecast Accuracy

In Figure 3.2, I display the *ex post* forecast of total addressed admail for the period February 1996 through September 1997. This forecast is based on the revised regression for total addressed admail presented in Table 3.3, column 3. This is an *ex post* forecast since it relies on the realized volumes of the explanatory factors. The plot in Figure 3.2 reveals some variation between forecasted and actual values.

Figure 3.1: Shift in Demand Between Fiscal Years 1995 and 1996



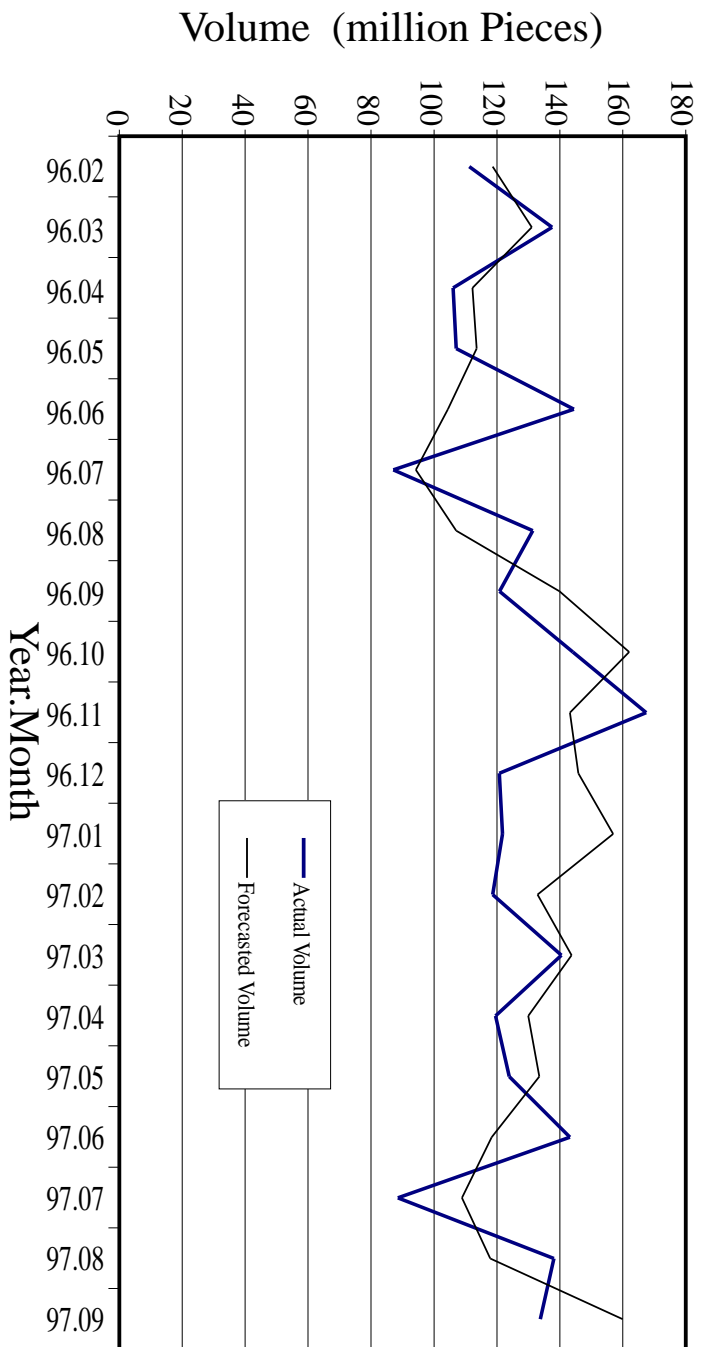


Figure 3.2: Forecasted Volume Versus Actual Volume

3.4.1. Forecasting Addressed Admail Volume

To predict the value of y^0 associated with a regressor vector x^0 under the model

$$y^0 = \beta'x^0 + \varepsilon^0,$$

recall that $\hat{y}^0 = b'x^0$ is the minimum variance linear unbiased estimator of $E[y^0]$. The forecast error is

$$e^0 = y^0 - \hat{y}^0 = (\beta - b)'x^0 + \varepsilon^0.$$

The forecast variance is:

$$\begin{aligned}\text{Var}[e^0] &= \sigma^2 + \text{Var}[(\beta - b)'x^0] \\ &= \sigma^2 + x^{0'}[\sigma^2(X'X)^{-1}]x^0.\end{aligned}$$

I estimate the forecast variance using s^2 in place of σ^2 . The estimated standard error of the forecast variance is denoted by $se(\hat{y}^0)$. A confidence interval for y^0 is formed using $\hat{y}^0 \pm t_{\lambda/2}se(\hat{y}^0)$. For the log-linear demand models used in this study, the forecast confidence interval is:

$$\text{prob} \left[e^{\hat{y}^0} e^{-t_{\lambda/2}se(\hat{y}^0)} \leq e^{y^0} \leq e^{\hat{y}^0} e^{t_{\lambda/2}se(\hat{y}^0)} \right] = 1 - \lambda$$

Salkever (1976) suggested a method for combining the computation of forecasts and standard errors using an expanded regression. Suppose that the estimation is based on n observations and that we desire to forecast n^0 observations,

$$y^0 = X^0\beta + \varepsilon^0.$$

Construct an augmented regression:

$$\begin{bmatrix} Y \\ \dots \\ 0 \end{bmatrix} = \begin{bmatrix} X & \vdots & 0 \\ \dots & \dots & \dots \\ X^0 & \vdots & -I \end{bmatrix} \begin{bmatrix} \beta \\ \dots \\ \gamma \end{bmatrix} + \begin{bmatrix} \varepsilon \\ \dots \\ \varepsilon^0 \end{bmatrix}$$

or

$$y^* = X^*\beta^* + \varepsilon^*.$$

In the augmented regression, there are n^0 new observations and a set of n^0 new variables. Each column in the second part of X^* is a dummy variable that takes the value minus one for the new observations and zero otherwise. Using this expanded regression model, Salkever (1976) shows that the least squares regression of y^* on X^* produces the coefficient vector $[b, c]$, where b is the original OLS coefficient vector and c are the predictions of y^0 ; the residuals

from this regression are, for the first n observations, the original least squares residuals and for the last n^0 , zero; and the estimated covariance matrix for the expanded vector of coefficient estimates contains, in its upper left block, the covariance matrix for the least squares estimates of β and, in its lower block, the covariance matrix for the forecasts.

The i th diagonal element of the covariance matrix in the lower block is:

$$\text{Est Var } [c_i] = \text{Est Var } [\hat{y}_i^0] = s^2 \left[1 + x_i^{0'} (X'X)^{-1} x_i^0 \right].$$

3.4.2. Measures of Forecast Accuracy

Various measures have been proposed for assessing the predictive accuracy of forecasting models. Most of these measures are designed to evaluate *ex post forecasts*, that is, forecasts in which the exogenous variables do not themselves have to be forecasted. A common measure based on the residuals from the forecasts is the root mean squared error:

$$\text{RMSE} = \sqrt{\frac{1}{n^0} \sum_i (y_i - \hat{y}_i)^2}$$

where n^0 is the number of forecasted periods. This has an obvious scaling problem as it depends on the magnitude of y_i . Several measures without scaling problems are based on the Theil U statistic:²

$$U = \sqrt{\frac{(1/n^0) \sum_i (y_i - \hat{y}_i)^2}{(1/n^0) \sum_i y_i^2}}.$$

Large values of Theil's U statistic indicate poor forecasting performance. In order to examine the model's ability to track turning points it is possible to replace the levels y_i , with either absolute change, $y_i - y_{i-1}$, or percentage changes $(y_i - y_{i-1})/y_{i-1}$. The U statistics are calculated below.

3.4.3. Forecast Results—Confidence Intervals

In Table 3.5, I show the estimated 95 percent confidence intervals for the addressed admail model. The estimation period ends in January 1996, which was the last available data point in my previous study. The period from February 1996 through September 1997 was forecasted *ex post* (i.e., using the realized values of exogenous factors appearing in the model, including paper and printing costs, retail sales, admail prices, and so forth).

In Figure 3.3, I show the 95 percent confidence interval, the forecasted point estimates, and the realized actual values for this period. I note that only two

²See Theil (1961).

of twenty values fall outside the estimated intervals. Therefore, the estimated confidence intervals contained the true values 90 percent of the time.

3.4.4. Forecast Results—Conclusions

The 95 percent confidence intervals range in volume by plus 23 percent and minus 19 percent of typical monthly volumes. The total addressed admail volume for the 20-month period was 2,505,951,000 pieces. The model's prediction for this period was 2,574,181,000 pieces. The difference is approximately a 2.7 percent error. Hence, the model is quite good at forecasting the overall volume even if there are some substantial monthly errors. The total addressed admail volume for fiscal year 1996/1997 was 1,510,016,000 pieces. The model's prediction for this period was 1,556,070,000 pieces. Hence, the model over-predicted the results for fiscal year 1996/1997 by approximately 3.0 percent.³

Finally, I have calculated Theil U statistics for the forecast period based on levels, differences, and percentage changes. The U statistic for levels is on the order of 11 percent, while the U statistics for differences or percentage changes is about 55 percent. These U statistics reveal that the direct mail demand model does a better job at predicting levels than turning points.

³The point estimates were within as little as 2.3 percent of actual monthly volume and were off by as much as 28 percent in one month. Apparently, the model is particularly poor in forecasting the extreme swings that occur in the summer months. The model tends to under-predict the June highs and the July lows.

Figure 3.3: Estimated Confidence Interval for Addressed Admail

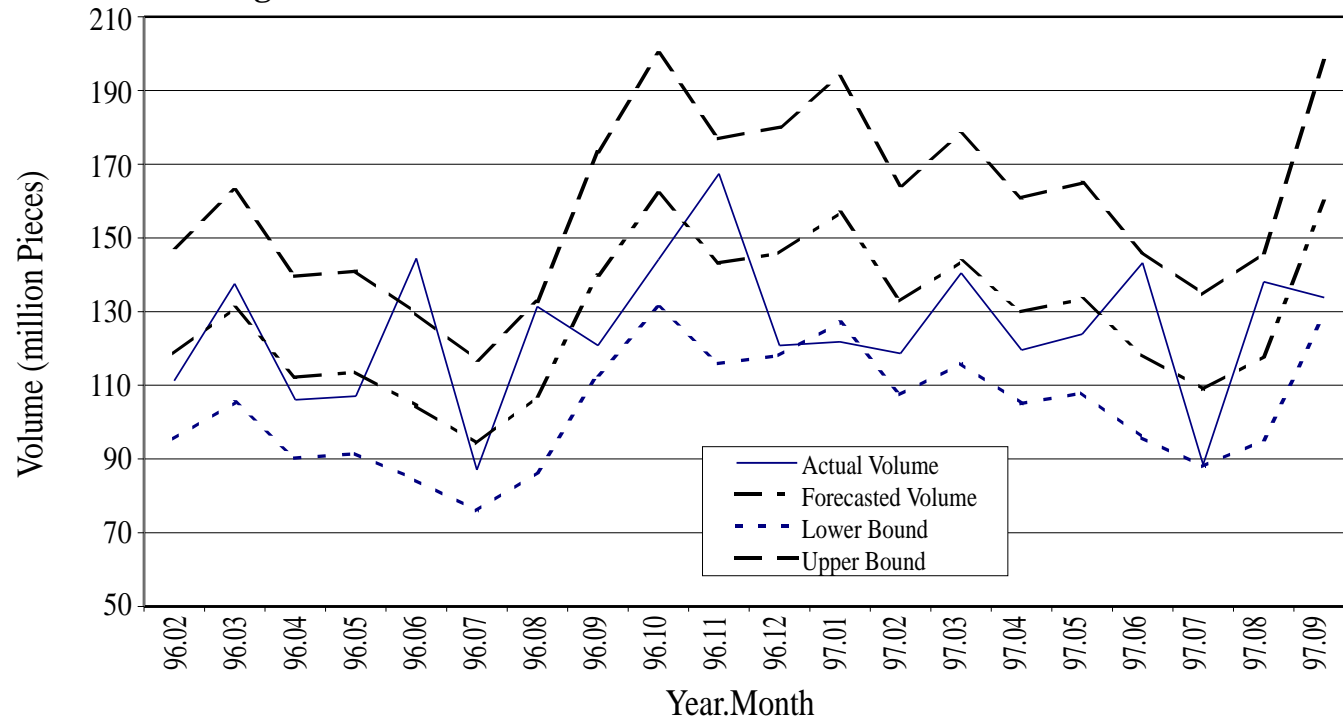


Table 3.2. Addressed Admail (Original)

Dep Variable:	Admail S/L (log(all_sl.v))	Admail O/S (log(all_os.v))	Admail All (log(all_ad.v))
(1)	-24.006 (-7.66)	-15.753 (-5.60)	-20.133 (-8.23)
s1	0.141 (2.70)	0.367 (7.04)	0.253 (5.98)
s2	0.007 (0.12)	0.240 (4.18)	0.118 (2.56)
s3	0.102 (2.85)	0.147 (3.71)	0.122 (4.09)
s4	-0.021 (-0.64)	-0.086 (-2.32)	-0.051 (-1.82)
s5	-0.037 (-1.01)	-0.214 (-5.29)	-0.121 (-3.90)
s6	-0.112 (-3.12)	-0.316 (-8.04)	-0.2045 (-6.75)
s7	-0.325 (-9.65)	-0.226 (-6.06)	-0.281 (-9.99)
s8	-0.207 (-6.16)	-0.127 (-3.40)	-0.175 (-6.21)
s9	0.097 (2.90)	0.167 (4.48)	0.127 (4.49)
s10	0.220 (6.51)	0.184 (4.93)	0.201 (7.13)
s11	0.161 (4.59)	-0.113 (-2.98)	0.042 (1.41)
(log(pi_mnppr/cpi))	-0.874 (-2.41)	-0.317 (-2.59)	-0.321 (-3.27)
(log(all_ad.p/cpi))	-0.267 (-1.12)	-	-
(log(all_os.p/cpi))	-	-0.516 (-1.90)	-
(log(all_ad.p/cpi))	-	-	-0.423 (-1.87)
(log(ret_m.al/cpi))	-2.709 (1.23)	1.140 (5.45)	0.702 (4.05)
(log(trend))	7.065 (20.22)	3.905 (10.47)	5.655 (20.69)
Observations	142	142	142
Corrected R-squared	0.872	0.751	0.878
Mean of Dep Variable	10.826	10.522	11.385

t-statistics in parenthesis

Table 3.3. Addressed Admail (Revised)

Dep Variable:	Admail S/L (log(all_sl_v))	Admail O/S (log(all_os_v))	Admail All (log(all_ad_v))
(1)	-23.871 (-7.34)	-15.872 (-5.60)	-20.145 (-8.00)
s1	0.137 (2.67)	0.365 (6.98)	0.250 (5.94)
s2	0.003 (0.05)	0.238 (4.13)	0.114 (2.49)
s3	0.101 (2.83)	0.146 (3.70)	0.121 (4.05)
s4	-0.021 (-0.63)	-0.086 (-2.32)	-0.051 (-1.80)
s5	-0.036 (0.98)	-0.213 (-5.27)	-0.119 (-3.85)
s6	-0.111 (-3.10)	-0.316 (-8.03)	-0.203 (-6.71)
s7	-0.324 (-9.63)	-0.227 (-6.08)	-0.282 (-9.97)
s8	-0.207 (-6.15)	-0.127 (-3.42)	-0.175 (-6.21)
s9	-0.097 (2.89)	-0.166 (4.46)	-0.126 (4.45)
s10	0.220 (6.52)	0.183 (4.92)	0.201 (7.12)
s11	0.162 (4.62)	-0.113 (-2.98)	0.043 (1.44)
(log(pi_mnppr/cpi))	-0.268 (-2.37)	-0.299 (-2.46)	-0.298 (-3.15)
(log(all_sl_p/cpi))	-0.244 (-1.03)	-	-
(log(all_os_p/cpi))	-	-0.512 (-1.89)	-
(log(all_ad_p/cpi))	-	-	-0.404 (-1.79)
(log(ret_m_al/cpi))	0.253 (1.17)	1.128 (5.38)	0.684 (3.99)
(log(trend))	7.088 (19.11)	3.951 (10.45)	5.703 (20.11)
Observations	142	142	142
Corrected R -squared	0.872	0.752	0.877
Mean of Dep Variable	10.826	10.522	11.385

t-statistics in parenthesis

Table 3.4. Addressed Admail (Updated)

Dep Variable:	Admail S/L (log(all_sl_v))	Admail O/S (log(all_os_v))	Admail All (log(all_ad_v))
(1)	-22.115 (-7.01)	-16.535 (-5.90)	-19.662 (-7.48)
s1	0.097 (1.98)	0.347 (6.74)	0.222 (5.30)
s2	-0.035 (-0.66)	0.223 (3.95)	0.089 (1.95)
s3	-0.095 (2.79)	-0.158 (4.12)	0.123 (4.14)
s4	-0.037 (-1.16)	-0.080 (-2.22)	-0.057 (-2.04)
s5	-0.027 (-0.77)	0.226 (-5.74)	-0.120 (-3.88)
s6	-0.063 (-1.85)	-0.276 (-7.24)	-0.159 (-5.30)
s7	-0.341 (-10.60)	-0.237 (-6.55)	-0.296 (-10.57)
s8	-0.178 (-5.55)	-0.090 (-2.50)	-0.144 (-5.12)
s9	0.076 (2.38)	0.150 (4.14)	0.106 (3.78)
s10	0.222 (6.67)	0.169 (4.53)	0.196 (6.77)
s11	0.183 (5.30)	-0.098 (-2.57)	0.061 (1.98)
(log(pi_mnppr/cpi))	-0.221 (-2.20)	-0.315 (-2.98)	-0.280 (-3.31)
(log(all_ad_p/cpi))	-0.157 (-0.66)	-	-
(log(all_os_p/cpi))	-	-0.654 (-2.95)	-
(log(all_ad_p/cpi))	-	-	-0.430 (-1.94)
(log(ret_m_al/cpi))	0.111 (0.54)	1.103 (5.28)	0.611 (3.54)
(log(trend))	6.973 (20.86)	3.930 (10.46)	5.638 (20.39)
Observations	162	162	162
Corrected <i>R</i> -squared	0.877	0.722	0.868
Mean of Dep Variable	10.881	10.546	11.427

t-statistics in parenthesis

Table 3.5. Addressed Admail 95% Confidence Intervals

Period	Confidence Interval (lower bound)	Actual Volume (1,000's)	Forecasted Volume (1,000's)	Confidence Interval (upper bound)
February 96	95,476	111,230	134,917	147,394
March 96	105,461	137,573	118,628	162,952
April 96	90,222	106,041	131,092	139,583
May 96	91,472	107,056	112,221	140,998
June 96	84,232	144,462	113,567	129,529
July 96	75,987	87,051	104,452	116,809
August 96	86,417	131,374	94,213	132,684
September 96	112,895	120,815	107,080	173,337
October 96	131,211	144,097	139,889	200,115
November 96	115,903	167,412	162,041	176,926
December 96	118,106	120,788	143,200	180,077
January 97	127,136	121,785	145,836	193,705
February 97	107,523	118,632	156,930	164,226
March 97	115,958	140,503	132,884	178,152
April 97	104,991	119,576	143,730	160,861
May 97	107,964	123,892	129,957	164,969
June 97	95,748	143,230	133,457	146,155
July 97	88,024	88,526	118,297	134,661
August 97	95,228	138,103	108,873	145,964
September 97	129,137	133,805	117,897	198,087

Chapter 4

VALUATION OF A TECHNOLOGY PATENT—SCOPE, DURATION, AND ROYALTY

4.1. Introduction

The owner of a patent receives value from his patent in several ways. First, the patent owner can exploit the patent themselves to the legal exclusion of others. Second, the patent owner may sell all rights to the patented invention. Third, he may license the patent to others. The value of the patent depends on several factors including its scope of application (economic, technological, and legal), its duration of application (legally limited to a fixed period of time but often constrained by non-infringing superior substitutes, and by its royalty or profit rate. The value of a patent to its owner is clearly constrained by the availability of substitutes. Substitutes limit the value of the patent by affecting the scope, duration, and royalty a patent owner may expect to receive. The determination of these influences on a particular patent is hardly ever a generic exercise. More often than not the value of a patent is determined on a case by case basis.

The purpose of this chapter is to establish the value of a particular patent and illustrate the role of scope, duration, and royalty rate in determining this value. The patent I examine was awarded to the computer manufacturer Acer in 1991.

By 1990, upgrading of personal computers ("PCs") had become common.¹ Consumers wanted a method to protect their investment in PCs against technological obsolescence. In order to meet this need, most major manufacturers developed their own schemes for providing upgrades. In 1991, Acer applied for and received two U.S. patents (5,455,927, con't of Aug. 22, 1991, "dual-

¹Upgrading computers from XT/286 to Pentium machines is discussed in "Tips from the NERD," July 7, 1996 at <http://www.computernerd.com/tips4old.htm>. This author does not recommend upgrading 386 systems because of modest speed improvements and serious residual limitations.

socket") and (5,551,012, con't of Apr. 22, 1991, "single-socket").² Acer has referred to these patents in its product documentation as the "Chip Up" technology.³

The Chip-up technology pertains to upgrading a computer by adding a second (typically faster) control processing unit (CPU) in a socket reserved for the upgrade. A user could upgrade his computer by adding the newer CPU in the reserved socket at any time. As I discuss below, using a primary socket for the main CPU and a secondary socket reserved for the upgrade implies that Acer's system is a dual-socket technique. The sockets that receive the processors in Acer's design are contained on the main circuitry board (motherboard) of the computer.

The Chip-up technology patent specifically was designed to upgrade Intel 80386 (386) computers by a specific method. This patent would not have applied to computer upgrades using non-infringing designs (e.g. modular, daughterboard, card, cartridge, overclocking, etc.) and would not have applied to upgrades of Intel 80486 (486) and Intel 80586 (Pentium) computers. I address these points in turn. First I discuss non-infringing designs which provide 386 computer owners with the ability to upgrade their computers. Second, I discuss upgrades of 486 and Pentium computers.⁴

The value of the Chip-up technology is demonstrated to be small due to the availability of substitutes and the limited applicable scope. The remainder of this chapter is divided into several sections. In Section 4.2, I discuss the scope of the Acer patent. In Section 4.3, I discuss duration issues. In Section 4.4, I discuss the reasonable royalty rate for the Chip-up technology. In Section 4.5, I present my conclusions.

4.2. Scope Limitations

Upgrading a computer might involve enlarging disk drive storage, adding Random Access Memory (RAM) or cache-memory to the motherboard, adding a math coprocessor to speed up numerical calculations or changing the 386/387 pair to specialized chips for computer aided design. A consumer could also upgrade a computer by overclocking the existing CPU.⁵

²The validity of the Acer patents has not been established to the best of my knowledge. Each Acer patent refers to substantial prior art. Additionally, Intel's patent (5,410,726, con't of Apr. 18, 1991, dual-socket upgrades) and Advanced Logic Research's patent (5,297,277, August 2, 1989, automatic enablement) and Advanced Micro Devices' (4,967,346, March 14, 1988, automatic enablement) might establish prior art relevant to the Acer patents and were not cited by the patent office in connection with the Acer patent.

³The name "Chip Up" is apparently trademarked (Trademark 1,730,271, 11/3/92, Trademark Register, 39th Edition, 1997).

⁴I refer to the Intel 80386 CPU as a 386 and the Intel 80486 CPU as a 486. The math coprocessor versions of these chips have model numbers ending in 7. Thus, the Intel 80387 coprocessor (387) is the math coprocessor for the 386.

⁵Overclocking is explained in <http://www.sysoopt.com/ocexp.html>.

I mention these non-CPU upgrade paths because different consumers would find some upgrades more advantageous than others.⁶ For instance, many applications were slowed by poor hard disk performance or insufficient disk storage space. Other applications operated slowly due to too little RAM in the computer. Some consumers did require a speed increase in their CPU. However, independent of processor speed, the preferred upgrade path for many of these consumers was a non-CPU upgrade. For example, for a great many consumers, a coprocessor to do mathematical calculations would have been the most logical upgrade. In fact, most 386 motherboards contained one or two empty sockets for coprocessors. In most cases, these coprocessors were used for arithmetic processor upgrades (adding a 387 or Weitek chip). However, in some cases the sockets were used for other CPU upgrades (e.g. the Intel CAD pair⁷

4.2.1. The Variety of Computer Upgrade Technologies

Since the first generation of IBM PCs, it was common practice to use empty sockets for motherboard upgrades (e.g. the 8087 paired with the 8086, the 80287 paired with the 80286 and the 80387 or the Weitek paired with the 80386). Replacing existing chips with higher speed versions or overclocking (running a CPU at a higher speed than it was rated) were other common upgrade paths for consumers who did not choose a full "generation skip." A generation skip is defined as a movement between families of CPUs such as upgrading from a 386 to a 486. It is a generation skip upgrade that Acer's design contemplated.

The Acer design was not the only available design to accomplish a generation skip upgrade. Some manufacturers used a modular upgrade that utilized a module or separate component on which the CPU and its replacement could be interchanged in the computer. Some manufacturers relied on cards which plugged into the computer's bus, often using proprietary high-speed bus interfaces to maximize the computer's throughput. Other manufacturers relied on third-party daughterboard upgrades that required removing the existing CPU and fitting a small upgrade component into the old socket (this is known as the single-socket upgrade method). Still others adopted a technology such as Acer's that left an additional empty socket for the upgrade.

⁶Upgrading a PC (especially a motherboard upgrade) was never considered an easy process. For instance, *Byte* (April 1, 1991, Vol. 16, No. 14, pp. 283–286), noted that replacing a CPU is a "nightmare," recommending only "the most grizzled hardware veterans" attempt it. Problems with user upgrades include breaking pins on the old or new CPU, destroying a part with static discharge, BIOS incompatibilities, failure of the system to function after the upgrade, software incompatibilities, etc. These issues made CPU upgrades a non user-friendly task. Consequently, the market for upgrades was always a small one.

⁷See e.g. "Intel Coprocessor to Boost CAD performance on 386 PCs", (*PC Week*, Feb 17, 1992, Vol. 9, No. 7, p. 30).

Single-socket upgrades have been around since the beginning of personal computers. Daughterboards were designed to replace the existing CPU, in its socket, with upgraded CPUs.⁸

Acer's technique was identical to a coprocessor upgrade in that he left an empty socket for CPU additions to enhance performance. The 8086 system generally left one empty socket for the coprocessor. The 80286 system left one empty socket. The first 386 computers generally left two empty sockets - one for the existing 287 math coprocessor and another for the yet to be released 387 math coprocessor. Most, 386 systems also left two empty sockets. The first empty socket was for a 387 math coprocessor while the second was left for the Weitek coprocessor. Some dual CPU systems provided both I860 (Intel) and 80X86 (80286 and 80386) CPUs on the same motherboard, although only one processor had control at a time. In fact, there was nothing new in Acer's patent with regard to having empty sockets on the motherboard for the permanent or transitory use of a second processor.⁹

Mirecki (1987) discusses a method by which a personal computer's micro-processor is replaced by a more capable one. The author notes that if the original CPU remains, "the accelerator is called a coprocessor. This does not imply that the two CPUs dynamically share the workload. In most cases only one CPU is active at any given time, and the switch between them is performed by the user." This description is precisely the coprocessor solution adopted by Intel for the original PC and for the IBM XT and for all subsequent Intel based products.¹⁰

4.2.2. Coprocessors and Empty Slots

The coprocessor socket, available in nearly all PCs from the 8086 generation onward,¹¹ had full access to the CPU bus and adequate control lines to disable the main CPU.¹²

Intel's "vacant socket" technique for upgrading 486 systems relied solidly on prior art to provide an upgrade for 486 systems and Pentium systems.¹³

⁸Acceleration upgrades for the 8088 are discussed, for instance, in Manildi (1988).

⁹Another example is the Intel hyper-cube. Intel's hyper-cubes use many (16, 64, 256) CPUs in their own sockets in a parallel processing system. A user would upgrade this architecture by adding additional CPUs.

¹⁰Mirecki (1987) also discusses daughterboards and other plug-in single socket coprocessor upgrade solutions.

¹¹The first 8086 was released in the early to mid 1980s.

¹²See *Electronic Engineering Times* (July 1, 1991).

¹³Intel (1997) describes its Intel 486 Processor Upgrades by socket type (<http://www.intel.com/overdrive/unbrand/486.htm>), its Intel 386 systems upgrades at <http://www.intel.com/overdrive/unbrand/386.htm>. Its upgrade for Pentium systems is described at <http://www.intel.com/overdrive/unbrand/pentpsa.htm>. Intel notes that "As the market has transitioned to Pentium processor technology and demand for upgrades based on older technologies has declined, Intel OverDrive processors based on Intel 486 technology have been phased out." The key point is that the newer technology makes the older technology less attractive even as an upgrade.

Intel's upgrade method reused the existing coprocessor socket for coprocessor or CPU upgrades. By contrast, Acer's technique leaves the original CPU and coprocessor sockets and adds an additional socket reserved for the upgrade. Acer's dual-socket approach is therefore minorly different from Intel's approach. Intel's approach reused an existing empty socket reserved for coprocessor upgrades whereas Acer's approach added an additional empty socket.¹⁴

A review of the technology shows that 486 systems and Pentium systems were upgraded by using well established prior art and by single-socket upgrade solutions.¹⁵ It is for this reason that the Acer patent would not apply to any system beyond 386 based systems.¹⁶

4.2.3. Proprietary Upgrade Methods

Most brand manufacturers had their own schemes for computer upgrades. It is true that, in many cases, the consumer was locked into the original manufacturer for the upgrade. However, the consumer nonetheless trusted that the brand manufacturer would provide a working upgrade solution and would support the PC under its original warranties and conditions.

In *Electronic Engineering Times* (1991)¹⁷ the technology known as EISA II was touted as defining an emerging standard for plug-in CPU boards. According to this article, Compaq, AST, ALR, Northgate, NEC, Hyundai, Blackship Computer Systems, Tandon, Acer, Altos, Copam, Arche Technologies, and CAF had all announced proprietary modular ISA and/or EISA systems. Additionally, Alpha Systems, Hauppauge Computer Works, American Megatrends and Fujikama had announced modular upgradeable motherboards. Further, Intel is reported in this article to have introduced a standard modular connector for CPU cards (the P3 interconnect). The article lists five additional motherboard manufacturers that were reported to have products or promises to develop motherboards around the emerging Intel standard. The Intel standard (P3) was *not* a dual-socket design. The *Electronic Engineering Times* article

¹⁴Intel Senior Vice President David House noted that "starting with the new 486SX CPU board, the coprocessor is no longer just a math coprocessor socket—it has become a universal upgrade socket." House also noted that Intel would introduce several products that fit into this socket to increase not only math performance but also overall system performance. House said, "Some systems today have an upgrade path. Today you can insert a 487 math chip to boost math performance. Next year, you'll be given a second upgrade choice: a two-times clock CPU upgrade that will run at 40 MHz internally."

¹⁵See "Has OverDrive Outlived Its Usefulness?" *PC Week* (October 13, 1997) for a discussion of the Overdrive technology which Intel offered to fill its empty socket and one author's view of the motivation to provide this infrequently used upgrade option.

¹⁶Intel's patent on the vacant socket technique is apparently contained at United States Patent 5,410,726, Bouquai, *et al.* "Upgrading the microprocessor of a computer system without removal by placing a second microprocessor in an upgrade socket." Intel's original application date for this patent is April 18, 1991.

¹⁷See *Electronic Engineering Times* (October 28, 1991, Issue 665).

also states that 65 percent of manufacturers were providing a modular upgrade system.¹⁸

4.2.4. Dual-Socket Technology Was Rare

My extensive searches of *Byte* magazine and my review of articles published in other computer magazines from the 1989-1993 period showed few designs for upgradeable systems using the dual-socket approach. This is a very significant finding. If PC manufacturers or motherboard manufacturers used dual-socket upgrade designs and if this were an important selling feature to consumers then I would have expected to see extensive disclosures in product advertisements of this upgrade method. On the contrary, very few manufacturers based an upgrade on the dual-socket technology and none of the major manufacturers other than Acer adopted this technique. For new technologies, such as computer upgrades, the reputation or brand name of the manufacturer is a very important aspect of the purchase decision for the consumer. Without the strong reputation or brand name of a major PC manufacturer, the dual-socket upgrade technique would in and of itself be of limited value as it would play a limited role in promoting consumer sales.

Atman systems used a dual-socket design in 1992 while Lodestar used a dual-socket design during the period 1991-1993 (*Byte* advertisements). Deico's "predator" motherboard used a dual-socket in April 1992, as did the HIPPO upgradeable server motherboard (JYS Enterprise), Blue Star 386/33 system in May 1992, and Poly 3186/Zen system in Feb 1992. Additionally, *PC Sources*, April 1992 discloses that Peach Computer Systems had an upgradeable dual-socket motherboard. Acer America's Acer Power 386SX is disclosed to use a dual-socket upgrade system in *PC World* (January 1992). An earlier product announcement for the Acer 386SX appears in *InfoWorld*.¹⁹ However, this article states that "the new system features an upgrade socket that can accept either the 20 MHz 486SX or the 487SX from Intel." To the extent that Acer's dual-socket upgrade method uses the coprocessor socket, it would appear to be closer to the Intel coprocessor upgrade method and would therefore be covered under Intel's patent.

In sum, Acer Computer and Lodestar Computer were the only manufacturers with any name recognition that sold dual-socket upgradeable systems. Of lesser stature were Deico (motherboards after April 1992), Blue Star (systems

¹⁸An article in *PC Week Buyer's Guide* (July 29, 1991), reports that ALR, AST, Club American, Compaq, Dell, Digital Scientific, Micro Express, Northgate, Reply, Touch Micro Technologies, and Unisys all use a modular architecture. Everex is disclosed to use a daughterboard upgrade system in "20 Top Upgrades-Has Your PC Seen Better Days" and American Megatrends (AMI) is disclosed to use a daughterboard upgrade system in "AMI Offers Upgradeable 386 Motherboard," *InfoWorld* (July 2, 1990, Vol. 12, No. 27, p. 25).

¹⁹See *InfoWorld* (October 14, 1991, Vol. 13, No. 41, p. 28)

after May 1992), Peach Computer Systems (motherboards after April 1992) and Atman Systems (motherboards after September 1992).

Data from StoreBoard reveal that AST, ALR, Compaq, NEC, Hyundai, IBM, and Everex had the highest sales volume among PC manufacturers in late 1991 and 1992. None of these manufacturers used a dual-socket system. Instead, each provided their own modular proprietary upgrades. Other "brand" manufacturers of the day included EPSON, Leading Edge, Hewlett-Packard, Toshiba, and Zenith. Available evidence suggests that their computers provided no upgrade path for their systems and instead relied on after-market designs for customer upgrades. Of the top twenty personal computer companies in unit sales, Acer was the only company to use a dual-socket design for some of its personal computer systems.

As I discussed above, 65 percent of PCs provided upgrade solutions. Of those that I have identified as providing some upgrade path, only two of twenty-three used a dual-socket design.²⁰ Furthermore, the only users of dual-socket designs appear to have been Taiwanese motherboard manufacturers. Approximately eleven percent of U.S. imports of all computer products came from Taiwan (U.S. Department of Commerce and EIA Market Research Department). If Acer's design had been adopted by 50% of Taiwanese manufacturers, then roughly 5.5 percent of the non-branded PC systems of the day would have used the dual-socket design.

4.3. Duration Limitations

In this section, I review the rapid advances of CPU technology, and show how newer technology quickly displaces older technology. Consumers faced with rapid technological changes come to expect significant price decreases in older technology as newer technology is introduced. The consequence of the rapid decline in CPU prices coupled with newer systems with substantial increases in computing power is that the value of the upgrade option embodied in the dual-socket motherboard is diminished over time. Indeed, the second-socket option in the dual-socket motherboard will have its highest value as an upgrade from 386 to 486 technology when the price of 486 CPUs is also at its highest point and expected to decline the most. As the price of 486 CPUs decline, the value a consumer places on a second socket system in a dual-socket

²⁰I refer to approximately 23 systems (whether motherboards or full computers) in the summaries above. I have given less weight to Dataexpert, ATMAN, DEICO, Bluestar, and PEACH and combined them into one manufacturer as their products appear only in mid to late 1992 or 1993. Also, none of the latter were disclosed in the *PC Week Buyers Guide* (July 29, 1991) review article. However, I give them collectively the same weight as Lodestar. Lodestar had systems for sale during the full period 1991-1993. Lodestar also had advertising in *Byte* magazine even though they were missed in the *PC Week Buyer's Guide* table.

motherboard also declines.²¹ This short interval during which consumers place a significant value for the second socket in the dual-socket system, in conjunction with the rapid decline in the 486 personal computer systems price, leads me to conclude that the time interval when dual-socket motherboards would have been desired by consumers was short. This implies that the appropriate or reasonable royalty value for the design would have declined rapidly over time as the option value faded rapidly.

In Table 4.1, I display a selected history of microprocessor introduction and other significant dates from August 1986 through November 1995. In Figure 4.1, I show the timeline for Intel's processors. Figure 4.1 is based on the information presented in Table 4.1.

Table 4.1. Selected Processor History*

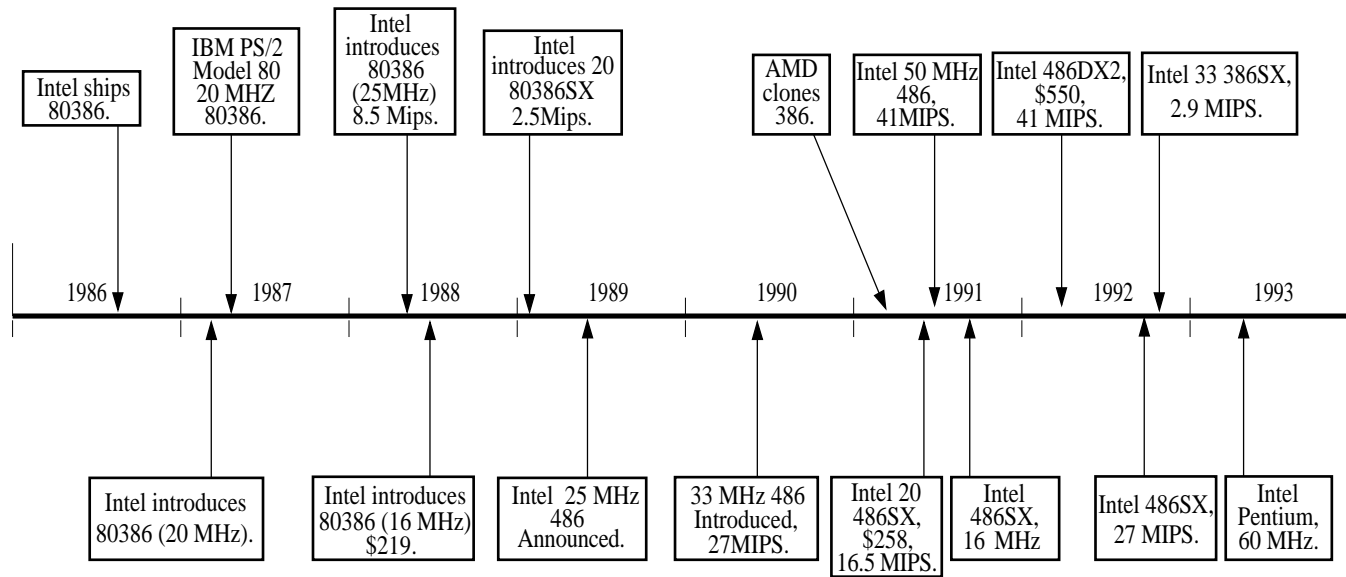
August 1986	Intel Ships 80386
February 1987	Intel Introduces 80386 (20 MHz)
April 1987	IBM PS/2 Model 80 20 MHz 80386
April 1988	Intel Introduces 80386 (25 MHz) 8.5 MIPS
June 1988	Intel Introduces 80386X (16 MHz) \$219
January 1989	Intel Introduces 20 80386X 2.5 MIPS
April 1989	Intel 25 MHz 486 Announced
April 1989	Intel 33 MHz 386 DX Announced
June 1989	First 486 PC Introduced, 25 MHz, 486, and IBM MCA Bus, \$18,000
May 1990	33 MHz 486 Introduced, 27 MIPS
March 1991	AMD Clones 386
April 1991	Intel 20 MHz 486SX, \$258, 16.5 MIPS
June 1991	Intel 50 MHz 486, 41 MIPS
September 1991	Intel 486SX, 16 MHz
March 1992	Intel 486DX2, \$550, 41 MIPS
September 1992	Intel 486SX, 27 MIPS
October 1992	Intel 33 386SX, 2.9 MIPS
March 1993	Intel Pentium, 60 MHz
March 1994	Intel 486 DX4 75, 100 MHz
November 1995	Intel Pentium Pro, 100 MHz

* Source: "Chronology of Events in the History of Microcomputers 1994-1997," and "Processor Timeline," <http://zdnet.com/pcmag/features/cpu/cpu12.htm>.

Importantly, between April 1991 and June 1991, Intel introduced its 50 MHz 486 processor. This occurred only one year after it introduced the previous standard, the 33 MHz 486 processor. The 486 CPU had been announced as early as 1989. But, in 1991, the introduction of the AMD 386 clone and faster

²¹ Moore's Law states that microprocessor CPU speeds will double every 18 months. Thus far this empirical regularity has remained fairly accurate. The consequence of Moore's Law is that the price of a CPU of constant speed should nearly halve every 18 months as well. Often the price declines occur more rapidly.

Figure 4.1: Intel Processor Timeline



Source: Based on Table 4.1

486 CPUs caused the previous technologies' prices to decline rapidly. In Figure 4.2, I display the specifications and speeds of successive generations of Intel processors using the standard iCOMP index.

Each CPU generation brought faster and faster processing power, making older technologies obsolete. This pattern of newer technologies supplanting older technologies and causing their prices to fall is discussed in *InfoWorld* (April 8, 1991).²² The effect on the value of a rapidly changing market of accelerator boards (and, more generally, upgrade options) is discussed in *Personal Computing* (August 1, 1989).²³

4.3.1. A Product Lifecycle Model for CPU Sales

Frank Bass (Bass (1969)), in a pioneering study, introduced a simple and now standard model of consumer durable purchases. In the Bass model, there are innovators (those who buy first and for whom price is not a substantial barrier) and adopters (those who follow based on the innovator's experience with the product). Given a fixed number of potential buyers, Bass shows that product sales will generally rise at an increasing rate, eventually plateau and then fall to zero as the stock of potential buyers declines.

The Bass model has been modified by other authors to include advertising and price. In the case of price, it has been shown that it is optimal for the seller to "skim-price." In skim pricing, the product price, when first introduced, is set at a high level (since demand is inelastic for the innovators) and is eventually lowered to increase the number of potential buyers. Many consumer goods and computer products exhibit this pattern.

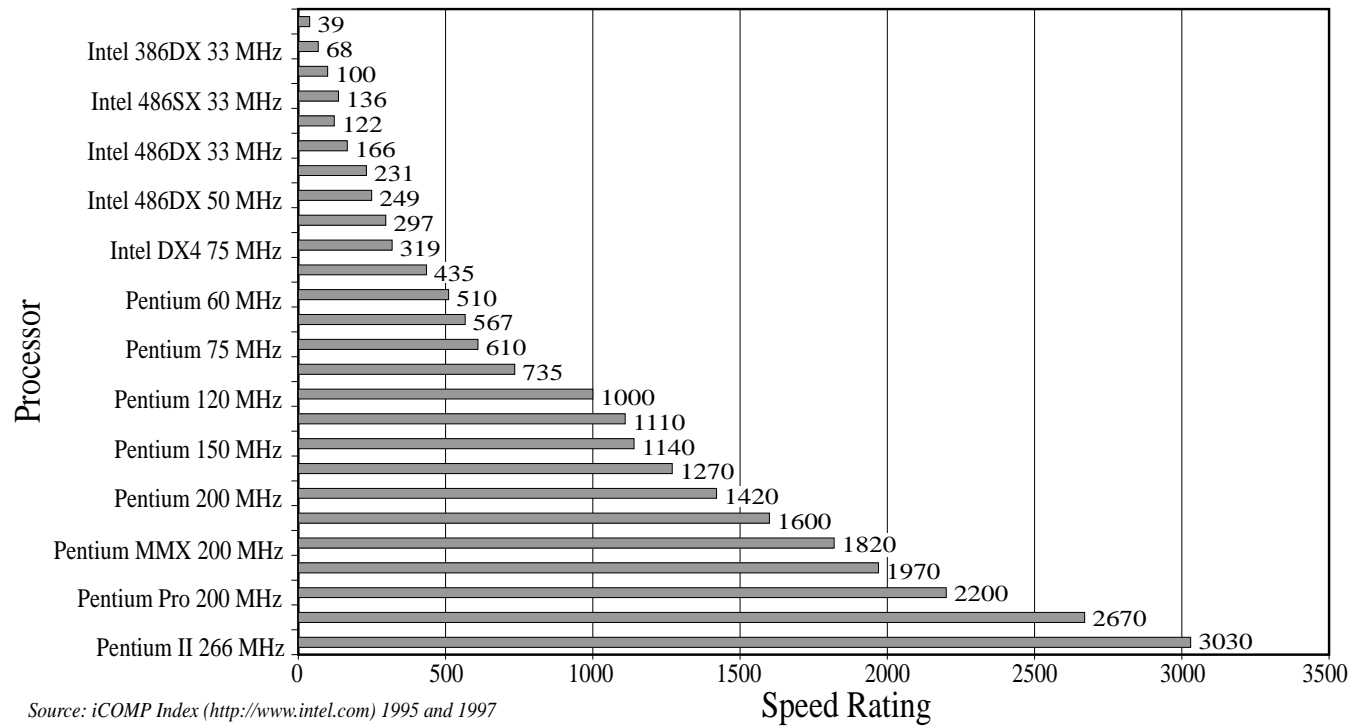
In Norton and Bass (1987) a variant of the product lifecycle model was introduced to explain the sales path for older technology when newer technologies are introduced.²⁴ Figure 4.3 is adapted from the Norton and Bass paper. This figure shows that sales of the first generation product rise and then fall. A similar pattern is seen for the second-generation product. However, the decline in the first generation is affected by introducing the second-generation technology. This is known as the overlapping product lifecycle model. Interestingly, when Norton and Bass applied their model to an actual empirical example they used successive generations of dynamic random access memory

²²See "486SX Will 'Cannibalize' 386 Sales, Leaving the 386SX Market Strong," *InfoWorld*, April 8, 1991, Vol. 13, No. 14, pg. 24.

²³See "What's New in Accelerator Boards—Challenged by Low-Cost Computers, Accelerator Boards have Become Less Expensive and Much Easier to Operate," *Personal Computing*, August 1, 1989, Vol. 13, No. 8, pp. 103-108.

²⁴See Norton and Bass (1987).

Figure 4.2: Speed Comparison of Intel Processors



(DRAM), including the 4K, 16K, 64K, and 256 DRAM. DRAMs are computer components which provide random access memory.²⁵

In Figure 4.4, I use data from InfoCorp for the 1981 to 1991 period to show how the rise in 286 CPU sales caused the sales of 8088/8086 CPUs to decline. I also show how the sales of 386 and 486 CPUs caused the 286 CPU sales to decline.²⁶ In Figure 4.5, I use data from the *Computer Industry Almanac* to show (on a market-share basis) how 486 sales cannibalized the 386 market and how successive chip sales thus followed closely the Norton and Bass overlapping generations model. Inspecting Figure 4.5 reveals that, by 1993, 386 CPU sales and 386-based computer systems had peaked and were declining significantly. Further, the prices for these systems and CPUs had also declined significantly. From various issues of *Byte* magazine, I have seen, for example, that the 386/33 MHz CPU sold for approximately \$345 in 1989. By 1990, its price had fallen to \$225, to \$140 by 1992, and to \$90 by 1993. Price declines also occurred in 486 CPUs and systems. In May 1991, a 486/33 MHz CPU sold for \$1150. By October 1992, its price had declined by two-thirds, to approximately \$340–\$380.

In summary, I find that the market for 386 computers had peaked by 1991. While sales of SX-based systems (low-end, 16-bit interface processors) continued somewhat in 1992 and 1993, 386 DX sales were exhausted by 1993. I also observe that 486-based system sales were on the increase. These observations follow exactly the classical product lifecycle pattern. 486 system prices and CPU prices had their biggest price decline during 1990–1991 and 1991–1992.²⁷

4.4. Limitations in the Royalty's Value

4.4.1. The Presence of Non-Infringing Substitutes Limits the Value of a Patent

Several non-infringing third-party upgrade solutions were available for consumers to purchase even if they did not purchase proprietary upgradeable systems from brand manufacturers. Cyrix provided an upgrade daughterboard, which utilized a single-socket upgrade design.²⁸ TransComputer offered a sim-

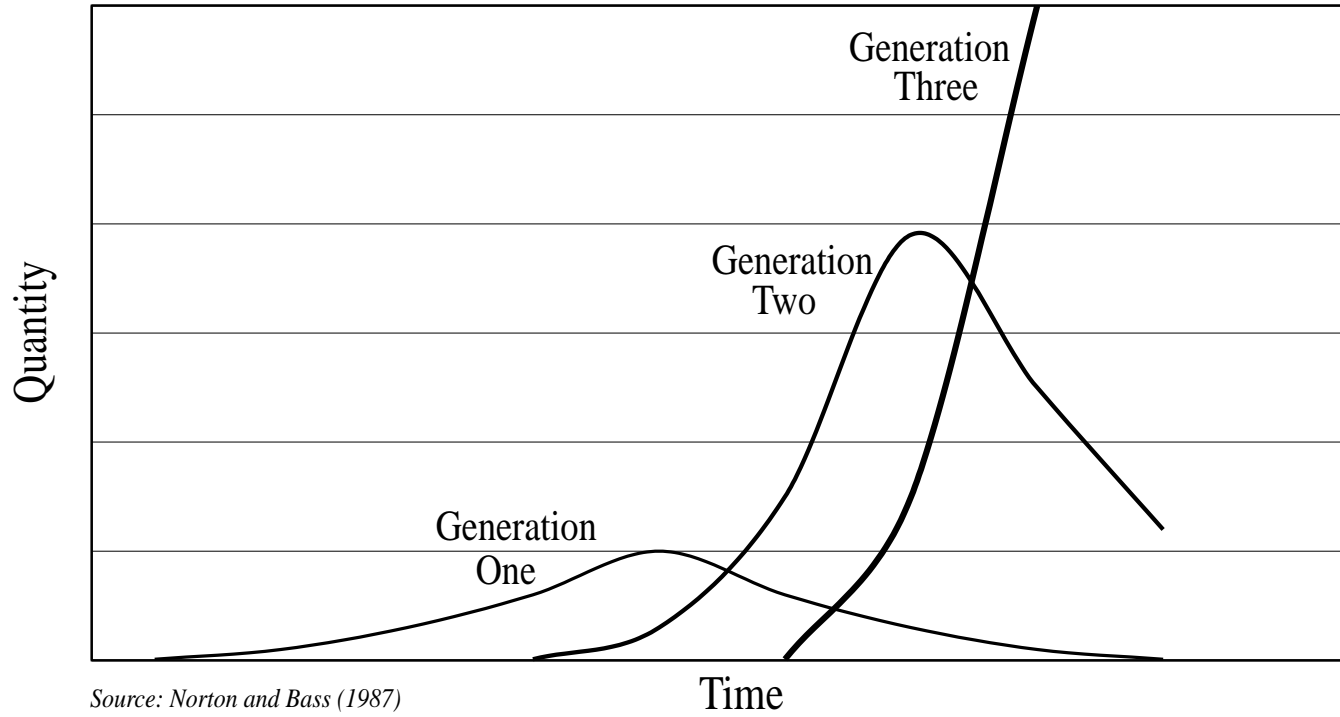
²⁵I note that in terms of technological progress, circa 1987, the high-end DRAM technology contained 256,000 bytes of memory on a chip. The standard today is 128 megabytes per chip at similar lower unit cost.

²⁶An example of the Pentium II processor eroding the markets for the Pentium Pro and Pentium MMX is given in *Intel Microprocessor Forecast*, <http://www.shipanalyst.com/techlib/intel/index.html>.

²⁷For instance, motherboard prices with 486/33 MHz CPUs were approximately \$2,095 in May 1991. These same motherboards were sold for \$849 in May 1992 and for \$499 in May 1993.

²⁸*PC Magazine*, May 31, 1994, Vol. 13, No. 10, p. 52; *PC World*, May 1, 1992, Vol. 10, No. 5, p. 34.

Figure 4.3: A Series of Technological Changes



Source: Norton and Bass (1987)

Figure 4.4: U.S. Microsystems Market, 1981 - 1991

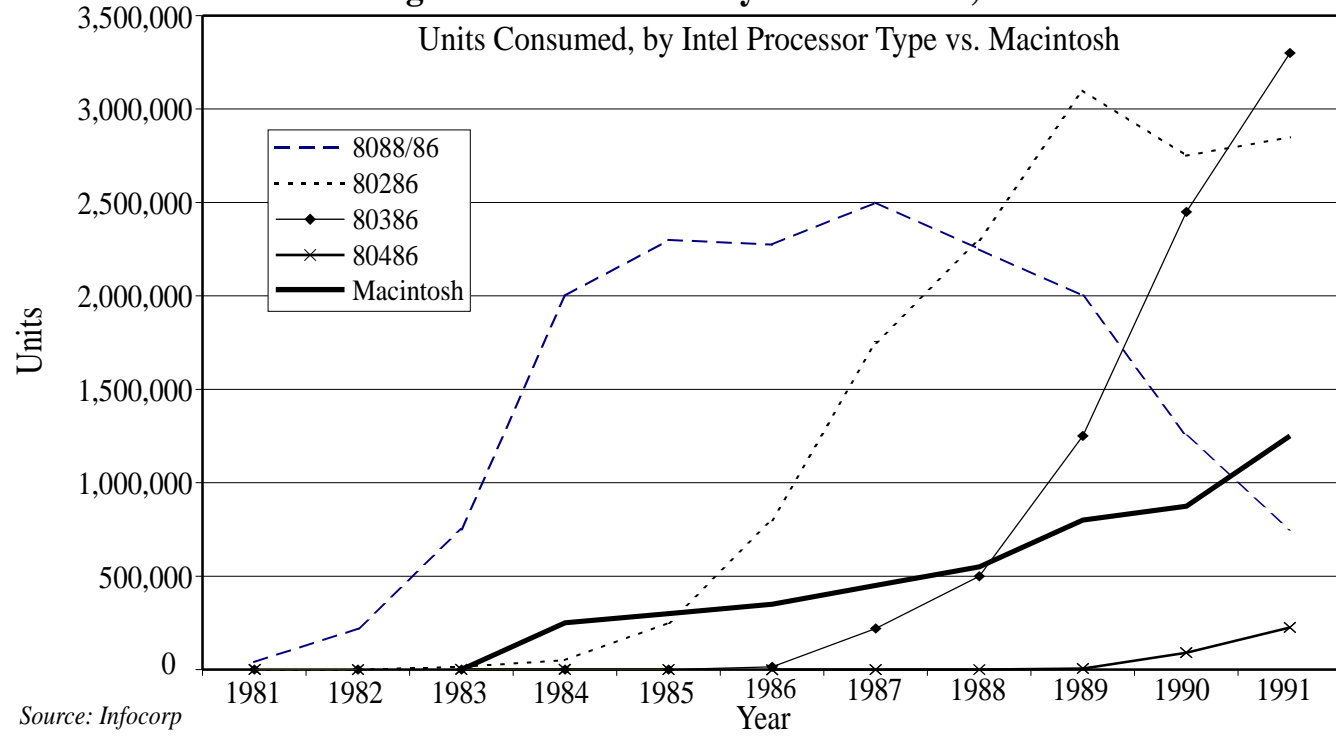
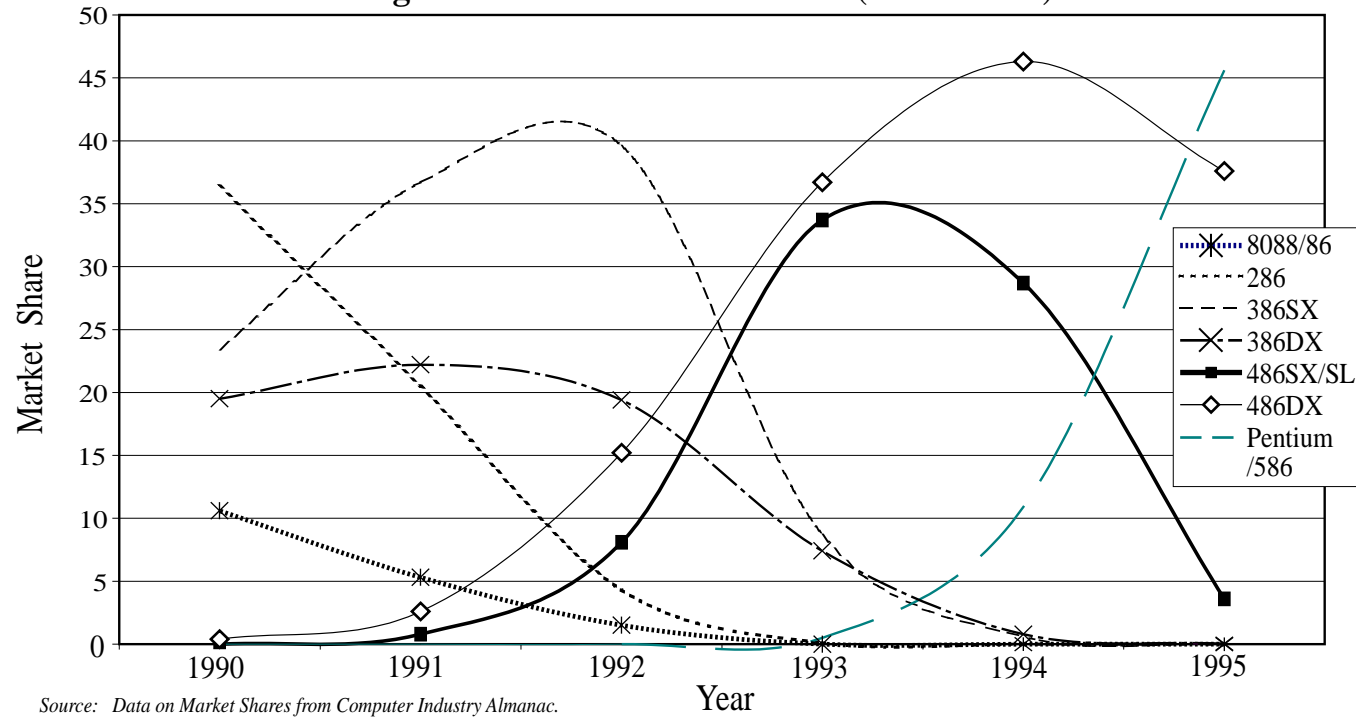


Figure 4.5: CPU Market Share (1990 - 1995)



Source: Data on Market Shares from Computer Industry Almanac. The 8th Annual Computer Industry. The Reference Press. 1996

ilar product.²⁹ Evergreen Rev to 486,³⁰ MicroMaster 386/33,³¹ 486/25 Power Platform³² all provided aftermarket upgrades for the 386 system.

The presence of third-party products, which were of similar quality, as early as 1992 significantly crowded the upgrade market. Many competitors would lead consumers to elastically demand the dual-socket motherboard design. The presence of competition would have led to lower royalty rates.

4.4.2. The Value of a Consumer Upgrade Option

In this section, I analyze how consumers value an upgrade option.³³ This model shows that consumers will value the upgrade option greatest when the price of the next-generation processor is also at its peak. This is also the time when the next generation processor's price is expected to decline most. As the next generation processor's price declines, so does the value of the upgrade option.

The pricing model is based on a number of assumptions that reasonably describe how consumers value computer systems according to their performance. I assume in my model that the 386 processor has a speed of 1 unit (this is a normalization that is not consequential). Consumers can increase their computer's speed by switching to a processor in the 486 generation and receive a speed increase of γ percent. The system costs an amount $\$C$ and the additional cost for the second socket (the upgrade option in the dual-socket design) costs $\$S$ to the consumer. Finally, today's cost of a 486 chip will be denoted by P_0 and the cost in one year (the upgrade period) is \tilde{P}_1 . (The symbol above the price P_1 denotes that P_1 is unknown at time period 0 and is treated as a random variable). I assume that the system cost includes a 386 CPU and that the 386 has little or no residual value in one year's time. Finally, I assume that a given consumer values speed in computers at the rate of λ per unit of speed.

Different consumers will have different values, λ , leading some to purchase a 486 today rather than waiting. Other consumers will choose to wait to get a 486 via an upgrade. I now ask the question: "what is the maximum a consumer would value the second socket in the dual-socket before they would prefer to buy the 486 today rather than later?"

The value today for a risk-neutral consumer who upgrades in one year is:

²⁹*PC Week*, June 18, 1990, Vol. 7, No. 24, p. 16.

³⁰*PC Week*, November 8, 1994, Vol. 13, No. 19, p. 110-151.

³¹*PC World*, May 1, 1990, Vol. 8, No. 5, p. 88.

³²*PC Week*, June 19, 1989, Vol. 6, No. 24, p.5.

³³Similar models for replacement decisions in economics appear in Eckstein and Wolpin (1989).

$$\begin{aligned}
V_u &= \lambda + \lambda(1 + \gamma) - C - S - \tilde{P}_1 \\
&= \lambda + \lambda(1 + \gamma) - C - S + \tilde{r}P_0 - P_0
\end{aligned}$$

where $\tilde{r} = (\tilde{P}_0 - P_1)/P_0$ is rate of decline in 486 prices between year 0 and year 1. The value of upgrading is equal to the value of processing in the first year, λ , plus the value of processing in the second year at higher speed, $\lambda(1 + \gamma)$, less the cost of system C, less the cost of the socket S, less the cost of the buying a new 486 chip at price \tilde{P}_1 in the second year.

The value of buying a 486 today is $V = 2\lambda(1 + \gamma) - C - P_0$. It is equal to two years of value at the higher processor speed less the system cost and less the cost of buying a 486 today. I note that there is no socket cost if the consumer buys the 486 personal computer straight away.³⁴

The maximum a consumer would pay to upgrade would equate these two values since if the consumer paid more for the socket, it would be better for that consumer to upgrade immediately. Therefore the price of the upgrade socket cannot exceed: $S \leq -\lambda\gamma + \tilde{r}P_0$. In words, this equation implies that consumers may pay more for the upgrade socket when the decline in prices \tilde{r} is greatest between 486 chips today and next year. Consumers will pay less for the socket when the speed difference is larger because they will value the speed in the first year to a higher degree, *i.e.*, there is a greater opportunity cost of not using the faster processor.

Consumers who are averse to risk will pay less for the option due to the uncertainty in the rate of decline, \tilde{r} . This can be proven mathematically using a utility function $U(w)$, which is concave in the consumers' wealth level. The upgrade option has expected utility: $EU(W_0 + \tilde{Z})$ where W_0 is initial wealth and $\tilde{Z} = 2\lambda + \lambda\gamma - C + \tilde{r}P_0 - P_0 - S$.

Using a first-order Taylor-series expansion, expected utility is $U(W_0 + \bar{Z}) + 1/2U''(W_0 + \bar{Z})\sigma^2$. Hence, the consumer's expected utility is the utility at the mean wealth level plus a term that lowers expected utility by the randomness in \tilde{Z} as measured by its variance σ^2 . Comparative static analysis demonstrates that the maximum socket value S falls due to the uncertainty term σ^2 . Additionally, a positive time of rate preference will make a consumer prefer the faster speeds today rather than next year and lead to lower socket value. Finally, the consumer who purchases a computer with an upgrade option may not exercise the option to upgrade if 486 prices have not declined enough in the second period. This further reduces the socket's value.

In sum, consumers value the upgrade option most when the prices of the next generation processor are highest and expected to fall the fastest. Based on the evidence of price declines in 486 CPUs reported in the previous section, I

³⁴The salvage value (if any) from not purchasing the 386 CPU would further limit the socket value.

conclude that the upgrade socket would have some value to consumers in 1990 and 1991, lower value in 1992, and virtually no value by 1993. Since the values consumers place on the socket determine its usefulness as a product feature and thereafter determine the reasonable royalty paid by re-sellers, I conclude that royalties would be positive in 1991 and 1992 and would be zero or negligible for a dual-socket, upgrade system by 1993.

4.4.3. Trademark and Patent Valuation

There are several methods to value trademarks and patents. One method is known as the cost approach where the analyst attempts to estimate the cost involved in developing the patented item. In a second approach, termed the comparables approach, the analyst attempts to compare the item at issue to other similar items for which royalties have been established in the marketplace. In the market approach, the analyst analyzes the price premium or profitability of the item and then bases the royalty rate on a split of the profits between the licensor and the licensee to reflect the differential risks experienced by the manufacturer and the patent owner. Another related technique is known as the income approach where the analyst values the royalty based on the income stream produced by ownership of the patented item. Here, I rely on the comparables and market approaches to determine a reasonable royalty for the Acer patent.

First to establish a royalty rate based on comparables, I used the royalty rates in Smith (1997). Smith considers ranges of royalties by the licensee's use of various trademarks (whether pertaining to commercial/industrial products, consumer products and services, food and apparel, or toys/game/entertainment) and by the type of the trademark (whether institutional, corporate, fashion, celebrity, character, or college/sports). In establishing a reasonable royalty for the Acer patent, I used the commercial/industrial use for corporate type trademarks. The royalty values established by Smith (Smith, p. 178) are in the range of 0.5 to 1 percent (royalty rate relative to sales).³⁵

The second method I discuss to establish a reasonable royalty is based on market transactions. Using advertisements by Atman, a manufacturer of motherboards that used a dual-socket upgrade, I analyzed the prices for motherboards in September 1992. Atman's motherboard with a 386/40 MHz CPU then sold for \$200. Using *Byte* magazine, I determined that Intel 386/40 MHz CPU chips were selling for \$80 to \$100 at this time. At the same time, Atman offered the same motherboard with a 486/33 MHz CPU for \$480. During this period, 486/33 MHz CPUs sold in the range of \$340 to \$380.

³⁵Arguably, the patent at issue would have a higher valuation than its trademark value but not exceedingly so given the availability of close substitutes.

Hence, I ascertained that the ATMAN motherboard was worth approximately \$100 to \$140 (the total price less the CPU cost).³⁶ This price is little different from the price that Acer sold its motherboards for in 1992. Furthermore, prices for the motherboard with 386 CPU imply a value of \$100 to \$120 together for the motherboard and the empty socket. Using M to represent the value of the motherboard and S to represent the value of the socket, it follows that $\$100 \leq M \leq \140 and $\$100 \leq M+S \leq \120 , so that $\$0 \leq S \leq \20 . Since both products are sold in market equilibrium, the socket has an implicit value between \$0 and \$20. I take the average value in this range of \$10. Furthermore, manufacturing is not free and involves an incremental cost of \$1 or more. Thus, incremental profits are no more than \$9.00 for the socket at retail. Assuming a retailer or dealer mark-up of approximately 33 percent implies that the profit to the manufacturer is about \$6.00. Of this amount, the licensee will typically receive 25 percent while the licensor will receive the remainder (Goldscheider's Rule). Hence, a patent owner might expect roughly \$1.50 from the sale of each motherboard. Using the 386 motherboard price with CPU as the denominator for calculation of the royalty percentage, the royalty rate would be approximately 0.75 percent. This figure is in the middle of range given by Smith (1997).

4.5. Conclusions

In the conclusion, I calculate Acer's lost royalties from possible patent infringement due to sales made by generic PC manufacturers. Even if generic manufacturers had licensed the "Chip-up" technology, the royalties would have been small given the degree of scope, duration, and royalty rates as we have seen.

Sales data for the top twenty computer manufacturers from 1990 through 1994 is presented in Table 4.2 and was derived from InfoCorp sources. Figure 4.6 presents the sales by manufacturer in 1991 based on the InfoCorp information.

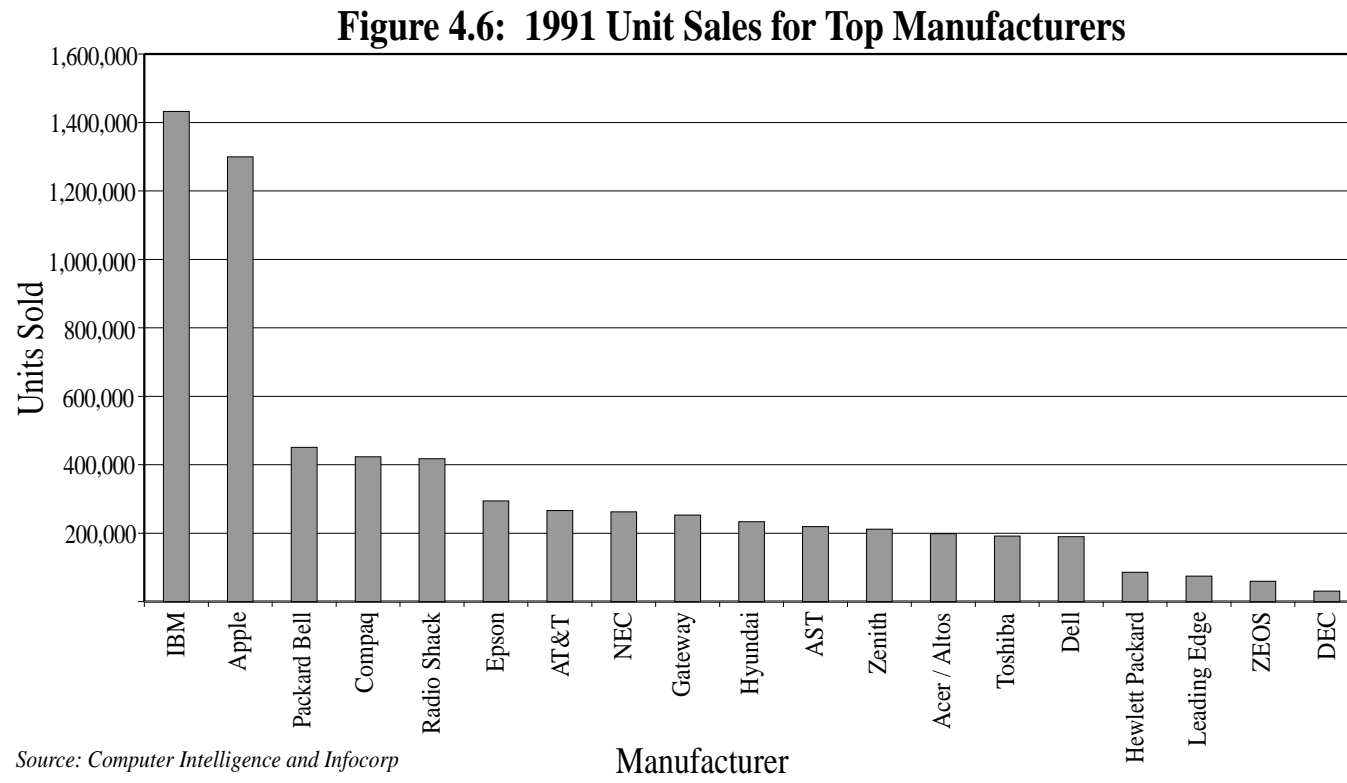
From *Computer Industry Almanac*, I determined that approximately 58% of computer systems during this period were of 386SX or 386DX type. I apply this percentage to Acer's sales, and further assume that all units of 386 type would use the dual-socket design. Using the *InfoCorp* data for 1991 and 1992 sales, I determined that the collective non-top twenty, *i.e.*, no-brand PC sales were 2,437,521 units in 1991 and 2,865,853 units in 1992. I apply a fraction of fifty percent to sales in 1991 since the Acer patent would not have covered all of 1991. I do not include any sales for 1993 or beyond in the calculation

³⁶In this comparison, there is a purposeful asymmetry. The 486 motherboard also contains the empty socket. But the socket is of no value to the consumer. The consumer will only upgrade and will never downgrade the motherboard.

Table 4.2. Top Twenty Computer Manufacturers

1994 Rank	Manufacturer	1990	1991	1992	1993	1994	1994 M.S.
1	Compaq	435,596	423,390	673,242	1,475,205	2,388,736	12.54%
2	Apple	945,086	1,299,398	1,433,287	1,946,291	2,226,736	11.69%
3	IBM	1,516,786	1,432,034	1,296,687	1,726,354	1,874,071	9.84%
4	Packard Bell	303,551	450,519	643,470	998,384	1,370,435	7.19%
5	Gateway 2000	102,000	253,103	659,890	749,369	1,118,982	5.87%
6	Dell	126,745	189,583	421,499	737,495	954,308	5.01%
7	Toshiba	185,008	191,671	206,421	309,489	533,016	2.80%
8	NEC	241,038	262,048	241,206	255,020	491,665	2.58%
9	AST Research	162,021	219,364	262,618	423,058	458,556	2.41%
10	Hewlett Packard	89,259	85,781	84,142	198,937	410,616	2.16%
11	DEC	32,872	30,774	50,543	174,150	322,763	1.69%
12	Acer/Altos	169,490	198,708	236,452	241,864	320,246	1.68%
13	AT&T GIS	218,884	266,251	255,816	271,891	315,507	1.66%
14	Epson	342,992	293,947	261,131	242,588	273,185	1.43%
15	Tandy/Radio Shack	487,637	417,302	403,650	309,814	258,418	1.36%
16	Zenith Data Systems	206,858	211,483	186,834	140,957	170,344	0.89%
17	ZEOS International	36,180	59,427	68,557	59,620	120,047	0.63%
18	Ambra/IBM	—	—	—	47,159	106,691	0.56%
19	Leading Edge	47,388	74,374	145,448	117,165	100,223	0.53%
20	Hyundai Electronics	203,735	233,793	221,459	161,387	93,832	0.49%
	Total (Top 20)	5,853,126	6,592,950	7,752,352	10,586,197	13,908,377	73.01%
	Total (All)	8,016,883	9,030,201	10,618,206	14,499,653	19,049,962	
	Residual	2,163,757	2,437,251	2,865,854	3,913,456	5,141,585	26.99%

Source: *Computer Intelligence* and *InfoCorp* as cited in *Brandweek* (March 20, 1995).



of lost royalties for the reasons discussed above. Using half of 1991 sales and applying a percentage for the 386 machines at issue, I find that $(1,218,761 + 2,865,852) * .58 = 2,369,076$ units were in the 386 class. Based on my review of existing products in the relevant period, I find that only 5.5 percent of these 386 units would have used a dual-socket design. Therefore, royalties would be $\$1.50 * 2,369,076 * .055 = \$195,449$. I view the possibility of collecting this figure from third-party manufacturers as remote. Indeed, there is no evidence that Acer attempted to enforce their patent. The limited applicability of this particular patent and the availability of non-infringing substitutes to others constrained the value of the Chip-up technology.

Chapter 5

STATISTICAL ANALYSIS OF THE ADDITIVE AND MULTIPLICATIVE HYPOTHESES OF MULTIPLE EXPOSURE SYNERGY FOR COHORT AND CASE-CONTROL STUDIES

5.1. Introduction

In epidemiological studies, where there are multiple causes of a particular disease, the issue arises as to whether the multiple causes have a synergistic relationship so that their combined effect is both greater than that of either activity alone, and greater than what one would expect by the sum of their individual risk contributions. Two hypotheses are frequently tested. The first hypothesis states that when the sources of disease act independently, the relative risk of disease, given exposure, is an additive relationship. Thus, the relative risk of dying from cause A adds to the relative risk of dying from cause B to determine the combined relative risk of dying when exposed to both A and B . A second hypothesis states that the relationship between disease and the two causal factors is multiplicative. In this case, the combined risk is the product of the individual risks. Of course synergism is itself a concept that is model dependent. For instance, a lack of synergism in a logit model of risk, as demonstrated by the statistical insignificance of an interaction term, leads to a multiplicative model of relative risk. Consider the following example. Suppose that the probability of dying from a disease depends on two factors, A and B . Let δ_A denote exposure to A , and δ_B denote exposure to B . Suppose further that the probability of dying is logistic and given by:

$$P[D|\delta_A, \delta_B] = 1 / \left(1 + e^{-(X_0\beta_0 + \delta_A X_A\beta_A + \delta_B X_B\beta_B + \delta_A\delta_B X_C)} \right)$$

where X_A , X_B , X_C , and X_0 are vectors of explanatory factors, and β_j are true but unknown coefficient vectors. The presence of the term $\delta_A\delta_B$ allows

for synergism in this model, and specifies that the probability of disease may be different when causal factors A and B are both present. Now, assume that β_C is zero so that there is no synergistic relationship in the model. The relative odds of dying when exposed to both agents are:

$$\begin{aligned} RO_{AB} &= P[D|\delta_A = 1, \delta_B = 1]/P[\bar{D}|\delta_A = 1, \delta_B = 1] \\ &= \exp(X_0\beta_0 + X_A\beta_A + X_B\beta_B) \end{aligned}$$

Similarly, the relative odds of dying when exposed to A alone are:

$$\begin{aligned} RO_A &= P[D|\delta_A = 1, \delta_B = 0]/P[\bar{D}|\delta_A = 1, \delta_B = 0] \\ &= \exp(X_0\beta_0 + X_A\beta_A) \end{aligned}$$

and

$$\begin{aligned} RO_B &= P[D|\delta_A = 0, \delta_B = 1]/P[\bar{D}|\delta_A = 0, \delta_B = 1] \\ &= \exp(X_0\beta_0 + X_B\beta_B) \end{aligned}$$

and the relative odds of dying from background exposure is

$$\begin{aligned} RO_0 &= P[D|\delta_A = 0, \delta_B = 0]/P[\bar{D}|\delta_A = 0, \delta_B = 0] \\ &= \exp(X_0\beta_0) \end{aligned}$$

The relative risk is defined as the ratio of the relative odds between the exposure group and the baseline:

$$\begin{aligned} RR_{AB} &= \left(\frac{P[D|\delta_A = 1, \delta_B = 1]}{P[\bar{D}|\delta_A = 1, \delta_B = 1]} \right) \bigg/ \left(\frac{P[D|\delta_A = 0, \delta_B = 0]}{P[\bar{D}|\delta_A = 0, \delta_B = 0]} \right) \\ &= \left(\frac{P[D|\delta_A = 1, \delta_B = 1]}{P[D|\delta_A = 0, \delta_B = 0]} \right) \bigg/ \left(\frac{P[\bar{D}|\delta_A = 1, \delta_B = 1]}{P[\bar{D}|\delta_A = 0, \delta_B = 0]} \right) \end{aligned}$$

which says that the relative risk from combined exposure is equal to the ratio of the relative odds of dying in the exposed population to the relative odds of dying in the un-exposed population.

$$\begin{aligned} \text{Then } RR_{AB} &= (RO_{AB}/RO_0) \\ &= \exp(X_A\beta_A + X_B\beta_B) \text{ and} \\ RR_A &= (RO_A/RO_0) = \exp(X_A\beta_A), \\ RR_B &= (RO_B/RO_0) = \exp(X_B\beta_B) \end{aligned}$$

Thus $RR_{AB} = RR_A \cdot RR_B$ (the multiplicative model of risk) even though the logit model lacks “synergism” when $\beta_C = 0$.

This chapter considers several methods for determining the relative odds ratio, including the case-control method and the cohort method. The case-control method begins with a group of individuals who have an observed attribute (such as a given disease or death). To the cases are matched a set of control individuals. The matching typically is done at the individual level. For cases and controls, a retrospective determination is made of exposure to one or more contaminants. From the retrospective exposure, prospective odds of becoming a case given exposure are determined.

Cohort studies, by contrast, derive mortality and morbidity rates with reference to an external reference group. The method is based on the idea of comparing the incidence of disease in an exposed cohort to the number expected in a “normal” reference group. Finally, relative risks may be determined using the prevalence method. In prevalence studies it is common to analyze populations that all have some non-zero exposure to a contaminant. For example, in asbestosis studies all subjects, by definition, have some exposure to asbestos. In such cases, the issues of additivity and multiplicativity are not germane because one can consider the separate effect of each causal agent. These situations are nevertheless illuminating in discerning the relative contribution of a second contaminant as it affects the probability of contracting or dying from a disease. Another example is the analysis of the prevalence of a disease attribute (such as pleural plaques) in an exposed population.¹

This chapter focuses on testing the statistical hypotheses of additivity and multiplicativity for the relative risk measures. While other papers have considered the confidence intervals for relative risk measures, no systematic study has been made of the additivity and multiplicativity hypotheses as a matter of statistics. Therefore, while practice in epidemiology has been to say that one or more studies appear to support the multiplicative model, these studies have not, in general, been statistical statements; i.e., statements made with attendant levels of confidence.

This chapter is divided into six sections. In Section 5.2, I discuss the case-control method and Wald type tests for the multiplicative and additive hypotheses, derive and discuss Woolf’s method for determining the variance of log-odds ratios (Woolf (1955)), and discuss maximum likelihood methods for optimization subject to constraints following the methods of Gardner and Munford (1980). In Section 5.3, I discuss other approaches for determining confidence intervals, including Bonferroni Intervals and Monte Carlo simulation. In Section 5.4, I describe various synergy indices and how they relate to tests of

¹A prevalence model may be fitted with a logistic functional form. The outcome variable is usually the presence or absence of a disease characteristic where the explanatory factors will include control variables and an indicator for the level of contaminant. If the cohort provides some level of variance in the level of exposure of both contaminants, an interactive term can be used to test for synergy, even if this does not provide a test of additivity or multiplicativity.

hypotheses for additive and multiplicative statistics. In Section 5.5, I discuss cohort studies and derive hypothesis tests for the additive and multiplicative statistics. In Section 5.6, I present our conclusions.

5.2. Case-Control Studies

I begin with a table of case-control outcomes at differing exposure levels:

Exposure					
	None	A	B	A&B	TOTAL
cases	h_1	h_2	h_3	h_4	h
controls	k_1	k_2	k_3	k_4	k

We next express the row counts as fractions:

Exposure					
	None	A	B	A&B	TOTAL
cases	π_1	π_2	π_3	π_4	1
controls	θ_1	θ_2	θ_3	θ_4	1

where $\hat{\pi}_j = h_j/h$ and $\hat{\theta}_j = k_j/k$ are consistent estimates of the true cell probabilities.

First, we demonstrate that the retrospective odds-ratio from a case control method provides an approximate estimate of the relative risk of being a “case,” given exposure. To prove this, we examine the odds ratio $\pi_4 \cdot \theta_1 / \pi_1 \cdot \theta_4$, although the result clearly generalizes to other cases. I show that, given exposure level $A\&B$, the odds-ratio approximates the relative risk of being a case.

I denote cases as D (death from lung cancer for instance) and \bar{D} a control (death from other causes for instance). The combined exposure $A\&B$ is referred to as E (exposure). A case with no exposure is denoted \bar{E} (no exposure).

The odds-ratio $\pi_4 \cdot \theta_1 / \pi_1 \cdot \theta_4$ is equal to

$$\frac{P[E|D] \cdot P[\bar{E}|\bar{D}]}{P[\bar{E}|D] \cdot P[E|\bar{D}]} \quad (5.1)$$

since $\pi_4 = P[E|D]$, $\pi_1 = P[\bar{E}|D]$, $\theta_4 = P[E|\bar{D}]$, and $\theta_1 = P[\bar{E}|\bar{D}]$

As the notation implies, the probabilities π and θ are conditional probabilities indicating the respective likelihood of having been exposed, given an individual’s case-control status. Of interest is the prospective probability of being a case (i.e., dying) given exposure status.²

²Some research studies have used logit analysis to model the conditional probabilities shown above. This allows the introduction of covariates to provide additional controls in the analysis. For example a logit

Under a simplifying assumption, the odds-ratio approximates the prospective odds:

$$\begin{aligned}
 \frac{P[E|D] \cdot P[\bar{E}|\bar{D}]}{P[\bar{E}|D] \cdot P[E|\bar{D}]} &= \frac{P[E, D]/P[D]}{P[E, \bar{D}]/P[\bar{D}]} \cdot \frac{P[\bar{E}, \bar{D}]/P[\bar{D}]}{P[\bar{E}, D]/P[D]} \\
 &= \frac{P[D|E] \cdot P[E]}{P[D|\bar{E}] \cdot P[\bar{E}]} \cdot \frac{P[\bar{D}|\bar{E}] \cdot P[\bar{E}]}{P[\bar{D}|E] \cdot P[E]} \\
 &= \frac{P[D|E]}{P[D|\bar{E}]} \cdot \frac{P[\bar{D}|\bar{E}]}{P[\bar{D}|E]} \\
 &\doteq \frac{P[D|E]}{P[D|\bar{E}]} \tag{5.2}
 \end{aligned}$$

where the approximation results from the observation that $P[\bar{D}|\bar{E}] / P[\bar{D}|E]$ is close to one. Case control studies are useful as they provide estimates of the odds $P[E]/P[\bar{E}]$; i.e., the relative odds of exposure. The relative odds of being a case $P[D]/P[\bar{D}]$ are irrelevant as they are set by the researcher in the design. They do, however, have an influence on the confidence of the results.

5.2.1. Tests for Case-Control Studies—Multiplicative Case

The relative risk (prospective) of dying given exposure to contaminant A is $\pi_2\theta_1/\pi_1\theta_2$. The relative risk of dying given exposure to contaminant B is $\pi_3\theta_1/\pi_1\theta_3$. The relative risk of dying if exposed to both contaminants is $\pi_4\theta_1/\pi_1\theta_4$. The multiplicative hypothesis states that $RR_{A\&B} = RR_A \cdot RR_B$ so that:

$$\pi_4\theta_1/\pi_1\theta_4 = (\pi_2\theta_1/\pi_1\theta_2) \cdot (\pi_3\theta_1/\pi_1\theta_3)$$

Taking logarithms, this becomes:

$$\begin{aligned}
 \log \pi_4 + \log \theta_1 - \log \pi_1 - \log \theta_4 - \log \pi_2 - \log \theta_1 + \log \pi_1 + \log \theta_2 - \log \pi_3 - \log \theta_1 \\
 + \log \pi_1 + \log \theta_3 = 0
 \end{aligned}$$

This may be rewritten as

$$\log (\pi_4\theta_2\theta_3\pi_1) - \log (\theta_4\pi_2\pi_3\theta_1) = 0$$

or

$$M = (\log \pi_1 - \log \pi_2 - \log \pi_3 + \log \pi_4) - (\log \theta_1 - \log \theta_2 - \log \theta_3 + \log \theta_4) = 0$$

model may be used to specify the conditional probabilities: $P[A|D]$, $P[\bar{A}|D]$, $P[A|\bar{D}]$, and $P[\bar{A}|\bar{D}]$. A specification of such a model was illustrated in the introduction. The presence of additional covariates complicates the analysis presented below as the variances and covariances become dependent on the assumed probability model and on the precision of the parameter estimation.

A consistent estimate of this statistic is obtained by replacing π_j and θ_j with $\hat{\pi}_j$ and $\hat{\theta}_j$.

Deriving the variance of the resulting statistic is complicated by the fact that $h_1, h_2, h_3,$ and h_4 form a multinomial probability distribution. Similarly $k_1, k_2, k_3,$ and k_4 form a multinomial probability distribution, but one which is independent of the joint distribution of the h_j by assumption.

To derive the joint distribution of the $\log \pi_j$ and $\log \theta_j$, I begin with results for the joint distribution of the h_j . Similar results hold for the outcome of the k_j . For notational simplicity I present the results using a common symbol n_j where $n_1 + n_2 + n_3 + n_4 = n$.

Lemma 1 *Let $\delta_{jt} = 1$ if outcome j is realized in observation t . The probability that $\delta_{jt} = 1$ is denoted π_j . Let n_j denote the total number of outcome j 's that are observed in the sample of n independent draws, with*

$$\pi_1 + \pi_2 + \pi_3 + \pi_4 = 1, \quad n_1 + n_2 + n_3 + n_4 = n,$$

$$n_j = \sum_{t=1}^n \delta_{jt}, \quad n = \sum_{t=1}^n (\delta_{1t} + \delta_{2t} + \delta_{3t} + \delta_{4t}) = \sum_{t=1}^n 1.$$

Then $E(n_j) = n\pi_j$, $V(n_j) = n\pi_j(1 - \pi_j)$, and $\text{cov}(n_j, n_k) = -n\pi_j\pi_k$ for $j \neq k$.

Proof: $n_j = \sum_{t=1}^n \delta_{jt}$ implies $E(n_j) = \sum_{t=1}^n E(\delta_{jt}) = n\pi_j$ since $E(\delta_j) = 1 \cdot \pi_j + 0 \cdot (1 - \pi_j)$. Next $V(n_j) = \sum_{t=1}^n V(\delta_{jt})$. But $V(\delta_{jt}) = E(\delta_{jt}) - E(\delta_{jt})^2 = \pi_j - \pi_j^2 = \pi_j(1 - \pi_j)$. Hence $V(n_j) = n\pi_j(1 - \pi_j)$. Finally $\text{cov}(n_j, n_k) = E[(n_j - n\pi_j)(n_k - n\pi_k)] = E(n_j n_k) - n\pi_j n\pi_k - n\pi_k n\pi_j + n^2\pi_j\pi_k = E(n_j n_k) - n^2\pi_j\pi_k$. Now

$$E(n_j n_k) = E\left[\left(\sum_t \delta_{jt}\right)\left(\sum_t \delta_{kt}\right)\right] = E\left[\sum_t \delta_{jt}\delta_{kt} + \sum_{t \neq s} \delta_{jt}\delta_{ks}\right].$$

But $\delta_{jt}\delta_{kt} = 0$ if $j \neq k$ in observation t (only one unique outcome is realized in each trial) so that the first sum is exactly zero. The second sum consists of $(n^2 - n)$ terms, which are the products of independent random variables (since δ_{jt} and δ_{ks} are independent when $t \neq s$). The expectation of each term in the second sum is $E(\delta_{jt}\delta_{ks}) = \pi_j\pi_k$.

Hence $E(n_j n_k) = (n^2 - n)\pi_j\pi_k$. Combining these results we obtain

$$\begin{aligned} \text{cov}(n_j, n_k) &= (n^2 - n)\pi_j\pi_k - n^2\pi_j\pi_k \\ &= -n\pi_j\pi_k \end{aligned} \tag{5.3}$$

■

Combining these results into the variance covariance matrix for n_j we obtain:

$$E \begin{pmatrix} n_1 \\ n_2 \\ n_3 \\ n_4 \end{pmatrix} = n \begin{pmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \\ \pi_4 \end{pmatrix}$$

and

$$\begin{aligned} V \begin{pmatrix} n_1 \\ n_2 \\ n_3 \\ n_4 \end{pmatrix} &= n \begin{pmatrix} \pi_1(1-\pi_1) & -\pi_1\pi_2 & -\pi_1\pi_3 & -\pi_1\pi_4 \\ -\pi_2\pi_1 & \pi_2(1-\pi_2) & -\pi_2\pi_3 & -\pi_2\pi_4 \\ -\pi_3\pi_1 & -\pi_3\pi_2 & \pi_3(1-\pi_3) & -\pi_3\pi_4 \\ -\pi_4\pi_1 & -\pi_4\pi_2 & -\pi_4\pi_3 & \pi_4(1-\pi_4) \end{pmatrix} \\ &= n(I - \phi\phi') \end{aligned}$$

where $\phi = (\sqrt{\pi_1} \quad \sqrt{\pi_2} \quad \sqrt{\pi_3} \quad \sqrt{\pi_4})'$.

To derive the variance-covariance matrix for $\log \hat{\pi}_j = \log(n_j/n)$, we use a Taylor's series expansion to first-order for the logarithm. Then

$$\log \hat{\pi}_j \doteq \log \pi_j + \frac{1}{\pi_j}(\hat{\pi}_j - \pi_j)$$

where we have evaluated the Taylor's series expansion around the true but unknown π_j . Then:

$$\log \begin{pmatrix} \hat{\pi}_1 \\ \hat{\pi}_2 \\ \hat{\pi}_3 \\ \hat{\pi}_4 \end{pmatrix} = \begin{pmatrix} \log \pi_1 \\ \log \pi_2 \\ \log \pi_3 \\ \log \pi_4 \end{pmatrix} + \begin{pmatrix} 1/\pi_1 & & & 0 \\ & 1/\pi_2 & & \\ 0 & & 1/\pi_3 & \\ & & & 1/\pi_4 \end{pmatrix} \begin{pmatrix} (\hat{\pi}_1 - \pi_1) \\ (\hat{\pi}_2 - \pi_2) \\ (\hat{\pi}_3 - \pi_3) \\ (\hat{\pi}_4 - \pi_4) \end{pmatrix}$$

Hence

$$\begin{aligned}
V(\log \hat{\pi}_j) &= \begin{pmatrix} 1/\pi_1 & 0 & & \\ 0 & 1/\pi_2 & & \\ & & 1/\pi_3 & \\ & & & 1/\pi_4 \end{pmatrix} Var(\hat{\pi}_j - \pi_j) \\
&= \frac{1}{n} \begin{pmatrix} 1/\pi_1 & 0 & & \\ 0 & 1/\pi_2 & & \\ & & 1/\pi_3 & \\ & & & 1/\pi_4 \end{pmatrix} \cdot \\
&\quad \begin{pmatrix} \pi_1(1 - \pi_1) & -\pi_1\pi_2 & -\pi_1\pi_3 & -\pi_1\pi_4 \\ -\pi_2\pi_1 & \pi_2(1 - \pi_2) & -\pi_2\pi_3 & -\pi_2\pi_4 \\ -\pi_3\pi_1 & -\pi_3\pi_2 & \pi_3(1 - \pi_3) & -\pi_3\pi_4 \\ -\pi_4\pi_1 & -\pi_4\pi_2 & -\pi_4\pi_3 & \pi_4(1 - \pi_4) \end{pmatrix} \\
&\quad \begin{pmatrix} 1/\pi_1 & 0 & & \\ 0 & 1/\pi_2 & & \\ & & 1/\pi_3 & \\ & & & 1/\pi_4 \end{pmatrix}' \quad (5.4)
\end{aligned}$$

since $Var(\hat{\pi}_j) = \frac{1}{n^2} Var(n_j)$.

Theorem 2 For the multiplicative hypothesis,

$$Var(M) = \left(\frac{1}{h_1} + \frac{1}{h_2} + \frac{1}{h_3} + \frac{1}{h_4} \right) + \left(\frac{1}{k_1} + \frac{1}{k_2} + \frac{1}{k_3} + \frac{1}{k_4} \right)$$

Proof: The multiplicative hypothesis may be written as

$$M = \begin{pmatrix} 1 & -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} \log \hat{\pi}_1 \\ \log \hat{\pi}_2 \\ \log \hat{\pi}_3 \\ \log \hat{\pi}_4 \end{pmatrix} - \begin{pmatrix} 1 & -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} \log \hat{\theta}_1 \\ \log \hat{\theta}_2 \\ \log \hat{\theta}_3 \\ \log \hat{\theta}_4 \end{pmatrix}$$

Hence $Var(M) = \frac{1}{h} (1 \ -1 \ -1 \ 1)$.

$$\begin{aligned}
& \begin{pmatrix} (1 - \pi_1)/\pi_1 & -1 & -1 & -1 \\ -1 & (1 - \pi_2)/\pi_2 & -1 & -1 \\ -1 & -1 & (1 - \pi_3)/\pi_3 & -1 \\ -1 & -1 & -1 & (1 - \pi_4)/\pi_4 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} \\
& \frac{1}{k}(1 \ -1 \ -1 \ 1) \cdot \\
& \begin{pmatrix} (1 - \theta_1)/\theta_1 & -1 & -1 & -1 \\ -1 & (1 - \theta_2)/\theta_2 & -1 & -1 \\ -1 & -1 & (1 - \theta_3)/\theta_3 & -1 \\ -1 & -1 & -1 & (1 - \theta_4)/\theta_4 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} \\
& = \frac{1}{h} [(1)[(1 - \pi_1)/\pi_1 + 1 + 1 - 1] + (-1)[-1 - (1 - \pi_2)/\pi_2 + 1 - 1] + \\
& \quad (-1)[-1 + 1 - (1 - \pi_3)/\pi_3 - 1] + (1)[-1 + 1 + 1 + (1 - \pi_4)/\pi_4] \\
& \quad + \text{similar terms in } \theta \\
& = \frac{1}{h} \left[\frac{(1 - \pi_1)}{\pi_1} + 1 + \frac{(1 - \pi_2)}{\pi_2} + 1 + \frac{(1 - \pi_3)}{\pi_3} + 1 + \frac{(1 - \pi_4)}{\pi_4} + 1 \right] \\
& \quad + \text{similar terms in } \theta \\
& = \frac{1}{h} \left[\frac{1}{\pi_1} + \frac{1}{\pi_2} + \frac{1}{\pi_3} + \frac{1}{\pi_4} \right] + \text{similar terms in } \theta \tag{5.5}
\end{aligned}$$

Hence

$$\begin{aligned}
Var(M) & = \left[\left(\frac{1}{h\pi_1} + \frac{1}{h\pi_2} + \frac{1}{h\pi_3} + \frac{1}{h\pi_4} \right) + \left(\frac{1}{k\theta_1} + \frac{1}{k\theta_2} + \frac{1}{k\theta_3} + \frac{1}{k\theta_4} \right) \right] \\
& = \left(\frac{1}{h_1} + \frac{1}{h_2} + \frac{1}{h_3} + \frac{1}{h_4} \right) + \left(\frac{1}{k_1} + \frac{1}{k_2} + \frac{1}{k_3} + \frac{1}{k_4} \right). \tag{5.6}
\end{aligned}$$

■

5.2.2. Woolf's Method

A similar result for the variance of a log odds-ratio itself is derived as follows. Consider $\log(\pi_4\theta_1/\pi_1\theta_4)$, the log-odds ratio for the relative risk at the combined exposure level in a case control study. We have

$$\log RR_{A\&B} = \log(\pi_4\theta_1/\pi_1\theta_4) = \log(\pi_4/\pi_1) - \log(\theta_4/\theta_1)$$

Next, without loss of generality, assume that π_1 and π_4 have been normalized so that $\pi_1 + \pi_4 = 1$ (This may be accomplished by setting $\pi'_1 = \pi_1/(\pi_1 + \pi_4)$)

and $\pi'_4 = \pi_4/(\pi_1 + \pi_4)$. Now $\pi'_1 + \pi'_4 = 1$ and the log odds-ratio remains unchanged since

$$\log RR_{A\&B} = \log (\pi'_4\theta'_1/\pi'_1\theta'_4) = \log (\pi_4\theta_1/\pi_1\theta_4).$$

The expression for $\log (\pi_4/\pi_1)$ is in the form $\log \left(\frac{\rho}{1-\rho} \right)$ where $\rho = \pi_4$ and $(1 - \rho) = \pi_1$. A Taylor's series expansion of $\log \left(\frac{\rho}{1-\rho} \right)$ demonstrates that:

$$\begin{aligned} \log \frac{\rho}{1-\rho} &= \log \frac{\rho_0}{1-\rho_0} + \frac{1-\rho}{\rho} \left[\frac{(1-\rho) + \rho}{(1-\rho)^2} \right] \bigg|_{\rho_0} \cdot (\rho - \rho_0) \\ &= \log \frac{\rho_0}{1-\rho_0} + \frac{\rho - \rho_0}{\rho_0(1-\rho_0)}. \end{aligned} \quad (5.7)$$

Next

$$\begin{aligned} \text{Var} \left(\log \frac{\hat{\rho}}{(1-\hat{\rho})} \right) &= \left(\frac{1}{\hat{\rho}(1-\hat{\rho})} \right)^2 \frac{\hat{\rho}(1-\hat{\rho})}{N} \\ &= \frac{1}{N\hat{\rho}(1-\hat{\rho})} \end{aligned} \quad (5.8)$$

where $\hat{\rho} = \frac{1}{N} \sum_{t=1}^N \delta_t$ is the unbiased estimator of ρ , $E(\hat{\rho}) = \rho$ and $\text{Var}(\hat{\rho}) = \hat{\rho}(1-\hat{\rho})/N$ and N is the number of independent trials resulting in $\sum_{t=1}^N \delta_t$ exposure cases (as compared to non-exposure cases). Similar expressions follow for the theta distribution. Now

$$\begin{aligned} \text{Var}(\log (\hat{\pi}_4/\hat{\pi}_1)) &= \frac{1}{h\hat{\pi}_4\hat{\pi}_1} = \frac{h}{(h\hat{\pi}_4)(h\hat{\pi}_1)} \\ &= \frac{h_1 + h_4}{h_1h_4} \\ &= \frac{1}{h_4} + \frac{1}{h_1}. \end{aligned} \quad (5.9)$$

Then

$$\text{Var}(\log RR_{A\&B}) = \frac{1}{h_1} + \frac{1}{h_4} + \frac{1}{k_1} + \frac{1}{k_4}$$

Note that the repeated application of this result (assuming independence) to the multiplicative hypothesis would not produce the correct result in a case control setting because $RR_{A\&B}$, RR_A and RR_B are mutually correlated.

This result is also known as Woolf's method, and is sometimes written

$$\text{Var} \left(\log \frac{AD}{BC} \right) = \frac{1}{A} + \frac{1}{B} + \frac{1}{C} + \frac{1}{D}$$

where $RR = AD/BC$ and A denotes the number of cases with exposure, B denotes cases without exposure, C denotes controls with exposure, and D denotes controls without exposure.

It is also possible to derive the covariances of the relative risk measures. Consider $RR_A = \pi_2\theta_1/\pi_1\theta_2$ and $RR_{A\&B} = \pi_4\theta_1/\pi_1\theta_4$. Then

$$\log RR_A = (\log \pi_2 - \log \pi_1) - (\log \theta_2 - \log \theta_1)$$

and

$$\log RR_{A\&B} = (\log \pi_4 - \log \pi_1) - (\log \theta_4 - \log \theta_1).$$

Clearly, these are correlated because of the common components. Consider the π components first (analogous results apply to the θ components). Recall that $\text{Var}(\log \hat{\pi})$

$$= \frac{1}{h} \begin{pmatrix} (1 - \pi_1)/\pi_1 & -1 & -1 & -1 \\ -1 & (1 - \pi_2)/\pi_2 & -1 & -1 \\ -1 & -1 & (1 - \pi_3)/\pi_3 & -1 \\ -1 & -1 & -1 & (1 - \pi_4)/\pi_4 \end{pmatrix}$$

But $\log \pi_2 - \log \pi_1 = (-1, 1, 0, 0) \begin{bmatrix} \log \pi_1 \\ \log \pi_2 \\ \log \pi_3 \\ \log \pi_4 \end{bmatrix}$ so that

$$\begin{aligned} & \text{Var}(\log \pi_2 - \log \pi_1) \\ &= \left(\frac{1}{h}\right) (-1, 1, 0, 0) \cdot \\ & \quad \begin{pmatrix} (1 - \pi_1)/\pi_1 & -1 & -1 & -1 \\ -1 & (1 - \pi_2)/\pi_2 & -1 & -1 \\ -1 & -1 & (1 - \pi_3)/\pi_3 & -1 \\ -1 & -1 & -1 & (1 - \pi_4)/\pi_4 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \\ &= \left(\frac{1}{h}\right) (-1, 1, 0, 0) \begin{bmatrix} -(1 - \pi_1)/\pi_1 - 1 \\ 1 + (1 - \pi_2)/\pi_2 \\ 0 \\ 0 \end{bmatrix} \\ &= \frac{1}{h} \left(\frac{(1 - \pi_1)}{\pi_1} + 1 + \frac{(1 - \pi_2)}{\pi_2} + 1 \right) \\ &= \left(\frac{1}{h}\right) \left(\frac{1}{\pi_1}\right) + \left(\frac{1}{h}\right) \left(\frac{1}{\pi_2}\right) = \frac{1}{h_1} + \frac{1}{h_2} \end{aligned} \tag{5.10}$$

Combining this with the analogous result for $\log \theta_2 - \log \theta_1$, we obtain:

$$\text{Var}(\log RR_A) = \frac{1}{h_1} + \frac{1}{h_2} + \frac{1}{k_1} + \frac{1}{k_2}$$

This is exactly the Woolf result shown above. Similarly:

$$\text{Var}(\log RR_{A\&B}) = \frac{1}{h_1} + \frac{1}{h_4} + \frac{1}{k_1} + \frac{1}{k_4}$$

and

$$\text{Var}(\log RR_B) = \frac{1}{h_1} + \frac{1}{h_3} + \frac{1}{k_1} + \frac{1}{k_3}$$

Next consider the covariance between $\log RR_A$ and $\log RR_{A\&B}$. Again, we consider the π terms first. Using the fact that $\text{cov}(t'x, s'x) = t'\text{Var}(x)s$ for conformable column vectors, we have, (for the π terms only)

$$\begin{aligned} \text{cov}[\log RR_A, \log RR_B] &= \left(\frac{1}{h}\right) \begin{pmatrix} -1 & 1 & 0 & 0 \end{pmatrix} \cdot \\ &\begin{pmatrix} (1-\pi_1)/\pi_1 & -1 & -1 & -1 \\ -1 & (1-\pi_2)/\pi_2 & -1 & -1 \\ -1 & -1 & (1-\pi_3)/\pi_3 & -1 \\ -1 & -1 & -1 & (1-\pi_4)/\pi_4 \end{pmatrix} \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \\ &= \left(\frac{1}{h}\right) \begin{pmatrix} -1 & 1 & 0 & 0 \end{pmatrix} \begin{bmatrix} -(1-\pi_1)/\pi_1 - 1 \\ 1 - 1 \\ 1 - 1 \\ 1 + (1-\pi_4)/\pi_4 \end{bmatrix} \\ &= \left(\frac{1}{h}\right) \left(\frac{1}{\pi_1}\right) = \frac{1}{h_1} \end{aligned} \quad (5.11)$$

A similar covariance term can be derived for the θ terms. Thus

$\text{cov}[\log RR_A, \log RR_{A\&B}] = \frac{1}{h_1} + \frac{1}{k_1}$. Combining analogous results for all log-odds ratios we obtain:

$$\begin{aligned} &\text{Var} \begin{bmatrix} \log RR_A \\ \log RR_B \\ \log RR_{A\&B} \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{h_1} + \frac{1}{h_2} + \frac{1}{k_1} + \frac{1}{k_2} & \frac{1}{h_1} + \frac{1}{k_1} & \frac{1}{h_1} + \frac{1}{k_1} \\ \frac{1}{h_1} + \frac{1}{k_1} & \frac{1}{h_1} + \frac{1}{h_3} + \frac{1}{k_1} + \frac{1}{k_3} & \frac{1}{h_1} + \frac{1}{k_1} \\ \frac{1}{h_1} + \frac{1}{k_1} & \frac{1}{h_1} + \frac{1}{k_1} & \frac{1}{h_1} + \frac{1}{h_4} + \frac{1}{k_1} + \frac{1}{k_4} \end{bmatrix} \end{aligned}$$

We now apply these results to derive the variance of the multiplicative statistic, M . We have

$$\begin{aligned} M &= \log RR_{A\&B} - \log RR_A - \log RR_B \\ &= \begin{pmatrix} -1 & -1 & 1 \end{pmatrix} \begin{bmatrix} \log RR_A \\ \log RR_B \\ \log RR_{A\&B} \end{bmatrix}. \end{aligned}$$

Hence, $Var(M) = (-1, -1, 1)$.

$$\begin{bmatrix} \frac{1}{h_1} + \frac{1}{h_2} + \frac{1}{k_1} + \frac{1}{k_2} & \frac{1}{h_1} + \frac{1}{k_1} & \frac{1}{h_1} + \frac{1}{k_1} \\ \frac{1}{h_1} + \frac{1}{k_1} & \frac{1}{h_1} + \frac{1}{h_3} + \frac{1}{k_1} + \frac{1}{k_3} & \frac{1}{h_1} + \frac{1}{k_1} \\ \frac{1}{h_1} + \frac{1}{k_1} & \frac{1}{h_1} + \frac{1}{k_1} & \frac{1}{h_1} + \frac{1}{h_4} + \frac{1}{k_1} + \frac{1}{k_4} \end{bmatrix} \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} \quad (5.12)$$

$$= (-1, -1, 1) \begin{bmatrix} -\left(\frac{1}{h_1} + \frac{1}{h_2} + \frac{1}{k_1} + \frac{1}{k_2}\right) \\ -\left(\frac{1}{h_1} + \frac{1}{h_3} + \frac{1}{k_1} + \frac{1}{k_3}\right) \\ \left(\frac{1}{h_1} + \frac{1}{h_4} + \frac{1}{k_1} + \frac{1}{k_4}\right) - \left(\frac{1}{h_1} + \frac{1}{k_1}\right) - \left(\frac{1}{h_1} + \frac{1}{k_1}\right) \end{bmatrix} \quad (5.13)$$

$$= \left(\frac{1}{h_1} + \frac{1}{h_2} + \frac{1}{h_3} + \frac{1}{h_4}\right) + \left(\frac{1}{k_1} + \frac{1}{k_2} + \frac{1}{k_3} + \frac{1}{k_4}\right) \quad (5.14)$$

Hence, this formula for $Var(M)$ agrees with my previous derivation.

To test the multiplicative hypothesis, we note that $\log M$ should be zero if the multiplicative hypothesis is true. Therefore we can perform a Wald test using the ratio of M to its standard error $\sqrt{Var(M)}$. This will have an asymptotic normal distribution (Rao (1973)).

5.2.3. Tests for Case-Control Studies—Additive Case

We next consider the additive hypothesis, which may be stated:

$$A = RR_{A\&B} - (RR_A + RR_B - 1) = 0$$

i.e., that the relative risk of dying from contaminants $A\&B$ is equal to the sum of the relative risks from A and B separately less one. To derive a variance for the statistic A , we note that

$$\begin{aligned} Var(A) &= Var(RR_{A\&B}) + Var(RR_A) + Var(RR_B) \\ &\quad - 2cov(RR_{A\&B}, RR_A + RR_B) \\ &= Var(RR_{A\&B}) + Var(RR_A) + Var(RR_B) \\ &\quad + 2cov(RR_A, RR_B) - 2cov(RR_{A\&B}, RR_A) \\ &\quad - 2cov(RR_{A\&B}, RR_B) \end{aligned} \quad (5.15)$$

In the derivations presented thus far, we have found expressions for the variances and covariances of log relative risks. Clearly, the additive hypothesis requires variances and covariances of the relative risks themselves. One approach is to develop confidence intervals for the log relative risks, and translate them into confidence intervals for the relative risks by exponentiating the terms in the confidence interval inequality. In the presence of correlation, however, the best one can achieve with this technique are broad intervals based on the Bonferroni inequalities. A second approach uses the fact that if the log relative

risks are approximately normal, then the relative risks are approximately log normally distributed. Again, the joint distribution of log normal random variables is not straightforward. Consequently, this approach similarly becomes unworkable.

Instead, we follow Rothman (1976) and rely on a Taylor's series expansion. Specifically, we approximate the logarithm using: $\log y \doteq \log y_0 + (y - y_0)/y_0$ so that $Var(\log y) \doteq Var(y)/y_0^2$. Hence: $Var(y) = y_0^2 Var(\log y)$. The accuracy of the approximation improves for y close to y_0 , which we will achieve by taking y to be a consistent estimate of y_0 .

Collecting the terms required for $Var(A)$, we have:

$$Var(RR_{A\&B}) \doteq (RR_{A\&B})^2 \cdot \left[\frac{1}{h_1} + \frac{1}{h_4} + \frac{1}{k_1} + \frac{1}{k_4} \right]$$

$$Var(RR_A) \doteq (RR_B)^2 \cdot \left[\frac{1}{h_1} + \frac{1}{h_2} + \frac{1}{k_1} + \frac{1}{k_2} \right]$$

$$Var(RR_B) \doteq (RR_b)^2 \cdot \left[\frac{1}{h_1} + \frac{1}{h_3} + \frac{1}{k_1} + \frac{1}{k_3} \right]$$

For the covariance terms we employ similar expansions. Specifically let:

$$\log y \doteq \log y_0 + \frac{1}{y_0}(y - y_0) \text{ and}$$

$$\log z \doteq \log z_0 + \frac{1}{z_0}(z - z_0). \text{ Then}$$

$$\text{cov}(\log y, \log z) \doteq \frac{1}{y_0 z_0} \text{cov}(y - y_0, z - z_0) \text{ so that}$$

$$\text{cov}(y, z) \doteq (y_0 z_0) \cdot \text{cov}(\log y, \log z). \text{ Then:}$$

$$\text{cov}(RR_A, RR_B) = (RR_A \cdot RR_B) \cdot \left(\frac{1}{h_1} + \frac{1}{k_1} \right)$$

$$\text{cov}(RR_{A\&B}, RR_A) = (RR_{A\&B} \cdot RR_A) \cdot \left(\frac{1}{h_1} + \frac{1}{k_1} \right)$$

$$\text{cov}(RR_{A\&B}, RR_B) = (RR_{A\&B} \cdot RR_B) \cdot \left(\frac{1}{h_1} + \frac{1}{k_1} \right)$$

Then we have:

$$\begin{aligned} Var(A) \doteq & (RR_{A\&B})^2 \cdot \left[\frac{1}{h_1} + \frac{1}{h_4} + \frac{1}{k_1} + \frac{1}{k_4} \right] \\ & + (RR_A)^2 \cdot \left[\frac{1}{h_1} + \frac{1}{h_2} + \frac{1}{k_1} + \frac{1}{k_2} \right] \\ & + (RR_B)^2 \cdot \left[\frac{1}{h_1} + \frac{1}{h_3} + \frac{1}{k_1} + \frac{1}{k_3} \right] \end{aligned}$$

$$\begin{aligned}
& +2(RR_A \cdot RR_B) \cdot \left(\frac{1}{h_1} + \frac{1}{k_1}\right) \\
& -2(RR_{A\&B} \cdot RR_A) \cdot \left(\frac{1}{h_1} + \frac{1}{k_1}\right) \\
& -2(RR_{A\&B} \cdot RR_B) \cdot \left(\frac{1}{h_1} + \frac{1}{k_1}\right)
\end{aligned}$$

A Wald test may be conducted using the ratio of A to its standard error $\sqrt{Var(A)}$. Asymptotically, this will be standard normal, given the limiting distribution of the joint multinomial probabilities for π and θ . Since the Wald tests are valid only asymptotically we also consider a likelihood ratio approach.

5.2.4. Maximum Likelihood

The likelihood function for the case control study is $\prod_{i=1}^4 \pi_i^{h_i} \theta_i^{k_i}$ and is maximized subject to the constraint $\sum \pi_i - 1 = \sum \theta_i - 1 = 0$. The log likelihood function is

$$F = \sum h_i \log \pi_i + \sum k_i \log \theta_i .$$

This is maximized subject to the constraints:

$$F_1 = \sum \pi_i - 1 = 0 \quad (5.16)$$

$$F_2 = \sum \theta_i - 1 = 0 \quad (5.17)$$

$$\text{and } F_3 = \frac{\pi_4 \theta_1}{\pi_1 \theta_4} - \frac{\pi_2 \theta_1}{\pi_1 \theta_2} - \frac{\pi_3 \theta_1}{\pi_1 \theta_3} + 1 = 0 \quad \text{“additivity” or}$$

$$F_4 = \log (\pi_1 \theta_2 \theta_3 \pi_4) - \log (\theta_1 \pi_2 \pi_3 \theta_4) = 0 \quad \text{“multiplicativity”}$$

Note that F_3 may be rewritten:

$$\begin{aligned}
F_3 &= \frac{\pi_4}{\theta_4} - \frac{\pi_2}{\theta_2} - \frac{\pi_3}{\theta_3} + \frac{\pi_1}{\theta_1} \\
&= \frac{\pi_1}{\theta_1} - \frac{\pi_2}{\theta_2} - \frac{\pi_3}{\theta_3} + \frac{\pi_4}{\theta_4} = 0
\end{aligned} \quad (5.18)$$

Additive Constraint

For the additive model, we maximize the Lagrangian

$$F_A = F + \mu_1 F_1 + \mu_2 F_2 + \mu_3 F_3$$

where μ_1 , μ_2 , and μ_3 are Lagrange multipliers. The first order conditions are:

$$\frac{\partial F_A}{\partial \pi_i} = \frac{h_i}{\pi_i} + \mu_1 + \frac{\delta_i \mu_3}{\theta_i} = 0 \quad (5.19)$$

$$\frac{\partial F_A}{\partial \theta_i} = \frac{k_i}{\theta_i} + \mu_2 - \frac{\delta_i \pi_i \mu_3}{\theta_i^2} = 0 \quad \text{and} \quad (5.20)$$

$$\frac{\partial F_A}{\partial \mu_i} = F_i = 0 \quad i = 1, 2, 3 \quad (5.21)$$

where $\delta_1 = -\delta_2 = -\delta_3 = \delta_4 = 1$.

It follows that:

$$\sum \pi_i \frac{\partial F_A}{\partial \pi_i} = h + \mu_1 = 0 \quad \text{and} \quad \sum \theta_i \frac{\partial F_A}{\partial \theta_i} = k + \mu_2 = 0$$

Hence, $\hat{\mu}_1 = -h$ and $\hat{\mu}_2 = -k$ and the remaining conditions may be written:

$$(h_i - h \hat{\pi}_i) \hat{\theta}_i + \delta_i \hat{\pi}_i \hat{\mu}_3 = 0$$

$$(k_i - k \hat{\theta}_i) \hat{\theta}_i - \delta_i \hat{\pi}_i \hat{\mu}_3 = 0$$

$$\text{and} \quad \sum \delta_i \frac{\hat{\pi}_i}{\hat{\theta}_i} = 0$$

Writing $x_i = \hat{\pi}_i / \hat{\theta}_i$ and solving the first order conditions implies:

$$x_i = \frac{(\delta_i k_i h - k \hat{\mu}_3) \pm \sqrt{(\delta_i k_i h - k \hat{\mu}_3)^2 - 4 \delta_i h_i h k \hat{\mu}_3}}{2 \hat{\mu}_3 h} \quad (5.22)$$

Since $\sum_i \delta_i x_i = 0$, it follows that:

$$\begin{aligned} 0 &= \sum_i \left[(k_i h - k \hat{\mu}_3 \delta_i) \pm \delta_i \sqrt{(\delta_i k_i h - k \hat{\mu}_3)^2 - 4 \delta_i h_i h k \hat{\mu}_3} \right] \\ &= k h + \sum_i \pm \delta_i \sqrt{(\delta_i k_i h - k \hat{\mu}_3)^2 - 4 \delta_i h_i h k \hat{\mu}_3} \end{aligned} \quad (5.23)$$

This equation in $\hat{\mu}_3$ may be solved for each of 16 possible sign combinations (+ or - for each of the four terms in the sum).

Using $x_i = \pi_i / \theta_i$, the first two first order conditions may be written

$$(k_i - k \theta_i) = \delta_i x_i \mu_3 \quad \text{and} \quad (h_i - h \pi_i) = -\delta_i x_i \mu_3$$

Hence, $(k_i - k\theta_i) = -(h_i - h\pi_i)$, which implies

$$k_i + h_i = k\theta_i + h\pi_i = \theta_i(k + hx_i) \quad \text{or}$$

$$\theta_i = \frac{k_i + h_i}{k + hx_i}. \quad (5.24)$$

Now substitute into the first order condition:

$$k_i - k \left[\frac{k_i + h_i}{k + hx_i} \right] = \delta_i x_i \mu_3 \quad \text{or}$$

$$k_i(k + hx_i) - k(k_i + h_i) = (k + hx_i)\delta_i x_i \mu_3$$

$$k_i k + k_i x_i h - k(k_i + h_i) = k\delta_i x_i \mu_3 + x_i^2 h \delta_i \mu_3$$

$$-kh_i = x_i(-k_i h + k\delta_i \mu_3) + x_i^2 h \delta_i \mu_3$$

$$-kh_i = x_i(k\delta_i \mu_3 - k_i h) + x_i^2 h \delta_i \mu_3$$

$$-kh_i \delta_i = x_i(k\delta_i^2 \mu_3 - k_i h \delta_i) + x_i^2 h \delta_i^2 \mu_3$$

Now use $\delta^2 = 1$ as $\delta = 1$ or -1 . Then:

$$-kh_i \delta_i = x_i(k\mu_3 - k_i h \delta_i) + x_i^2 (h\mu_3) \quad \text{so that}$$

$$0 = x_i^2 (h\mu_3) + x_i(k\mu_3 - k_i h \delta_i) + kh_i \delta_i$$

The last equation establishes a bound on x_i since the discriminant of the quadratic equation must be positive. The discriminant is:

$$(k\mu_3 - k_i h \delta_i)^2 - 4(h\mu_3)(kh_i \delta_i) \geq 0$$

$$k_i^2 h^2 - 2\delta_i k_i h k \mu_3 + k^2 \mu_3^2 - 2 \cdot 2\delta_i h_i h k \mu_3 \geq 0$$

$$\mu_3^2 - 2\delta_i \left(\frac{h}{k} \right) \mu_3 (k_i + 2h_i) + k_i^2 \left(\frac{h}{k} \right)^2 \geq 0$$

Next, solving this quadratic at the point of equality to zero for μ_3 , we obtain:

$$\begin{aligned}
\mu_{3i}^* &= \frac{2\delta_i \left(\frac{h}{k}\right) (k_i + 2h_i) \pm \sqrt{4\delta_i^2 \frac{h^2}{k^2} (k_i + 2h_i)^2 - 4k_i^2 \left(\frac{h}{k}\right)^2}}{2} \\
&= \delta_i \left(\frac{h}{k}\right) (k_i + 2h_i) \pm \sqrt{\frac{h^2}{k^2} (k_i + 2h_i)^2 - k_i^2 \left(\frac{h}{k}\right)^2} \\
&= \left(\frac{h}{k}\right) \left[(k_i + 2h_i)\delta_i \pm \sqrt{k_i^2 + 4k_i h_i + 4h_i^2 - k_i^2} \right] \\
&= \left(\frac{h}{k}\right) \left[\delta_i (k_i + 2h_i) \pm 2\sqrt{h_i(h_i + k_i)} \right] \tag{5.25}
\end{aligned}$$

Since the quadratic has a positive second derivative, the inequalities are $\mu_3 \leq \min \mu_{3i}^*$ and $\mu_3 \geq \max \mu_{3i}^*$. Setting:

$$a_i = \left(\frac{h}{k}\right) \left[k_i + 2h_i - 2\sqrt{h_i^2 + h_i k_i} \right],$$

Gardner and Munford (1980) show that $-\min(a_2, a_3) \leq \hat{\mu}_3 \leq \min(a_1, a_4)$. Unfortunately, while these bounds bracket the true value of $\hat{\mu}_3$ they are not guaranteed to produce sign changes in the equation of interest. Therefore, an iterative solution is required to bracket each of the solutions for $\hat{\mu}_3$.³

Once $\hat{\mu}_3$ is found $\hat{\pi}_i$ and $\hat{\theta}_i$ are found from the first order conditions.

Multiplicative Constraint

For the multiplicative model, we maximize the Lagrangian:

$$F_M = F + \lambda_1 F_1 + \lambda_2 F_2 + \lambda_4 F_4$$

$$\begin{aligned}
\text{with } F &= \sum h_i \log \pi_i + \sum k_i \log \theta_i \quad \text{and} \\
F_1 &= \sum \pi_i - 1 \\
F_2 &= \sum \theta_i - 1 \\
F_4 &= \sum \delta_i \log \pi_i - \sum \delta_i \log \theta_i
\end{aligned}$$

we have:

$$\frac{\partial F_M}{\partial \pi_i} = \frac{h_i}{\pi_i} + \lambda_1 + \lambda_4 \frac{\delta_i}{\pi_i} = 0$$

³I note that $\hat{\mu}_3 = 0$ will always be a trivial solution to the equation above, and should be ignored.

$$\begin{aligned} \frac{\partial F_M}{\partial \theta_i} &= \frac{k_i}{\theta_i} + \lambda_2 + \lambda_4 \frac{-(\delta_i)}{\theta_i} = 0 \text{ . Then:} \\ \sum \pi_i \frac{\partial F_M}{\partial \pi_i} &= \sum h_i + \lambda_1 + \lambda_4 \sum \delta_i = 0 \Rightarrow \lambda_1 = -h \\ \sum \theta_i \frac{\partial F_M}{\partial \theta_i} &= \sum k_i + \lambda_2 + \lambda_4 \sum -(\delta_i) = 0 \Rightarrow \lambda_2 = -k \\ \frac{h_i}{\pi_i} - h + \frac{\lambda_4 \delta_i}{\pi_i} &= 0 \Rightarrow h_i - h\pi_i + \lambda_4 \delta_i = 0 \\ h\pi_i &= h_i + \lambda_4 \delta_i \\ \pi_i &= \frac{h_i + \delta_i \lambda_4}{h} \\ \frac{k_i}{\theta_i} + \lambda_2 + \lambda_4 \left(\frac{-\delta_i}{\theta_i} \right) &= 0 \Rightarrow \\ \frac{k_i}{\theta_i} - k + \lambda_4 \left(\frac{-\delta_i}{\theta_i} \right) &= 0 \Rightarrow \\ k_i - k\theta_i + \lambda_4(-\delta_i) &= 0 \\ -k\theta_i &= \lambda_4 \delta_i - k_i \\ \theta_i &= \frac{\lambda_4 \delta_i - k_i}{-k} = \frac{k_i - \lambda_4 \delta_i}{h} \end{aligned}$$

Finally, substituting into the constraint implies:

$$\left[\frac{h_1 + \delta_1 \lambda_4}{h} \right] \left[\frac{k_2 - \delta_2 \lambda_4}{k} \right] \left[\frac{k_3 + \delta_3 \lambda_4}{k} \right] \left[\frac{h_4 + \delta_4 \lambda_4}{h} \right] - \text{similar terms} = 0$$

which implies:

$$(h_1 + \lambda_4)(k_2 + \lambda_4)(k_3 + \lambda_4)(h_4 + \lambda_4) - \text{similar terms} = 0. \quad (5.26)$$

This equation may be solved for λ_4 from which follow solutions for θ_i and π_i .

Unconstrained Maximum Likelihood

The log likelihood under the constraint of additivity or multiplicativity is $\sum h_i \log \hat{\pi}_i + \sum k_i \log \hat{\theta}_i$. For the unconstrained case we maximize the Lagrangian

$$L = \sum h_i \log \pi_i + \sum k_i \log \hat{\theta}_i + \psi_1 \left[\sum \pi_i - 1 \right] + \psi_2 \left[\sum \theta_i - 1 \right]$$

The first order conditions are

$$\frac{\partial L}{\partial \pi_i} = \frac{h_i}{\pi_i} + \psi_1 = 0 \quad \text{and} \quad \frac{\partial L}{\partial \psi_1} = \sum \pi_i - 1 = 0$$

$$\frac{\partial L}{\partial \theta_i} = \frac{k_i}{\theta_i} + \psi_2 = 0 \quad \text{and} \quad \frac{\partial L}{\partial \psi_2} = \sum \theta_i - 1 = 0$$

These equations imply that $\hat{\pi}_i = h_i/h$ and $\hat{\theta}_i = k_i/k$ for the unconstrained maximum likelihood.

Hypothesis tests may be based on $-2(\log \text{likelihood unconstrained} - \log \text{likelihood constrained})$, which has a χ^2 distribution with one degree of freedom. While the additive and multiplicative models are non-nested, a comparison of the log likelihood values provides a basis for a non-nested hypothesis test.

5.3. Bonferroni Intervals and Simulations

5.3.1. Bonferroni Interval

The additive statistic $A = RR_{A\&B} - RR_A - RR_B - 1$ is composed of three random variables. A confidence interval for each component may be established using the variance of the log-odds ratio. Set at appropriate levels, these confidence intervals may be combined using basic results from probability theory. For a 95 percent confidence interval, chose a significance level such that one third of one half of 5 percent probability is in each tail of a normal distribution. Then:

$$\text{prob}[-2.39 \leq N(0, 1) \leq 2.39] = 1 - \frac{.05}{6} = 0.98334$$

Since $(\log \widehat{RR} - \log RR)/\sigma \sim^A N(0, 1)$ we have

$$\text{prob}[-2.39\sigma \leq \log \widehat{RR} - \log RR \leq 2.39\sigma] = 0.98334$$

or

$$\text{prob}[-2.39\sigma + \log \widehat{RR} \leq \log RR \leq 2.39\sigma + \log \widehat{RR}] = 0.98334$$

so that

$$\text{prob}[\widehat{RR}e^{-2.39\sigma} \leq RR \leq \widehat{RR}e^{2.39\sigma}] = 0.98334$$

Similarly,

$$\text{prob}[\widehat{RR}_{A\&B}e^{-2.39\sigma_{RR_{A\&B}}} \leq RR \leq \widehat{RR}_{A\&B}e^{2.39\sigma_{RR_{A\&B}}}] = 0.98334$$

and so forth for \widehat{RR}_A and \widehat{RR}_B . Similarly:

$$\text{Prob}[C_{low}^{A\&B} \leq RR_{A\&B} \leq C_{high}^{A\&B}] = .98334$$

$$\text{Prob}[C_{low}^A \leq RR_A \leq C_{high}^A] = .98334$$

$$\text{Prob}[C_{low}^B \leq RR_B \leq C_{high}^B] = .98334$$

Denoting the intervals within square brackets as A, B , and C , we have by the Bonferroni inequality:

$$\text{prob}[A \cap B \cap C] \geq 1 - (P(A^c) + P(B^c) + P(C^c))$$

Then

$$\begin{aligned} \text{prob}[C_{low}^{A\&B} \leq RR_{A\&B} \leq C_{high}^{A\&B} \cap \\ -C_{high}^A \leq -RR_A \leq -C_{low}^A \cap \\ -C_{high}^B \leq -RR_B \leq -C_{low}^B] \geq 1 - .05 = .95 \end{aligned} \quad (5.27)$$

so that

$$\begin{aligned} \text{prob}[C_{low}^{A\&B} - C_{high}^A - C_{high}^B \leq RR_{A\&B} - RR_A - RR_B \\ \leq C_{high}^{A\&B} - C_{low}^A - C_{low}^B] \geq .95 \end{aligned}$$

and

$$\text{prob}[C_{low}^{A\&B} - C_{high}^A - C_{high}^B - 1 \leq A \leq C_{high}^{A\&B} - C_{low}^A - C_{low}^B - 1] \geq .95$$

As noted before, given the tendency of the intervals to be broad and imprecise, these intervals should be rejected in favor of Wald or Likelihood Ratio tests.

5.3.2. Simulation Methods

Consistent estimates of the π_j and θ_j are formed using h_j/h and k_j/k respectively. A Monte Carlo technique draws a random multinomial deviate with marginal probabilities π_j and θ_j . Then, the empirical distribution of the statistics M and A are formed using repeated simulations. The empirical distributions establish confidence intervals centered around the realized value of the statistic. If these confidence intervals contain zero, then the hypothesis is not rejected.

5.4. Synergy Indices

This section considers three synergy indices: the Rothman synergy index, the attributable proportion index, and the gamma statistic which establishes a spectrum of synergy between the additive and multiplicative models.

5.4.1. Rothman's S Index

Rothman (1976) considers the independently-acting agents A and B and a background effect C . C is assumed to act independently of A and B .

Let P_T denote the probability that disease develops when both A and B are present in addition to the background C . P_A is the probability that disease develops if A were to act in isolation (without background). I define P_B similarly. P_C is the probability of getting disease from background only. Then

$$\begin{aligned} P_T &= P[A \cup B \cup C] \\ &= P[A] + P[B] + P[C] - P[A \cap B] - P[A \cap C] - \\ &\quad P[B \cap C] + P[A \cap B \cap C] \end{aligned} \quad (5.28)$$

Now, under independence we have:

$$\begin{aligned} P_T &= P[A] + P[B] + P[C] - P[A]P[B] - P[A]P[C] - P[B]P[C] \\ &\quad + P[A]P[B]P[C] \end{aligned}$$

Let $R_{AB} = P_T$ denote the combined risk.

$$\text{Let } R_A = P[A \cup C] = P[A] + P[C] - P[A]P[C]$$

$$\text{Let } R_B = P[B \cup C] = P[B] + P[C] - P[B]P[C]$$

$$\text{Let } R_0 = P[C]$$

Then, under independence:

$$\begin{aligned} R_{AB} - R_0 &= (R_A - R_0) + (R_B - R_0) - \frac{P_A P_B (1 - P_C)(1 - P_C)}{(1 - P_C)} \\ &= (R_A - R_0) + (R_B - R_0) - \frac{(R_A - R_0)(R_B - R_0)}{(1 - R_0)} \end{aligned} \quad (5.29)$$

Rothman's synergy index is defined as the ratio of the left-hand side of this equation to the right-hand side.

$$S = \frac{(R_{AB} - R_0)}{(R_A - R_0) + (R_B - R_0) + \frac{(R_A - R_0)(R_B - R_0)}{(1 - R_0)}}$$

Under independence, the numerator and denominator will be equal and the synergy index will equal one. Ignoring the product terms in the denominator, which are likely to be small, Rothman's index becomes:

$$S = \frac{(R_{AB} - R_0)}{(R_A - R_0) + (R_B - R_0)} = \frac{RR_{AB} - 1}{RR_A + RR_B - 2}$$

where $RR_{AB} = RR_{AB}/R_0$ etc. When $S = 1$, we obtain:

$$\begin{aligned} RR_{AB} - 1 &= RR_A + RR_B - 2 \quad \text{or} \\ RR_{AB} &= RR_A + RR_B - 1 \end{aligned} \quad (5.30)$$

which we recognize as the additive hypothesis.

An alternative expression for Rothman's S index is

$$S = \frac{ERR_{AB}}{ERR_A + ERR_B}$$

where $ERR_{AB} = RR_{AB} - 1$ and $ERR_A = RR_A - 1$ etc. Here, ERR denotes excess relative risk.

5.4.2. Attributable Proportion Index

The attributable proportion index is defined as the excess relative risk compared to the additive model divided by the combined relative risk. Formally,

$$\begin{aligned} AP &= \frac{ERR_{AB} - (ERR_A + ERR_B)}{(ERR_{AB} + 1)} \\ &= \frac{(R_{AB}/R_0 - 1) - [(R_A/R_0) + (R_B/R_0 - 1)]}{[R_{AB}/R_0 - 1 + 1]} \\ &= \frac{R_{AB} - R_0 - (R_A + R_B - 2R_0)}{R_{AB}} \\ &= \frac{R_{AB} - (R_A + R_B - R_0)}{R_{AB}} \\ &= \frac{RR_{AB} - (RR_A + RR_B - 1)}{RR_{AB}} \end{aligned} \quad (5.31)$$

When the additive model is correct, $AP = 0$.

Rothman's index S and the attributable proportion AP measure departure from additivity. They do not include the multiplicative hypothesis as a natural alternative. Therefore we consider an alternative which nests both hypotheses.

5.4.3. Additive-Multiplicative Measure

Define

$$\begin{aligned}\gamma &= \frac{(RR_{A\&B} - 1) - (RR_A - 1) - (RR_B - 1)}{(RR_A - 1)(RR_B - 1)} \\ &= \frac{RR_{A\&B} - RR_A - RR_B + 1}{(RR_A - 1)(RR_B - 1)}\end{aligned}\quad (5.32)$$

Note that when $\gamma = 0$ the additive hypothesis is true. When $\gamma = 1$ we have:

$$RR_{A\&B} = RR_A - RR_B + 1 = RR_A RR_B - RR_A - RR_B + 1$$

which implies: $RR_{A\&B} = RR_A \cdot RR_B$, i.e. the multiplicative hypothesis.

While difficult, a confidence interval may be derived by examining the distribution of $\log \gamma$. Note that

$$\log \gamma = \log A - \left(\log (RR_A - 1) + \log (RR_B - 1) \right)$$

where A is the additive statistic. Then

$$\begin{aligned}Var(\log \gamma) &= Var(\log A) - Var\left(\log (RR_A - 1)\right) \\ &\quad + Var\left(\log (RR_B - 1)\right) \\ &\quad + 2cov\left[\log (RR_A - 1), \log (RR_B - 1)\right] \\ &\quad - 2cov\left[\log A, \log (RR_A - 1)\right] \\ &\quad - 2cov\left[\log A, \log (RR_B - 1)\right].\end{aligned}\quad (5.33)$$

For case-control studies, we have previously derived these components. However, the utility of the expansion is questionable given that when the additive hypothesis is true, the log transformation is not defined.

5.5. Cohort Studies

Cohort studies derive standardized morbidity or mortality rates with reference to an external reference group. The standardized mortality rate (SMR) is also known as an observed to expected ratio because it is constructed by computing the expected number of outcomes (deaths) based on the external

reference group's rates. Given the large samples from which they are typically based, the latter rates are assumed to be known without error.

The cohort method compares the death rates between groups for those exposed to contaminant *A* (with or without exposure to *B*) and for those not exposed to contaminant *A* (with or without exposure to *B*). For present purposes, contaminant *A* will be smoking, while contaminant *B* will be asbestos. Death rates are calculated and given in the following 2×2 table:

	non-smoking	smoking
asbestos	d_A^{NS}	d_A^S
non-asbestos	d_{NA}^{NS}	d_{NA}^S

The cohort method follows a group of individuals with some exposure to asbestos. Death rates are determined over time for this cohort. A sample of individuals from a non-asbestos exposed population is matched to the exposed population at the aggregate level (i.e., there is a similar number of individuals of each age group).

Before discussing the derivation of the death rates d_j^i , we note that cohort studies make each cell of the 2×2 table independent by design. This greatly simplifies the hypothesis testing and determination of confidence intervals. Relative risks are determined as follows:

$$RR_A = \text{relative risk of asbestos exposure} = d_A^{NS} / d_{NA}^{NS}$$

$$RR_S = \text{relative risk of smoking exposure} = d_{NA}^S / d_{NA}^{NS}$$

$$RR_{AS} = \text{the relative risk of combined exposure} = d_A^S / d_{NA}^{NS}$$

The additive hypothesis is stated as:

$$RR_{AS} - RR_A - RR_S + 1 = 0$$

or

$$\frac{d_A^S}{d_{NA}^{NS}} = \frac{d_A^{NS}}{d_{NA}^{NS}} + \frac{d_{NA}^S}{d_{NA}^{NS}} - 1$$

or

$$d_A^S = d_A^{NS} + d_{NA}^S - d_{NA}^{NS}$$

or

$$A^* = d_A^{NS} + d_{NA}^S - d_A^S - d_{NA}^{NS}$$

Under additivity $A^* = 0$.

The multiplicative hypothesis is stated as:

$$RR_{AS} = RR_A \cdot RR_S = 0$$

or

$$\frac{d_A^S}{d_{NA}^{NS}} - \frac{d_A^{NS}}{d_{NA}^{NS}} \cdot \frac{d_{NA}^S}{d_{NA}^{NS}} = 0$$

or

$$d_A^S \cdot d_{NA}^{NS} - d_A^{NS} \cdot d_{NA}^S = 0 \quad (5.34)$$

or

$$\log d_A^S + \log d_{NA}^{NS} - \log d_A^{NS} - \log d_{NA}^S = 0$$

or

$$M^* = \log d_A^{NS} + \log d_{NA}^S - \log d_A^S - \log d_{NA}^{NS} = 0$$

Note that the multiplicative statistic is similar to the additive statistic with the exception that it is stated as a sum of logarithms. This suggests that the two hypotheses may be nested using a Box-Cox transformation.

It is worth noting that (32) implies

$$\frac{d_A^{NS}}{d_A^S} = \frac{d_{NA}^{NS}}{d_{NA}^S}$$

which states that the columns in the table are proportional to one another. Similarly, the rows are in proportion if the multiplicative hypothesis is correct. These are common statements of independence and can be tested via Pearson Chi-squared statistics for such tables. Finally, given the relationship between contingency tables and the log-linear model, we should expect a direct test of the multiplicative hypothesis from the log-linear model.

Suppose $\log (P[Y_1, Y_2]) = \mu_0 + \mu_1 Y_1 + \mu_2 Y_2 + \mu_{12} Y_1 \cdot Y_2$ Then

$$\begin{aligned} \log (P(0, 0)) &= \mu_0 \\ \log (P(0, 1)) &= \mu_0 + \mu_2 \\ \log (P(1, 0)) &= \mu_0 + \mu_1 \\ \log (P(1, 1)) &= \mu_0 + \mu_1 + \mu_2 + \mu_{12} \end{aligned}$$

If $P(0, 0)$ is estimated by d_{NA}^{NS} , $P(1, 0)$ by d_A^{NS} , $P(0, 1)$ by d_{NA}^S , and $P(1, 1)$

by d_A^S (after suitable normalization), then the multiplicative hypothesis may be stated as:

$$M^* = (\mu_0 + u_1) + (\mu_0 + \mu_2) - (\mu_0 + \mu_1 + \mu_2 + \mu_{12}) - (\mu_0) = -\mu_{12}$$

Then, $M^* = 0$ (the multiplicative hypothesis) if and only if the interaction parameter $\mu_{12} = 0$ in the log-linear model.

5.5.1. Determination of Death Rates

The death rate is defined as the number of deaths per 100,000 person years. This is typically measured by the number of deaths observed in the cohort divided by the number of person years multiplied by 100,000.

For example, suppose that a particular cohort has N_A^{NS} individuals who are non-smokers but who are exposed to asbestos. Suppose that these N_A^{NS} individuals are followed for Y_A^{NS} person years (on average Y_A^{NS}/N_A^{NS} years per person). Suppose that h_A^{NS} of these individuals die during the period of observation. Then

$$d_A^{NS} = \left(\frac{h_A^{NS}}{N_A^{NS}} \right) \left(\frac{N_A^{NS}}{Y_A^{NS}} \right) \cdot 100,000.$$

The stochastic component in the expression is the binomially distributed random variable h_A^{NS} that denotes the number of observed deaths in N_A^{NS} trials. Let P_A^{NS} denote the true but unobserved probability of dying. Then $\hat{P}_A^{NS} = h_A^{NS}/N_A^{NS}$ is a consistent estimate of P_A^{NS} .

$$\text{Now } E(\hat{P}_A^{NS}) = P_A^{NS} \text{ and } \text{Var}(\hat{P}_A^{NS}) = \frac{P_A^{NS}(1 - P_A^{NS})}{N_A^{NS}}.$$

Then

$$\text{Var}(d_A^{NS}) = \text{Var}(\hat{P}_A^{NS}) \cdot \left[\frac{N_A^{NS}}{Y_A^{NS}} \right]^2 \cdot 100,000^2$$

When logarithmic transformations are employed we have

$$\log d_A^{NS} = \log \hat{P}_A^{NS} + \log \left[\frac{N_A^{NS}}{Y_A^{NS}} \right] + \log (100,000).$$

Recall that a Taylor's series expansion shows that $\log \hat{P} \doteq \log P_0 + \frac{1}{P_0}(\hat{P} - P_0)$ so that

$$\text{Var}(\log \hat{P}) = \frac{1}{P_0^2} \frac{P_0(1 - P_0)}{N} = \frac{(1 - P_0)}{P_0 N}$$

Then

$$\text{Var}(\log (d_A^{NS})) \doteq \frac{(1 - \hat{P}_A^{NS})}{\hat{P}_A^{NS} N_A^{NS}}$$

Before proceeding with the formula for the variance of the A^* and M^* statistics, we note that replacing P_A^{NS} by \hat{P}_A^{NS} in the variance formula is valid asymptotically. Some researchers have noted that it may be more accurate in small samples to use a chi-square approximation.

To do this, we set $\chi^2 = \frac{(\hat{P}-P)}{P(1-P)/N}$. Then we set the χ^2 value to a critical level for the appropriate size test. Let χ_r^2 be the critical value. Then

$$\chi_r^2 = \frac{(\hat{P} - P)^2}{P(1 - P)/N}$$

so that

$$\begin{aligned} \hat{P} - 2\hat{P}P + P^2 &= \chi_r^2 P(1 - P)/N \\ &= \frac{P}{N}\chi_r^2 - \frac{P^2}{N}\chi_r^2 \end{aligned} \quad (5.35)$$

Then

$$P^2 \left(\frac{\chi_r^2}{N} + 1 \right) + P \left(\frac{-\chi_r^2}{N} - 2\hat{P} \right) + \hat{P}^2 = 0$$

is a quadratic equation that may be solved for P . A confidence bound is derived using the two solutions of the quadratic equation.

5.5.2. Variance of the Additive and Multiplicative Statistics

Next, we derive the variance of the additive and multiplicative statistics for cohort studies. Recall that

$$A^* = (d_A^{NS} - d_A^S) - (d_{NA}^{NS} - d_{NA}^S)$$

For the non-asbestos exposed cohort, the rates d_{NA}^{NS} and d_{NA}^S are determined from large samples and are considered non-stochastic. Therefore the variance is determined from the components d_A^{NS} and d_A^S , which are stochastic but independent. In this case,

$$\begin{aligned} Var(A^*) &= Var(d_A^{NS}) + Var(d_A^S) \\ &= \left[\frac{\hat{P}_A^{NS}(1 - \hat{P}_A^{NS})}{N_A^{NS}} \right] \left(\frac{N_A^{NS}}{Y_A^{NS}} \right)^2 \cdot (100,000)^2 + \\ &\quad \left[\frac{\hat{P}_A^S(1 - \hat{P}_A^S)}{N_A^S} \right] \left(\frac{N_A^S}{Y_A^S} \right)^2 \cdot (100,000)^2 \end{aligned} \quad (5.36)$$

For the multiplicative statistic,

$$M^* = [(\log d_A^{NS}) - (\log d_A^S)] - [(\log d_{NA}^{NS}) - (\log d_{NA}^S)]$$

so that

$$Var(M^*) = \frac{(1 - \hat{P}_A^{NS})}{\hat{P}_A^{NS} N_A^{NS}} + \frac{(1 - \hat{P}_A^S)}{\hat{P}_A^S N_A^S}$$

These variances are used to calculate standard errors, confidence intervals, and Wald tests for the additive and multiplicative hypotheses. For instance, $M^*/\sqrt{Var(M^*)}$ is asymptotically standard normally distributed under the null hypothesis that $M^* = 0$.

5.5.3. Variance of the Synergy Index

To derive the variance of S , $Var(S)$ we first find $\log S$.

$$\log S = \log(R_{AB} - R_0) - \log[(R_A - R_0) + (R_B - R_0)]$$

For cohort studies, we have:

$$\begin{aligned} \log S &= \log(RR_{AB} - 1) - \log[(RR_A - 1) + (RR_B - 1)] \\ &= \log\left(\frac{d_A^S}{d_{NA}^{NS}} - 1\right) - \log\left[\left(\frac{d_A^{NS}}{d_{NA}^{NS}} - 1\right) + \left(\frac{d_{NA}^S}{d_{NA}^{NS}} - 1\right)\right] \\ &= \log(d_A^S - d_{NA}^{NS}) - \log(d_A^{NS} + d_{NA}^S - 2d_{NA}^{NS}) \end{aligned}$$

$$Var(\log S) = \frac{Var(d_A^S)}{(d_A^S - d_{NA}^{NS})^2} + \frac{Var(d_A^{NS})}{(d_A^{NS} + d_{NA}^S - 2d_{NA}^{NS})^2} \quad (5.37)$$

where we have used the fact that $Var(d_{NA}^S) = Var(d_{NA}^{NS}) = 0$ in cohort studies since these variables are assumed to be non-stochastic.

For case-control studies, we have:

$$\begin{aligned} Var(\log S) &= \frac{Var(RR_{AB})}{(RR_{AB} - 1)^2} + \\ &\quad \frac{Var(RR_A) + Var(RR_B) + 2cov(RR_A, RR_B)}{(RR_A + RR_B - 2)^2} \end{aligned} \quad (5.38)$$

where the relevant components were derived above in the case-control section.

5.6. Conclusion

Case-control, cohort, and prevalence studies provide varying types of information to determine relative risks and attendant confidence levels. I have considered several methods for testing additivity and multiplicativity hypotheses using Wald and likelihood ratio techniques. In these cases, we have relied on asymptotic expectation for which the small sample populations are unknown. My empirical results are reported in the next chapter and, generally, I find agreement regarding the additivity or multiplicativity hypothesis whether the analysis is conducted using Wald or likelihood ratio methods.

Chapter 6

TESTS OF THE ADDITIVE AND MULTIPLICATIVE HYPOTHESES OF MULTIPLE EXPOSURE

6.1. Introduction

In this chapter, I consider the possible synergistic relationship between tobacco smoking and asbestos exposure. Asbestos refers to a group of naturally occurring flexible fibers that may be separated and woven. There are two major types of asbestos. The first type is amphibole fibers and appears commercially as amosite and crocidolite. The second type is chrysotile. Studies have demonstrated that amphibole fibers are much more dangerous to humans than chrysotile fibers. The fibers are resistant to heat and fire and do not conduct electricity. Asbestos fibers break easily if separated from their bonded finished products and can turn into dust particles that float in the air or stick to clothing.

Asbestos has been used since the late 1800s, but its use was greatly increased during World War II. Asbestos was used for strengthening cement and plastics as well as for fiber insulation, steam pipe insulation and in fireproofing. The automobile industry used asbestos in brake shoes and clutch pads. As is well known, literally thousands of products were manufactured which contained some form of asbestos. By the late 1970s, the United States barred the use of asbestos in wallboard products and in gas fireplaces. The asbestos industry has withdrawn asbestos from the majority of many products due to health concerns. Asbestos exposure, through inhalation, may lead to several diseases. These diseases include asbestosis, lung cancer, mesothelioma, and other cancers of the larynx and gastrointestinal tract.

Since the early 1970s, millions of workers were exposed to asbestos in shipyards, mining trades, manufacturing trades, insulation work, and automotive trades. Generally, the risk of disease increases with the amount and duration of exposure. Moreover, while brief periods of exposure are in some cases sufficient to produce disease, few signs of illness may develop for several decades.

The possible link between smoking and asbestos, was observed by Selikoff, *et al.* (1968). In this early study, Selikoff and his co-authors noted that combined exposure to asbestos and smoking produced more lung cancer deaths than might be expected under the simple addition of the risks taken separately. Thus, a potentially synergistic relationship was identified between the two carcinogens while the form of that synergy was left to be determined. Selikoff's original interpretation of his data suggested that a multiplicative model of risk existed or that asbestos would be harmful only in combination with smoking.

Selikoff, *et al.* (1968) used records from the International Association of Heat and Frost Insulators and Asbestos Workers to form his population. The population included 632 insulation workers as of 1942, that had been occupationally exposed to asbestos dust prior to 1922 or that had been exposed between 1922 and 1942. By 1962, 262 individuals had died and seven had died prior to their twentieth anniversary of first exposure to asbestos. For the remainder, Selikoff collected smoking habit history and clinical information. Selikoff's cohort followed 370 male insulation workers for a total of 52 months from 1963 through 1967. Of these, 283 were smokers or ex-smokers while 48 had never smoked regularly. An additional 39 were pipe or cigar smokers and eliminated from the study. The cohort method was used to compare the rate of death in the study population to a baseline rate of death taken from the American Cancer Society (ACS) for smokers and non-smokers. All together, 24 asbestos exposed smokers died, while no non-exposed smokers died during the fifty-two month follow-up period. For the comparison groups, ACS statistics revealed that 0.26 deaths were expected to occur among the 87 non-smokers and 2.78 deaths were expected to occur among the smokers (absent asbestos exposure).

Using the person years of follow-up for the exposed population, it is possible to calculate the death rate per 100,000 non-years as outlined in the previous chapter. This comparison of death rates is made to the ACS population so that relative risks can be calculated. While the Selikoff, *et al.* (1968) study was suggestive of synergism, the small sample size and lack of deaths among exposed non-smokers in the sample, lead to relative risk of lung cancer death from asbestos exposure equal to zero. Variances, as derived in the previous chapter, are thus, also zero and the tests of synergy become problematic. Selikoff noted this issue and subsequently studied a much larger population. Though Selikoff, *et al.* (1968) was later updated in Hammond, *et al.* (1979), it remains a classic article in this field and demonstrates the major elements of a cohort methodology.

In this chapter, I review 14 studies published between 1979 and 1992, which have applied either the cohort or case-control method to analyze synergy effects between smoking and asbestos exposure. In Section 6.2, I review these studies and summarize their findings in a standardized fashion to allow the

calculation of relative risks and variances. The 14 studies had sufficient data to analyze the synergy issue statistically.¹ It is important to remember that in almost all cases, the authors of the papers found evidence for either the additive or multiplicative hypothesis, but did so without basing their conclusions on formal statistical theory. Therefore, in some cases, it is not possible to accept the authors' conclusions as statistically significant or to reject the possibility that both the additive and multiplicative hypothesis are simultaneously true.²

In Section 6.3, I review two papers in more detail and demonstrate the cohort and case-control statistical tests in detail. I also implement the statistical tests derived in Chapter 5. I then summarize my findings for the additive and multiplicative hypothesis. In Section 6.4, I perform a meta-analysis using the data from the 14 studies to see whether the specific conclusions regarding synergy are more typical in some study formats or other characteristics than in other cases. Finally, I summarize my conclusions in Section 6.5.

6.2. Review of Studies

In this section, I review the 14 studies selected for analysis. My review of each article is necessarily brief and is intended to identify the population, the comparison group and the classification of exposure attributes.

6.2.1. Hammond, Selikoff, and Seidman (1979)

The Hammond, *et al.* (1979) paper considered 17,800 male insulation workers. This was the entire membership of the insulation workers union in the United States and Canada. These men were followed from the period of 1967 through 1976 in a cohort analysis. There were some 166,853 man-years of follow-up. A smoking questionnaire was administered in 1966 and 8,220 answered the questionnaire. 871 individuals were classified as non-smoking and 6,871 individuals were classified as smokers. 488 individuals were pipe smokers and excluded from the cohort study. In total, there were 51,397 man-years in the sample. The comparison study was the ACS survey of 73,761 men with matched smoking histories. All asbestos workers with at least twenty years of exposure were included in the study. The smoking attribute was provided as a "history of smoking" versus no smoking history; levels of smoking were excluded by the authors in their final tabulations.

In Table 8 of the original study, the lung cancer death rates per 100,000 man-years standardized by age, were computed for the study group and the

¹I have reviewed over 45 additional studies pertaining to lung cancer, but conclude that the relevant information is not present in these studies to test the synergy hypothesis.

²By far, the bulk of the effort in producing the data for the statistical analysis presented below, was in reviewing the articles and combining their findings in a systematic fashion.

ACS matched population. I summarize the results in Table 6.1. This table was derived from several sources within the Hammond, *et al.* (1979) paper including their Table 5 for the sample sizes, Table 7A for the number of deaths, and Table 8 for the death rates. I derived the number of person years for the exposed population as well as the matched number of deaths from ACS. For instance, I found that 0.78 deaths among 6849.0 person years of observation would imply a death rate of 11.30 as reported in the Hammond, *et al.* study.

Table 6.1. Hammond, *et al.* (1979)

	Non-Smokers (<i>N</i> = 891)		Smokers (<i>N</i> = 6841)	
	No	Yes	No	Yes
Asbestos Exposure				
Deaths (Lung Cancer)	0.78	4.00	54.60	268.00
Person Years		6849.00		44548.00
Death Rate / 100,000	11.30	58.40	122.60	601.60

6.2.2. Selikoff, Seidman, and Hammond (1980)

This study followed 933 male amosite factory workers between 1961 and 1977. These individuals had at least twenty years of asbestos exposure during their employment. Of the 933 men, 351 were excluded from the analysis because they were dead at the time of study or lost to follow-up. Among the non-smokers, there were 79 who had never smoked regularly and 59 who were pipe or cigar smokers. Finally, 15 were removed as their histories could not be determined. All together, the authors had 430 smokers and 132 non-smokers in their final cohort analysis. While death certificate deaths and best evidence deaths were both considered in Selikoff, *et al.* (1980), I follow their final tabulations and rely only on the “best evidence” information. The summary of death rates is provided in Table 6.2.

Table 6.2. Selikoff, *et al.* (1980)

	Non-Smokers (<i>N</i> = 137)		Smokers (<i>N</i> = 430)	
	No	Yes	No	Yes
Asbestos Exposure				
Deaths (Lung Cancer)	0.40	5.00	9.60	55.00
Person Years		1396.00		4811.00
Death Rate / 100,000	28.60	358.17	199.50	1143.21

For death rates of smokers who were not exposed to asbestos, Selikoff, *et al.* (1980) used the ACS sample. For non-smokers, they relied on a comparison to the population of New Jersey white males as the asbestos factory, at issue, was located in New Jersey. The authors concluded that a multiplicative model for lung cancer best explained the data. However, Steenland, and Thun (1986), in reviewing their paper, noted that as effect modification was occurring with non-smokers at higher rates than smokers, an additive model was more likely than a multiplicative model.

6.2.3. Berry, Newhouse, and Turok (1972)

This study followed asbestos factory workers between 1960 and 1970. The study period was 123 months. Stratification by exposure to asbestos (low or heavy) and by smoking (smoker, non-smoker, ex-smoker) and by gender (men and women) was conducted within the sample of 1300 men and 480 women. The comparison group was the London general population and expected deaths were based on smoking levels. The authors were not able to distinguish the additive and multiplicative effects from their data. Baker (1986) observed that this study proved that smoking was not indispensable in combination with asbestos to produce lung cancer. This is based on the sample of women non-smokers with severe asbestos exposure. Saracci (1977) finds that the relative risk from asbestos alone does not vary by smoking group and that a multiplicative model best explains the severe exposure category. My summary is based on males for comparison to the other studies. I combine the moderate and severe exposure groups into a heavy exposure group and combine male smokers and ex-smokers into a “smoking” category. However, as there were no deaths for non-smoking men without asbestos exposure I was unable to calculate the relevant risk for asbestos exposure alone. The summary information is provided in Table 6.3.

Table 6.3. Berry, *et al.* (1972)

	Non-Smokers (N = 85)		Smokers (N = 1279)	
	No	Yes	No	Yes
Asbestos Exposure				
Deaths (Lung Cancer)	0.00	0.00	18.60	42.00
Person Years		775.00		12083.00
Death Rate / 100,000	0.00	0.00	153.90	347.60

6.2.4. Berry, Newhouse, and Antonis (1985)

This article was a follow-up to the Berry, *et al.* (1972) study. The period of study was from 1971 through 1980. In this case, the authors conclude that their study favors the multiplicative hypothesis but they cannot reject the additive model. Steenland and Thun (1986), in their review, found that the differences in mortality rates for lung cancer among non-smokers, ex-smokers, and smokers were not significant due to the sample sizes. Vainio and Boffetta (1994) observed that the results apparently varied by type of asbestos, with workers exposed to amosite or a combination of amosite, chrysotile, and crocidolite showing varied degrees of interaction. My summary table is based on males where I combine the moderate and severe exposure groups into “heavy” exposure and combine ex-smokers together with smokers in our smoker category. Summary information is provided in Table 6.4.

Table 6.4. Berry, *et al.* (1972)

	Non-Smokers (<i>N</i> = 74)		Smokers (<i>N</i> = 1179)	
	No	Yes	No	Yes
Asbestos Exposure				
Deaths (Lung Cancer)	0.16	1.00	28.24	63.00
Person Years		669.00		10046.00
Death Rate / 100,000	23.90	149.50	281.10	627.10

6.2.5. McDonald, Liddell, *et al.* (1980)

This study followed 10,939 male chrysotile miners. While the authors conducted both case-control and cohort analyses using their data, we concentrate on their results using the case-control method. In the population of men, 245 lung cancer deaths were observed. Smoking was stratified into groups defined by (non-smokers, moderate smokers, ex-smokers, heavy smokers, and unknowns). This study was able to stratify asbestos exposure by the amount of dust and considered four ranges (none, low, moderate and heavy). Smoking history and available data allowed further study of 145 lung cancer cases. For the cases, 145 controls were selected from comparable persons with similar age and smoking histories. The results of this case-control study are summarized in Table 6.5.

The controls were selected at random among miners known to have survived to a greater age than that at which the case died and to have been born in the same year as the case. For ex-post classification of asbestos exposure, smoking history was matched between cases and controls. Smoking history was matched by number of cigarettes smoked per day (15-25, 26-35, 36-50).

Table 6.5. McDonald, Liddell, *et al.* (1980)

	No Exposure	Asbestos No-Smoking	No Asbestos Smoking	Asbestos Smoking
Cases (Lung Cancer)	5	15	41	84
Controls (No Lung Cancer)	10	10	49	76

The authors conclude that their data supports a synergism close to multiplicative but that neither model can be rejected confidently. Vainio and Boffetta (1994) interpreted their results as supporting the additive interaction theory. Baker (1986) found that the data did not discriminate between additive and multiplicative models.

6.2.6. Liddell, Thomas, *et al.* (1984)

This article was an extension of the McDonald, Liddell, *et al.* (1980) study. Briefly, 11,379 individuals born between 1891 and 1920 who had worked in asbestos mines and mills in Quebec were followed from 1967 through 1975. Among the 10,939 men, there were 4463 deaths. Among lung cancer deaths, referents were selected among men born in the same year and known to have survived to a greater age. This article used fiber counts rather than dust to measure accumulated exposure to asbestos. The study also used unmatched controls without regard to smoking status. I summarize the results in Table 6.6.

Table 6.6. Liddell, Thomas, *et al.* (1984)

	No Exposure	Asbestos No-Smoking	No Asbestos Smoking	Asbestos Smoking
Cases (Lung Cancer)	6	17	69	131
Controls (No Lung Cancer)	103	98	240	274

In Table 6.6, I have combined the moderate and heavy smokers and have combined the moderate and heavy asbestos exposures. The authors conclude

that the data is closer to multiplicative but that neither the additive or multiplicative model can be rejected with confidence. Steenland and Thun (1986) find that the outcome lies somewhere between additive and multiplicative while Vainio and Boffetta (1994) find that an additive was observed in this data from their review of the study.

6.2.7. Martischnig, Newell, *et al.* (1977)

This study, published in the British Medical Journal, examined 201 lung cancer patients admitted to a thoracic surgical center in the United Kingdom. Occupational exposure was determined and 58 individuals were identified to have had asbestos exposure in their jobs. The 201 men were matched to an equal number of controls based on age and residential area. Smoking was classified by the greatest number of cigarettes smoked in a day. Among controls, 29 men were found to have been exposed to asbestos on the job. The level of asbestos exposure was not stratified in this study. I summarize the results in Table 6.7. In the summary table, I have combined the moderate and heavy smokers into one category.

Table 6.7. Martischnig, Newell, *et al.* (1977)

	No Exposure	Asbestos No-Smoking	No Asbestos Smoking	Asbestos Smoking
Cases (Lung Cancer)	28	7	115	51
Controls (No Lung Cancer)	52	12	120	17

The authors interpret their results as showing that neither synergy hypothesis can be rejected with confidence. Saracci (1977) found that the data conform to a multiplicative model, as the relative risk was nearly constant across the different smoking categories.

6.2.8. Pastorino, Bearino, *et al.* (1984)

This study examined 204 lung cancers from the Lombardy region of Italy during the period 1976 through 1979. 70 individuals had some level of asbestos exposure of which 65 were smokers, and 61 were possibly exposed to asbestos of which 59 were smokers. A significant number of controls were matched to the cases. One control was selected per case for the deaths occurring in the first two years of the study and two controls were selected per case for the

deaths occurring in the last two years. The comparison group was males from the same study area. Smoking information was obtained from interviews with referents or from interviews with close family. As I have noted, exposure to asbestos was only “probable” in some cases. Cigarette smoking was defined using three categories (0-9 cigarettes per day, 10-19 cigarettes per day, and 20 or more cigarettes per day). I combined the moderate and heavy smoking groups into one category in Table 6.8.

Table 6.8. Pastorino, Bearino, *et al.* (1984)

	No Exposure	Asbestos No-Smoking	No Asbestos Smoking	Asbestos Smoking
Cases (Lung Cancer)	7	2	66	31
Controls (No Lung Cancer)	69	7	119	31

The authors do not provide a test of the additive or multiplicative hypothesis in this paper and do not reach any conclusions on the likelihood of synergy.

6.2.9. deKlerk, Musk, *et al.* (1991)

This study examined 2,928 Crocidolite factory workers and obtained detailed questionnaires from over 2,400 men and 149 from women. These individuals were followed between 1981 and 1986. In total, 40 lung cancer deaths were observed. Each case of lung cancer was matched to several controls of the same age. In total, there were 1,799 matched referents. The referents were taken from among surviving members of the same cohort. Information on occupational exposure from employment records allowed the authors to classify asbestos exposure as “high” versus “low.” Smokers were compared to non-smokers with the latter category including individuals for whom cessation was longer than 10 years. I summarize the results in Table 6.9.

The authors find a multiplicative effect for lung cancer. Vainio and Boffetta (1994) concur and note a pattern of results consistent with a multiplicative interaction. deKlerk, *et al.* (1991) used a logit model to explain the ex-post relative risks but failed to find a significant interaction term. As I discussed in Chapter 5, this is perfectly consistent with these findings of a multiplicative model and demonstrates only that super-multiplicativity is not present in this data.

Table 6.9. deKlerk, Musk, *et al.* (1991)

	No Exposure	Asbestos No-Smoking	No Asbestos Smoking	Asbestos Smoking
Cases (Lung Cancer)	2	4	9	25
Controls (No Lung Cancer)	399	357	522	521

6.2.10. Bovenzi, Stanta, *et al.* (1993)

This article considered 756 lung cancer deaths and a control population matched by age and location. The study area was Northeastern Italy over a five-year period. The analysis provided four different levels of smoking (no smoking, 1-19 cigarettes per day, 20-34 cigarettes per day, and greater than 40 cigarettes per day). Asbestos exposure was not stratified. In this study, cigar and pipe smoking were converted to cigarette smoking by assuming that one gram of tobacco was equivalent to one cigarette. The summary information is provided in Table 6.10.

Table 6.10. Bovenzi, Stanta, *et al.* (1993)

	No Exposure	Asbestos No-Smoking	No Asbestos Smoking	Asbestos Smoking
Cases (Lung Cancer)	10	4	245	143
Controls (No Lung Cancer)	103	19	249	75

The authors interpreted their results as a multiplicative model. They employed a logistic regression for the relative risks with three explanatory variables for smoking status, occupational exposure, and place of residence.

6.2.11. Blot, Harrington, *et al.* (1978)

This study examined 535 lung cancer cases with diagnosis made at hospitals in Georgia from 1970 through 1976. Only 458 individuals were included

in the final analysis due to interview refusals. The study assumed that all shipyard workers were exposed to asbestos but the amount of exposure was not determined. The smoking categories included non-smokers, light smokers, quits (greater than 10 years quit), moderate smokers, and heavy smokers (30 or more cigarettes per day). This study used two controls for individuals from Brunswick, Georgia and one control for cases from Savannah, Georgia. The controls were matched based on sex, race, age, and county of residence. I summarize the results in Table 6.11.

Table 6.11. Blot, Harrington, *et al.* (1978)

	No Exposure	Asbestos No-Smoking	No Asbestos Smoking	Asbestos Smoking
Cases (Lung Cancer)	50	11	313	84
Controls (No Lung Cancer)	203	35	270	45

The authors used a multivariate logit analysis to analyze the data although only the summary information was provided in the published paper. In preparing the summary table, I have combined the moderate and heavy smoking cases into a single category. Light smokers are further combined with non-smokers. The authors interpret the results as indicating strong synergism. However, Baker (1986), in his review, observed that the differences in odds ratios between smoking categories was not statistically significant.

6.2.12. Blot, Morris, *et al.* (1980)

This article examined 405 cases of lung cancer among shipbuilders in Virginia. 336 cases were included in the final analysis. All shipyard workers were assumed to have been exposed to asbestos if employment in the shipyards occurred prior to 1950. The study period was from 1972 through 1976 and deaths were ascertained from mortality listings in 1976. Smoking information was stratified identically to that in Blot, *et al.* (1978). There were 361 controls matched to the cases based on age, sex, race and county of residence. A summary of the information from this study is provided in Table 6.12.

In preparing the summary table, I combined the moderate and heavy smoking categories into a single group. I also combined light smoking with non-smoking and with those who had quit, as did the authors in their final results. The authors conclude that they cannot reject either hypothesis of synergism.

Table 6.12. Blot, Morris, *et al.* (1980)

	No Exposure	Asbestos No-Smoking	No Asbestos Smoking	Asbestos Smoking
Cases (Lung Cancer)	38	25	186	70
Controls (No Lung Cancer)	103	36	163	39

6.2.13. Kjuus, Skjaerven, *et al.* (1986)

This study examined 176 lung cancer cases taken from 1979 hospital data in Norway. The 176 male cases were matched to 176 controls. The comparison group came from the wards at the same two hospitals that provided the source population. Controls were matched by age and location. Asbestos history was obtained through personal interviews and questionnaires administered to the cases and controls. The asbestos exposure was grouped into four categories while smoking history was grouped into three categories (less than 10 cigarettes per day, 10-19 cigarettes per day, and 20 or more cigarettes per day). The summary data is provided in Table 6.13. In this table, I have combined moderate and heavy smokers into one category and non-smokers and light smokers into another category.

Table 6.13. Kjuus, Skjaerven, *et al.* (1986)

	No Exposure	Asbestos No-Smoking	No Asbestos Smoking	Asbestos Smoking
Cases (Lung Cancer)	29	8	103	36
Controls (No Lung Cancer)	96	11	63	6

The authors interpret their results as supporting the multiplicative model. Their statistical analysis was based on a conditional logit model.

6.2.14. Vena, Byers, *et al.* (1985)

The Vena, *et al.* article considered 1,002 lung cancers and established a set of 1,119 matched controls. Asbestos levels were stratified by three categories (none, 1-20 years, and 20 or more years). However, whether or not an individual had any asbestos exposure was determined by the individual's job title. Smoking was stratified into three categories (no-smoking, 1-39 pack years, and 40 or more pack-years).³ Lung cancers were identified by the type of cancer at the cellular level. A summary of the information from this study is provided in Table 6.14

Table 6.14. Vena, Byers, *et al.* (1985)

	No Exposure	Asbestos No-Smoking	No Asbestos Smoking	Asbestos Smoking
Cases (Lung Cancer)	97	41	164	86
Controls (No Lung Cancer)	240	72	86	25

The authors conducted a logit model to explain the relative risk of lung cancer. They were not able to reach a conclusion regarding additivity or multiplicativity in this data and, in fact, did not address this issue. Vainio and Boffetta (1994), in their review, concluded that multiplicative results held for squamous-cell carcinoma, and that additive results held for small-cell carcinoma, and for adenocarcinoma.

6.3. Tests of Hypothesis

In this section, I examine two of the studies in more detail. The first study is Hammond, Selikoff and Seidman (1979) and relies on the cohort method. The second study is deKlerk, Musk, Armstrong and Hobbs (1991) and relies on case control methods.

6.3.1. Hammond, Selikoff, and Seidman (1979)

From Table 6.1 in Section 6.2, we see that the number of non-smokers who were exposed to asbestos, N_A^{NS} , is 891. Of these, $h_A^{NS} = 4$ die of lung cancer during the study period. Further, the number of person-years for this group

³A pack-year is the smoking of one pack of cigarettes per day for one year.

is $Y_A^{NS} = 6849$ implying that the 891 individuals were followed for approximately 7.7 years per person. Recall that the study period was nine years so that the lower average reflects deaths in this cohort. The death rate is then:

$$d_A^{NS} = \left(\frac{h_A^{NS}}{N_A^{NS}} \right) \left(\frac{N_A^{NS}}{Y_A^{NS}} \right) \times 100,000 = \left(\frac{4}{891} \right) \left(\frac{891}{6849} \right) \times 100,000 = 58.4.$$

The probability of dying is denoted by P_A^{NS} in Chapter 5 and is estimated by $\hat{P}_A^{NS} = h_A^{NS}/N_A^{NS}$ which is $4/891$ or 0.00449 . The variance of this estimate is $V(\hat{P}_A^{NS}) = (0.00449) \times (1 - 0.00449)/891 = 5.02e - 6$. Thus $V(d_A^{NS}) = [(891/6849) \times 100,000]^2 \times (5.02e - 6) = 849.01$.

Similar calculations show that $d_A^S = 601.6320$ and $V(d_A^S) = 1297.54$. From equation (5.36), we see that the variance of the additive statistic is $849.01 + 1297.54 = 2146.5$. The additive statistic is:

$$A^* = (d_A^{NS} - d_A^S) - (d_{NA}^{NS} - d_{NA}^S) = (58.4 - 601.6) - (11.3 - 122.6) = -431.9.$$

Thus, the t -statistic for the additive hypothesis using a Wald test is $-431.9 / \sqrt{2146.6} = 9.32$. The t -value rejects the additive hypothesis at the 95 percent confidence level.⁴ Similar calculations were carried out for the multiplicative statistic M^* and its variance.

Following the analysis in the previous chapter, the relative risk of asbestos exposure is $RR_A = d_A^{NS}/d_{NA}^{NS} = 58.4/11.30 = 5.16$, the relative risk from smoking exposure is $RR_S = d_{NA}^S/d_{NA}^{NS} = 122.6/11.30 = 10.84$, and the combined relative risk is $RR_{AS} = d_A^S/d_{NA}^{NS} = 601.6/11.3 = 53.2$. The multiplicative hypothesis asks whether $5.16 \times 10.85 = 55.99$ is equal to the combined risk of 53.2 or whether $5.16 + 10.85 - 1 = 15.01$ is equal to the combined risk of 53.2. The t -statistic for the multiplicative hypothesis is 0.10, which implies that we cannot reject the multiplicative hypothesis.

The synergy index, in this case, is 3.7 and would be one if the additive hypothesis were true. The variance of the logarithm of the synergy index was calculated to be 0.0376 using equation (5.38). A 95 percent confidence interval for the synergy index is [2.53 - 5.49] and excludes the value one allowing us to reject additivity. Similar calculations produce the gamma statistic derived in the previous chapter. The gamma value is 0.93 in this case. Recall that gamma equals zero in the additive model, while gamma equals one in the multiplicative model. The attributable proportion statistic is 0.72 and would be zero if the model were additive. Thus, the additive model is rejected in favor

⁴Note that we have explicitly treated the death rates from the comparison group as non-stochastic as they are typically derived from a large population.

of the multiplicative model based on the statistical tests as well as the synergy indices.

Apparently, there is little controversy among authors who have reviewed this study. For example, the authors, themselves, interpret their results as a multiplicative model. Steenland and Thun (1986) find “no interaction as defined as a departure from the multiplicative model.” Vainio and Boffetta (1994) find that this study “showed a more uniformly multiplicative pattern.” However, Steenland and Thun were concerned that the rate of four deaths among non-smoking asbestos workers implied a large variance of the rate ratio and that “no firm conclusions” could be reached. To be precise, Steenland and Thun were concerned with the estimate of RR_A and its variance. However, as we have seen, the relevant variance is that for the death rate, d_A^{NS} , which has an estimated coefficient of variation of roughly 0.5. Therefore, this estimate is not as imprecise as Steenland and Thun had suspected. Of course, one remaining defect of this study is that the levels of smoking and asbestos exposure are not provided in the final analysis so that it is not possible to distinguish heavy smokers and light smokers or to gauge the importance of asbestos exposure on these results. This omission in the published article is unfortunately typical. Our next example is from a case control study.

6.3.2. deKlerk, *et al* (1991)

From Table 6.9 in Section 6.2, we see that the number of cases with no exposure (in this case “low” exposure) to asbestos and no exposure to smoking is 2. In Chapter 5.2, this outcome was labeled h_1 . Similarly, the number of non-smokers with exposures to asbestos (in this case, “high” exposure), h_2 , is 4. The number of cases without exposure to asbestos but with smoking exposure, h_3 , is 9 while the number of cases with smoking exposure and asbestos exposure, h_4 , is 25.

Next, from Table 6.9 in Section 6.2, the corresponding numbers of controls are $k_1 = 399$, $k_2 = 357$, $k_3 = 522$, and $k_4 = 521$. The estimated unconditional probabilities for the cases are $\pi_1 = 2/40 = 0.05$, $\pi_2 = 4/40 = 0.10$, $\pi_3 = 9/40 = 0.225$, and $\pi_4 = 0.625$. For the controls, the probabilities are $\theta_1 = 0.222$, $\theta_2 = 0.198$, $\theta_3 = 0.290$, and $\theta_4 = 0.290$.

As this is a case-control study, we estimate $RR_A = \pi_2\theta_1/\pi_1\theta_2 = 2.235$, $RR_B = \pi_3\theta_1/\pi_1\theta_3 = 3.440$, and $RR_{AB} = \pi_4\theta_1/\pi_1\theta_4 = 9.573$. Equation (5.15) provides the formula for the variance of the additive statistic. In this case, we estimate $\text{var}(RR_{AB}) = 49.89$, $\text{var}(RR_A) = 3.77$, $\text{var}(RR_B) = 7.28$, $\text{cov}(RR_A, RR_B) = 3.86$, $\text{cov}(RR_{AB}, RR_A + RR_B) = 27.30$, and therefore the combined variance is $\text{var}(T) = 14.08$. The value of the additive statistic is $T = 9.573 - (2.235 + 3.440 - 1) = 4.898$. Therefore, the additive t -statistic is $4.898/3.75 = 1.305$. Similarly, the multiplicative statistic is $M = \log RR_{AB} - \log RR_A - \log RR_B = 2.259 - 0.804 - 1.235 = 0.219$. Fi-

nally, equation (5.14) implies that $\text{var}(M) = 0.910$ so that the Wald t -statistic is 0.23. Thus the Wald tests permit one to accept both the additive and multiplicative hypothesis in this case.

Regarding the maximum likelihood results under the additive constraint, I find that $\hat{\mu}_3 = -2.9902$ is a solution for equation (5.25) when the sum is taken over terms with sign pattern $(-, +, +, -)$. Other than the trivial solution of $\hat{\mu}_3 = 0$ for the same pattern, there are no other solutions found for equation (5.25). The non-trivial solution for $\hat{\mu}$ implies $x_1 = 0.168$, $x_2 = 0.817$, $x_3 = 1.053$, and $x_4 = 1.702$ from equation (5.23) under the same sign pattern $(-, +, +, -)$. Finally, (5.24) provides solutions for θ_i and π_i with $\theta_1 = 0.222$, $\theta_2 = 0.197$, $\theta_3 = -.288$, $\theta_4 = 0.292$, $\pi_1 = 0.0374$, $\pi_2 = 0.1611$, $\pi_3 = 0.3037$, and $\pi_4 = 0.4976$. At these values, the additivity constraint is exactly satisfied. The likelihood ratio test has the value 3.4, which does not exceed the 95 percent critical value with one degree of freedom (the p -value is 0.065). Hence, additivity may be accepted in this study.

For the multiplicative constraint under maximum likelihood, I find using equation (5.26) that $\lambda_4 = -0.235$. At this value, I find $\theta_1 = 0.222$, $\theta_2 = 0.198$, $\theta_3 = 0.290$, $\theta_4 = 0.289$, $\pi_1 = 0.0044$, $\pi_2 = 0.105$, $\pi_3 = 0.231$, and $\pi_4 = 0.619$. At these values the implied relative risks satisfy the multiplicative constraint exactly. The likelihood ratio statistic is 0.05, which is smaller than the 95 percent critical value of the chi-square with one degree of freedom (3.84). Hence, the multiplicative hypothesis may also be accepted with this data.

Finally, the unconstrained estimates of relative risk produce values of the synergy indices of 1.625 for gamma, 2.330 for Rothman's synergy index and 0.512 for the attributable proportion statistic. I use equation (5.38) to derive the variance of the logarithm of Rothman's synergy index at 2.07. The 95 percent confidence bound is calculated to be $[0.13 - 41.44]$. As the interval includes $S = 1$, we may again accept the additive hypothesis.

The results of similar analysis for all 14 studies are displayed in Figures 6.1–6.5. Figure 6.1 summarizes the relative risks for smoking, asbestos exposure and the combined relative risks of smoking with asbestos exposure. Table 6.2 provides the t -statistics for the additive and multiplicative hypothesis. Five of the fourteen studies reject the additive hypothesis. None of the studies reject the multiplicative hypothesis. The Rothman synergy statistics are presented in Figure 6.3 along with their 95 percent confidence intervals. As the majority of these intervals include $S = 1$, we may accept the additive hypothesis in all but two studies. Figure 6.4 provides the estimate attributable proportion statistics. Finally, Figure 6.5 provides the gamma statistics. In these figures, the data for the Berry (1972) study has been omitted where it requires calculation of the asbestos relative risk for non-smokers as this is not defined when there are no lung cancer deaths.

Relative Risks

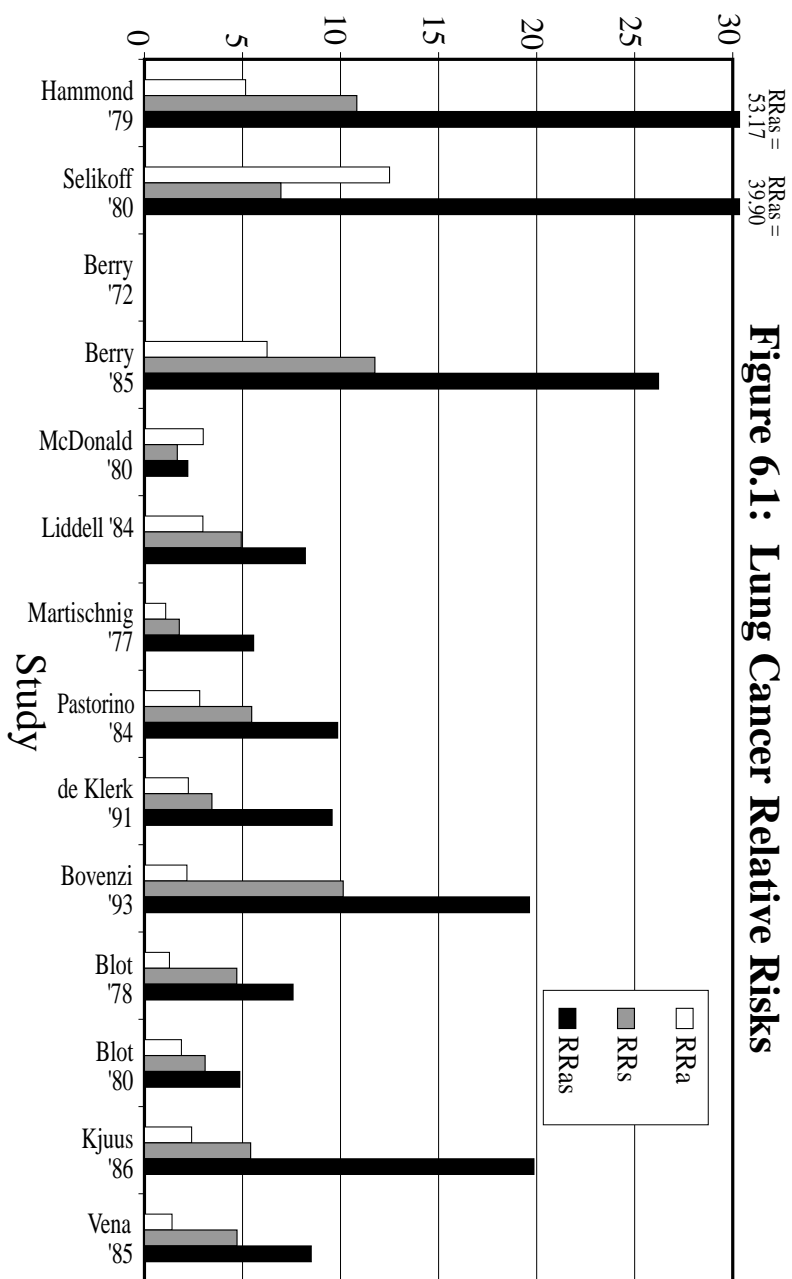


Figure 6.1: Lung Cancer Relative Risks

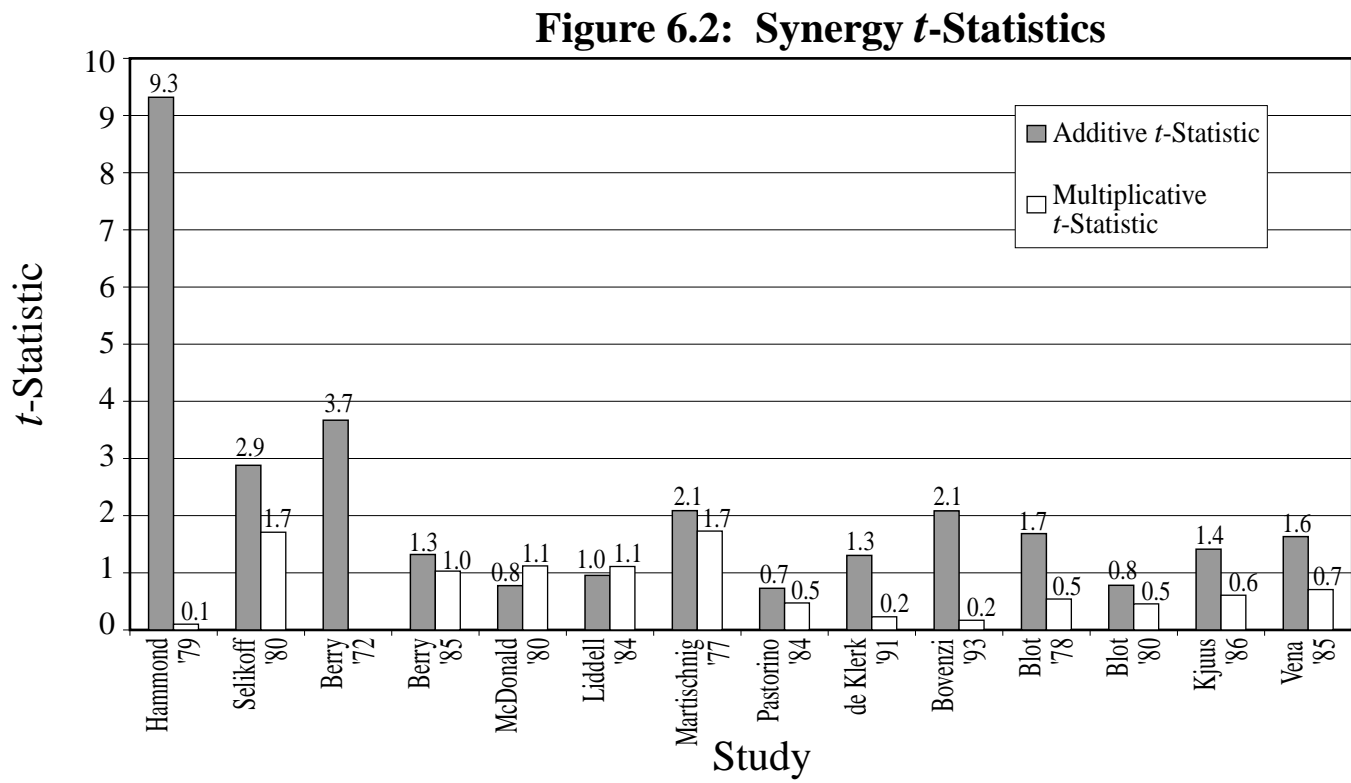


Figure 6.3: Synergy Index Values and Confidence Intervals

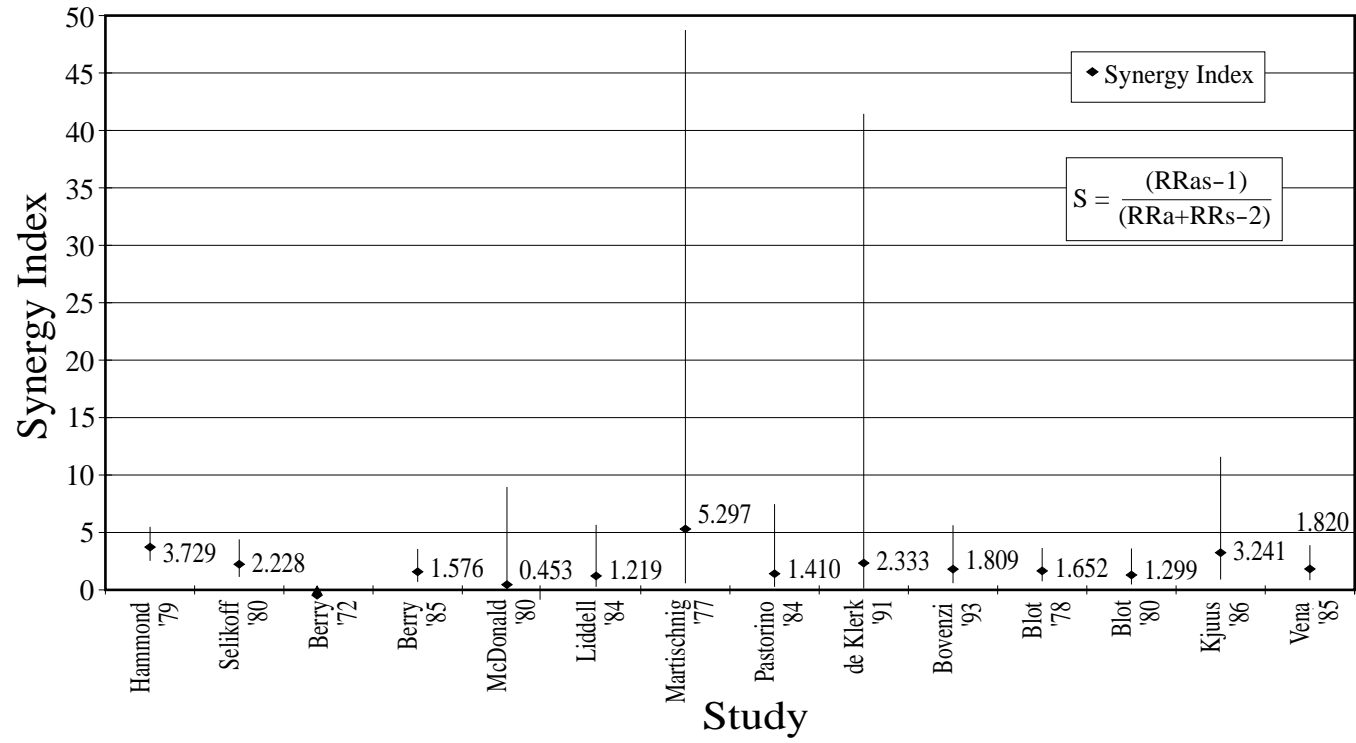
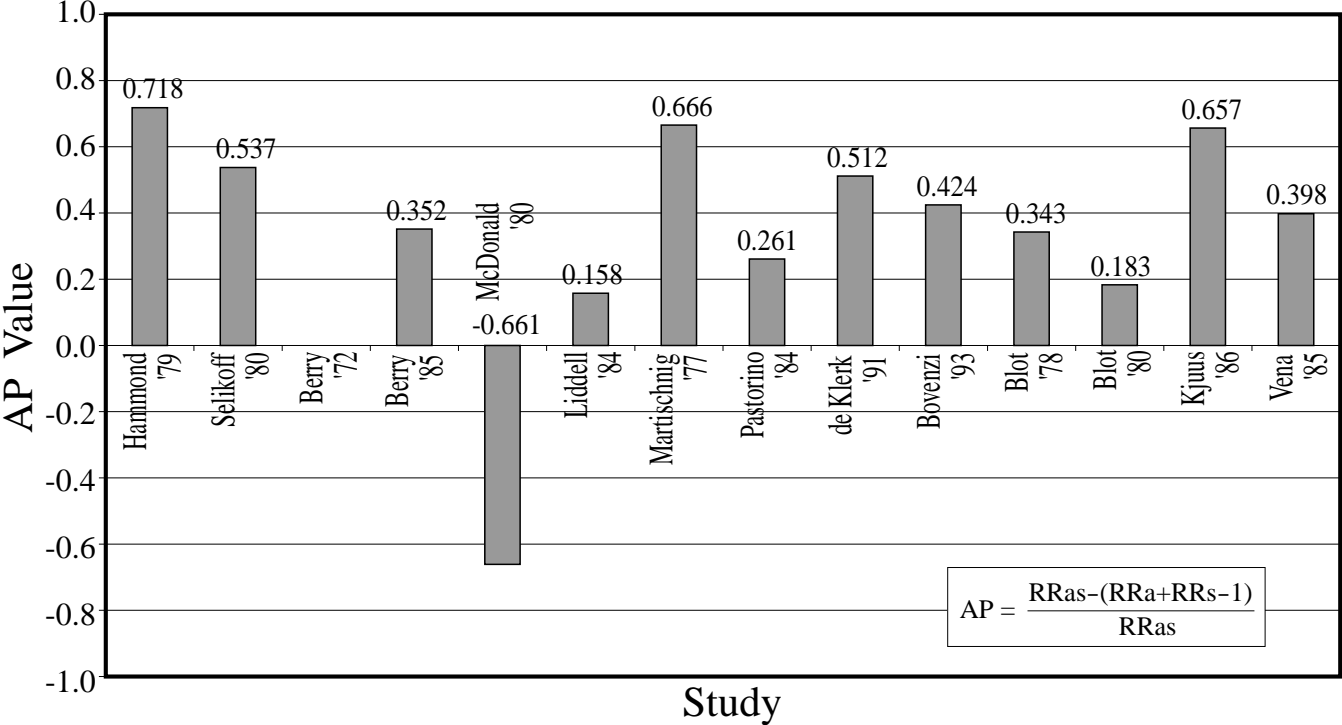


Figure 6.4: Attributable Proportion Statistic



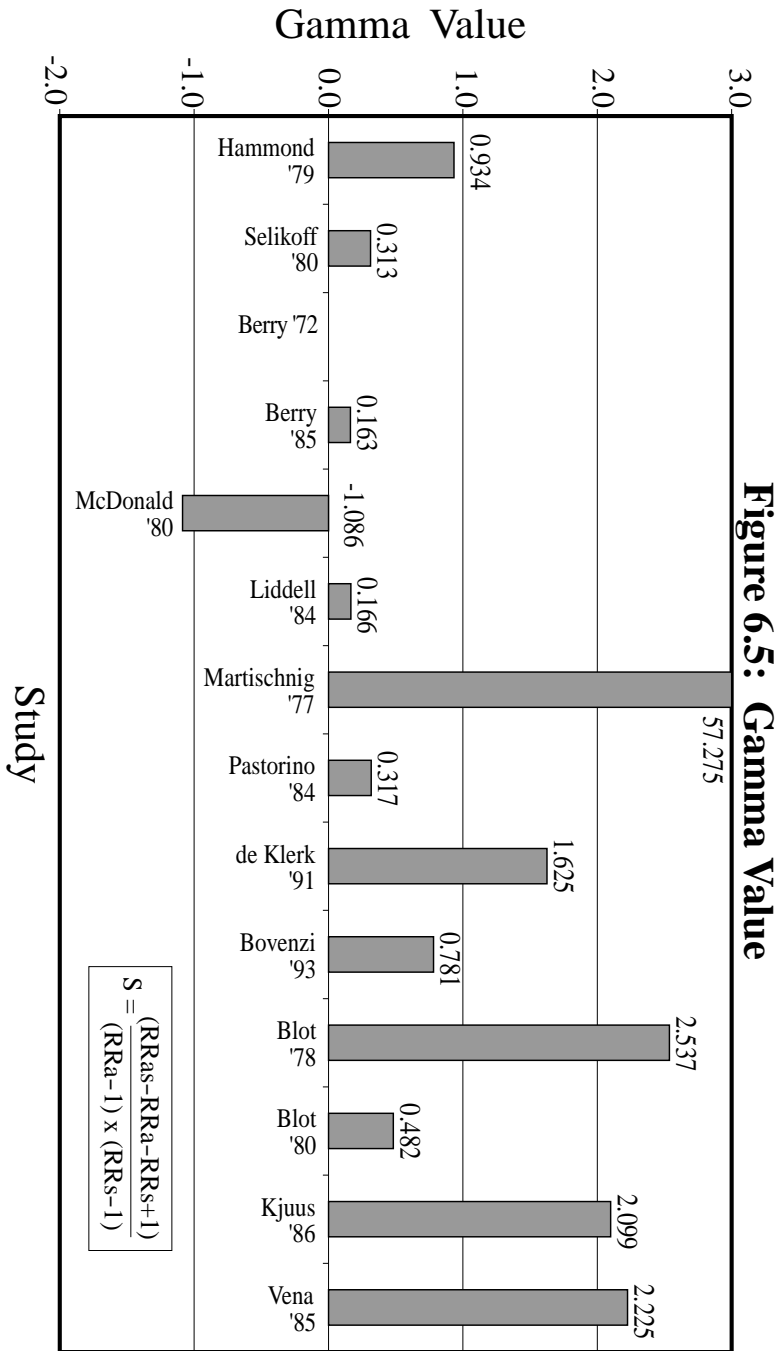


Figure 6.5: Gamma Value

A pooled hypothesis test for additivity may be based on the joint distribution of the t -test values. Each t -value is normally distributed and the studies are independent. Therefore, the sum of their squared values is chi-squared under the null hypothesis that each statistic has true value of zero under the additive or multiplicative statistic respectively. The sum of the squared t -values are 131.7 and 11.14 for the fourteen studies. The critical level for the chi-squared at 95 percent confidence and fourteen degrees of freedom is 23.7. Therefore, the pooled estimates reject additivity and accept multiplicativity. The deletion of just a few studies alters this result.

Likelihood ratio statistics for the additive hypothesis are calculated for the case-control studies. The statistics starting with the Liddell (1984) paper and presented in the order of the articles in Figure 6.1, are 1.247, 0.818, 6.56, 0.481, 3.413, 6.987, 3.62, 0.64, 5.466 and 9.51. For the multiplicative hypothesis, the likelihood ratio statistics are 1.289, 1.30, 3.07, 0.214, 0.0522, 0.028, 0.296, 0.207, 0.376, and 0.504. The likelihood ratio statistics for the additive hypothesis, reach the same conclusion as the Wald statistics with the exception of three studies where additivity was accepted under the Wald test but rejected under the likelihood ratio test. The likelihood ratio tests for the multiplicative hypothesis reach identical conclusions as the Wald tests. Additionally, the pooled likelihood ratio test for the joint hypothesis of additivity is rejected but is accepted under the multiplicative hypothesis as was the case using the Wald tests.

In the next section, I use a meta-analysis approach to further examine the pattern of results illustrated in these studies.

6.4. Meta-analysis of the Fourteen Synergy Studies

The fourteen studies discussed in the previous sections contain sufficient information to derive case-control or cohort results for a variety of sub-populations. For instance, the Selikoff, Seidman and Hammond (1980) study provides cohort matrices based on the best evidence for cause of death as compared with information taken from death certificates. Similarly, the Berry, *et al.* (1972) study allows asbestos levels to be distinguished by low exposure versus high exposure and further differentiates between men and women. This provides six cohort matrices and six possible results regarding synergy.

In another example, Berry, *et al.* (1985) provides eight separate cohort matrices as it stratifies by two levels of asbestos exposure (low and heavy) and by gender (male and female) and either compares non-smokers to smokers or to ex-smokers. In my summaries in Sections 6.2 and 6.3, above, I combined the results for these separate cohort matrices, in each study, into a single cohort matrix. In this section, I conduct the synergy analysis for 41 separate cohort and case-control matrices taken from the 14 studies. I do this to gain degrees of freedom in the meta-analysis and so that I may contrast several facets of

the studies. I am able to consider a variety of study attributes including, for example, whether or not the study stratifies by asbestos exposure and the level of that exposure (low or none at the lower boundary of no exposure versus general, low, moderate, high, or combined at the high boundary of exposure). These categories are necessarily crude given the information provided in the published studies.

I also consider the number of cases studied for case-control or cohort studies and the number of controls for case-control studies. I next determined the final year of the study period, the gender of the individuals in the study matrix, the type of study (whether case-control or cohort), whether Selikoff is an author of the study, whether the study is stratified by smoking levels and the boundaries for the exposed and non-exposed levels of smoking experience (non-smokers, light or moderate at the low-end and smokers, ex-smokers, light, moderate, heavy or combined for the high-end of exposure). I also calculate the relative risks and gamma synergy indices, as well as the additive and multiplicative t -values for each cohort or case-control matrix in each study.

Of the 41 cases, 17 come from the cohort studies while 24 come from the case-control studies. The gamma values calculated range from -6.7 to 60.4 with an average value of 3.86 . As the average is greater than one, it is possible that some subsets of the data indicate a synergistic effect greater than that in the multiplicative model.

Using an ANOVA analysis, I find that the t -values for the additive hypothesis are larger in cohort studies ($p = 0.027$). The t -values are also higher in cohort studies for the multiplicative hypothesis ($p = 0.0004$). The average of the t -values for the additive hypothesis is 1.28 and the average of the t -values for the multiplicative hypothesis is 0.37 . But the magnitude of the differences between cohort and case-control studies does not affect the result of the individual t -test at the 95 percent confidence level. Thus, changing a study from case-control to cohort would not be sufficient to change a result from acceptance to rejection of a synergy hypothesis.

The results of a multivariate regression analysis for the t -values are less clear. I find that controlling for case-control versus cohort, whether smoking exposure is stratified, whether asbestos exposure is stratified, and gender does not show that the cohort method results in higher t -values. However, the same regression analysis demonstrates that cohort methods lead to higher multiplicative t -values and that studies in which asbestos is stratified lead to lower t -values for the multiplicative hypothesis.

With respect to whether the results of a particular study are more likely or not to imply an additive or multiplicative synergy, I use the gamma statistic. Recall that gamma values of zero indicate additive models while gamma values of one indicate multiplicative models. Additionally, the gamma values may be less than zero or greater than one indicating sub-additivity or super-

multiplicativity. Based on ANOVA results, I find that the gamma values reveal statistically different results when the study stratifies by asbestos exposure level. The gamma values were next examined in a multivariate regression analysis. Here, I find that over 96 percent of the variation in the gamma values may be explained by very few factors. First, the cohort method lowers the gamma value. I also find that larger gamma values occur when men are studied (relative to women) and that studies that stratify by smoking level produce larger gamma values, but that studies which stratify by asbestos level produce lower gamma values. On balance, multiplicative models are more likely with men and in studies that stratify by smoking level. Further, conclusions regarding the level of smoking or asbestos exposure would be desirable but it is generally not possible to interpret the author's meaning of such terms as low, high, moderate and so forth.⁵

6.5. Conclusions

This chapter has reviewed 14 studies of tobacco asbestos synergy. The basic strategy has been to find a common set of results by calculating the implied relative risks from each study and calculating the synergy indices implied by the studies. In each case, I have estimated the precision of the underlying statistics. The average gamma values and the average Rothman synergy levels are all in the range of multiplicative outcomes while the average *t*-values are large enough to reject additivity and accept multiplicativity. The results indicate that the average attributable proportion is 0.35 and therefore implies that among smokers who are exposed to asbestos, approximately 35 percent of lung cancer cases can be attributed to a synergistic effect of the two agents as distinct from their individual effects. However, these averages are affected significantly by the detection of just a few studies. The meta-analysis demonstrates a lack of uniformity among the studies and shows that cohort studies produce higher additive and multiplicative *t*-values. Only five in fourteen studies have *t*-values large enough to reject additivity while the multiplicative hypothesis is not rejected by any of the studies using the Wald test. Relying on the author's separate stratified analyses, produced 41 synergy cases for study. In only six of these, was the additive hypothesis rejected while the multiplicative hypothesis was accepted all but one case. The average attributable proportion statistic for the 41 cases is 25 percent and leads to a lower estimate of the synergy effect as compared with the summarized results from the fourteen studies. The available evidence is supportive of the multiplicative model of synergy but, in the majority of cases, the additive result is equally tenable.

⁵To pursue this, I considered cases in which the "low" exposure level was at its lowest value and the "high" exposure level at its highest categorical value. However, these cases failed to distinguish the level of the gamma statistic.

Chapter 7

CONCENTRATION AND COMPETITION IN THE CHEMOTHERAPY DRUG MARKET

7.1. Introduction

The pharmaceutical industry is comprised of companies that discover and sell drugs that require a doctor's prescription. The pharmaceutical industry ranks among the top industries in the United States in sales and research and development. This chapter considers competition in this industry and investigates the relationship between market structure, patent protection, and concentration.

To illustrate several basic propositions in the pharmaceutical industry I narrow the focus of the analysis to the cancer drug market. According to estimates published in 1999 (*Med Ad News*, May, 1999), the cancer drug market is over \$7.82 billion in annual sales. Within the cancer drug market I analyze the anti-cancer paclitaxel drug Taxol manufactured by the Bristol-Myers Squibb company.¹ Taxol sales are over \$1.2 billion per annum or over \$3 million per day. Taxol is principally prescribed for breast cancer, lung cancer, and ovarian cancer. My analysis considers the market segment consisting of drugs sold to treat these three cancers. The breast cancer drug market has annual sales of \$2.66 billion, the lung cancer drug market has annual sales of \$125.9 million, and the ovarian cancer drug market has annual sales of approximately \$643 million. Therefore the segment of the cancer drug market for these cancers alone is over 40 percent of all cancer drug sales. More importantly these three cancers have very high incidences in the United States. Breast cancer and lung cancer have the top two incidence rates in the United States among women (113.2 per 100,000 and 41.3 per 100,000 respectively) and breast, lung, and ovarian cancer represent over 60 percent of all cancer incidence in women

¹Taxol is trademarked by the Bristol-Myers Squibb company.

according to the National Cancer Institute SEER program. Among men, lung cancer is second only in incidence to prostate cancer with over 80.7 cases per 100,000 men being diagnosed with lung cancer each year. Thus the breast, lung, and ovarian cancer drug market is both large in terms of annual sales and large in terms of incidence in the U.S. population. The common denominator in the lung, breast, and ovarian cancer drug market is Taxol, which represents nearly 35 percent of annual sales of drugs supplied in this market segment. One purpose of this chapter is to define the relevant economic market for Taxol and to consider concentration in that market.

The remainder of the chapter is organized as follows. In Section 7.2, I discuss the market structure of the pharmaceutical industry and consider research and development, brand loyalty, and factors which lead firms to create and maintain monopolies. In Section 7.3, I consider one important barrier to entry in the pharmaceutical industry which is the statutory exclusivity provided by law in the United States. In Section 7.4, I discuss the history of paclitaxel including its discovery and development by the Bristol-Myers Squibb Company. In Section 7.5, I discuss various market power measures including the Herfindahl-Hirschman statistic and modifications to this statistic used in a multi-market setting. In Section 7.6, I discuss the relevant product market for paclitaxel drugs and issues of demand and supply substitutability. I also calculate the various measures of concentration presented in the previous section. Finally, in Section 7.7, I present conclusions.

7.2. Market Structure

In the pharmaceutical industry, companies face several difficult barriers to entering the market. First is the enormous research and development (R&D) cost in developing a drug and securing FDA approval for a drug. Second is the monopoly status sometimes conferred on a drug company by statutes. Additional barriers to entry include brand loyalty, generic introduction by a patent holder, and an incumbent's attempts to maintain its monopoly position. I discuss each in turn below.

7.2.1. R&D Economies of Scale

The economies of scale in R&D combine with the difficult and time consuming trials necessary to secure FDA approval for an "indication"—a treatment regimen for a new drug for a particular disease. A firm must maintain a portfolio of R&D projects because only 25 percent of new drugs that enter clinical testing are ever marketed. And of these, only 30 percent ever cover their total costs.² In a 1991 article, the chief executive officer of Glaxo noted

²See Viscusi, *et al.* (1998, p. 853).

that "to be a big player, a company must spend somewhere north of \$500 million per year (on R&D) and grow it by more than 10 percent or 15 percent a year. Those who can't spend that will be left behind."³

The R&D process lends itself to traditional economies of scale. As noted by Viscusi, *et al.* (1998), "...the discovery process is characterized by significant fixed costs. Multidisciplinary teams of biologists, chemists, and other scientists are engaged in research. . . . The clinical development process also recognizes significant regulatory and legal expertise, which is also characterized by fixed costs and specialization.⁴ Thus, the enormous R&D costs pose a significant barrier to entry."

Those R&D costs combine with the cost and time necessary to secure the FDA approval for an indication. Without such approval, the drug cannot be brought to market. Without doubt, the R&D costs and the difficulties in securing FDA approval are significant barriers to entry.

7.2.2. Monopoly Power from Statutory Exclusivity

As I noted above, drug companies expend a great deal of money, time and effort in R&D in bringing a new drug to market. In order to protect a company's investment and prevent other companies from using its discoveries and work, certain statutes, including the patent laws and the Hatch-Waxman Act, give companies varying degrees of exclusivity to market and sell their products. For instance, in the pharmaceutical industry, awarding a patent to a pharmaceutical company for a drug with an approved indication erects the ultimate barrier to entry during the patent's life. During such a period of exclusivity, a drug company can price like a monopolist in order to recoup its investment by charging prices above what they would likely have been in a competitive market.

7.2.3. Brand Loyalty

Intuitively, one would think that once a drug came off-patent and cheaper generic substitutes were made available, the branded drug would need to reduce price to meet the generics or face losing all its market share. However, research has shown that this has not been the case and that the branded drugs have proven to be remarkably resilient. For example, among eighteen drugs studied in 1992, generics were introduced at 61 percent of the branded drug's price. Over the course of two years, this price fell to 37 percent of the branded drug's price. The price of the branded drug actually rose 11 percent in this

³*The Wall Street Journal*, June 25, 1991.

⁴See Viscusi, *et al.* (1998, pp. 853-54).

period and retained about 50 percent of the market share.⁵ There are several explanations for this phenomenon.

First, drug companies have successfully bifurcated the market into price sensitive and price insensitive segments. With the introduction of much less expensive generics, the non-branded bio-equivalent product successfully captures the price sensitive market segment. One would suspect that as managed health organizations grow and other cost containment measures gain in popularity and importance, the price sensitive segment will grow. However, given the price insensitivity exhibited by some market segments, drug companies find it more profitable to raise prices, sell less product and cut production costs for the branded product than to compete with generics head-to-head on price.

Additionally, doctors bolster the resilience of the price insensitive group. Often doctors are risk averse, insensitive to cost, creatures of habit and will prescribe a brand name drug without concern or even knowledge of its cost, even when lower cost generic substitutes are available.⁶ Doctors become accustomed to prescribing the brand name drug during its patent tenure and other periods of exclusivity. Consequently, doctors typically are slow to switch to generics. Partly, this is due to the fact that as the prescriber, and not the consumer, there is no price consequence to the doctor. Additionally, pharmaceutical companies spend a great deal of money in promoting their branded drugs to doctors. This process is known as "detailing." This promotion during the period the drug enjoys exclusivity also strengthens the ties the doctor has to a particular branded drug.⁷

Consumers also often lack the knowledge to evaluate alternatives. This renders them unable to weigh the small risk associated with substituting away from a prescribed brand-name drug to an equivalent generic. Pharmacists apparently do not substitute the generic for the brand drug even when they have the legal ability to do so.⁸ These factors all present a barrier to entry to generic drug manufacturers once a drug loses its exclusivity.

7.2.4. Generic Drug Introduction

An additional barrier to entry can occur when the brand drug producer introduces its own generic version of the drug before its exclusivity expires. While it may at first seem counterintuitive to bifurcate the market before it is necessary to do so, there are good economic reasons to do so. The relation of price to marginal cost level may be less important than the price required to provide decreased incentives for rivals to enter the market. Thus, a company fore-

⁵See Viscusi, *et al.* (1998, p. 853).

⁶See Scherer (1993) and Temin (1980).

⁷See Leffler (1981).

⁸See Kralewski, Pitt, and Dowd (1983).

goes some short-term monopoly profits for long-term generic market power and profits.

This strategy is related to first-mover advantage, which is the competitive advantage a firm has by being the first to introduce a product into the market. By beating competitors to the market, the first-mover enjoys a monopoly period. A first-mover in the market can make binding commitments that adversely affect its competitors' ability to profit in the market. This is particularly true in the pharmaceutical industry, where almost all first-movers are the market share leaders.⁹

When the brand drug producer introduces a generic drug during the period of exclusivity, it is not only a first-mover, it is an only-mover. Thus, by introducing a generic drug while the exclusivity is still in effect, the incumbent can secure a large share of the generic market before others are legally entitled to enter and compete. This establishes a formidable barrier to entry.

7.2.5. Creating and Maintaining a Monopoly

A firm may become or remain a monopolist by several distinct methods. First, a firm may have special knowledge or trade secrets that allow it to produce a new or superior product that others cannot imitate. This differentiated product results in a differentiated demand for that firm, and allows the firm to exercise monopoly power with respect to that unique product.

Second, a firm may have specialized knowledge about production that allows it to produce a given product at lower costs, and therefore to be more efficient. Again, the cost structure may allow the firm to set a superior price for the product, resulting in an effective barrier to entry.

If, on the other hand, the product's cost structure or uniqueness can be imitated, a third method by which a firm can create or maintain its monopoly is through a patent. As I discussed above, a valid patent will prevent other firms from copying the product and also protect the firm from competition for a given period of time.

A fourth mechanism that allows a firm to create and maintain monopoly power is through government restrictions on entry. These include the partnerships created under the CRADA, the additional exclusivity that a firm is afforded under the Orphan Drug Act, or the protection that a firm receives under the Hatch-Waxman Act for five years of exclusion from ANDAs (generic applicants).

Fifth, firms may be prevented from entering a market by a firm's strategic actions. These might include controlling essential inputs or being sufficiently large in size to dominate a market where there is room for one large firm and

⁹See Grabowski and Vernon (1992).

little room for multiple firms. Typically, this occurs in local geographic markets, but may also occur in specialized sub-markets for chemotherapy agents.

Strategic actions taken by a firm to keep out other firms may, in fact, be violations of Section 2 of the Sherman Act.¹⁰

7.3. Statutory Exclusivity in the United States

The pharmaceutical industry is comprised of the companies that discover, manufacture and sell drugs that require a doctor's prescription. Although it is a relatively young industry, it ranks at the top of American industry in terms of spending on R&D. In order to encourage pharmaceutical companies to continue their research, several laws provide exclusive rights to those individuals or companies that invent a product or process. I discuss these below.

7.3.1. Pure Food and Drug Act of 1906 and Food, Drug and Cosmetic Act of 1938

The pharmaceutical industry's birth and growth was accompanied by the government's interest in regulating new drugs. This interest resulted in Congress passing the Pure Food and Drug Act of 1906. This was followed with the Food, Drug and Cosmetic Act of 1938. This statute required drugs to be approved by the FDA prior to their introduction into interstate commerce. This act was amended in 1962 by the Kefauver-Harris Amendments.

These amendments required the drug company that wished to introduce the new drug to prove that, in addition to being safe, the drug was effective. This required extensive controlled scientific experiments before the FDA would grant approval to any new drug. These amendments proved to be extremely costly to the drug companies, both in time and money. The FDA eventually approves only one in four drugs that begin clinical trials. This lengthy process also ate into the effective patent protection enjoyed by a new drug, decreasing the potential profits for the drug company. This shortened period of profit made it more difficult for the drug companies to recoup their R&D investment costs.

7.3.2. 1983 Orphan Drug Act

The Orphan Drug Act provides an exclusive seven-year right to market a drug that is used to treat a disease that affects less than 200,000 people. The FDA must decide whether the drug would be brought to market without the protection offered by this exclusive right (and the tax benefits discussed be-

¹⁰Section 2 of the Sherman Antitrust Act, Title 35 of the United States Code, provides that: "Every person who shall monopolize, or attempt to monopolize, or combine or conspire with any other person or persons to monopolize, any part of the trade or commerce among the several States, or with foreign nations, shall be deemed guilty of a felony. . . ."

low). The Orphan Drug Act grants the FDA the authority to grant exclusive approval for treating the *specific* disease. This exclusivity is virtually indistinguishable from a patent, except that it is shorter in duration. In addition to the exclusivity granted under this Act, a subsidy to the pharmaceutical companies was considered necessary to encourage drug companies to develop such new drug treatments because the small markets would not likely be profitable. Thus, the Act also provides a 50 percent tax credit for clinical trial expenses.

7.3.3. Drug Price Competition and Patent Restoration Act of 1984

In 1984, Congress enacted the Drug Price Competition and Patent Restoration Act (Hatch-Waxman). This statute simultaneously lengthened patent life and lowered the barriers faced by generic drugs after the patent expires. Patents are granted to provide temporary exclusivity to the patent holder. Generally, patents are considered to encourage three things: (1) R&D, and investment; (2) widespread dissemination of technology; and (3) provide a viable basis for further technology transfer to third parties.¹¹ However, pre-1984, the clinical testing and trial stage necessary to secure FDA approval consumed 12 years of a patent's life, reducing the average effective patent period, post-FDA approval, to only seven or eight years. This shortened period of exclusivity negatively affected the three things that patents were intended to accomplish. The 1984 Act remedied this problem by extending the effective patent life by a time equal to the sum of the FDA review time plus one-half of the clinical testing period. These extensions are limited to a maximum of five years, and a maximum 14 years of effective patent life.¹²

The act also lowered barriers to entry for the generic drug manufacturers. Before the Act, generic drug manufacturers were required to duplicate many of the clinical tests that the patent holder had originally performed in securing FDA approval for the drug. Under the 1984 Act, the generic drug manufacturer only must prove that its drug is bio-equivalent to the patented drug. Thus, the generic drug manufacturer needs only to submit an Abbreviated New Drug Application (ANDA). However, the Act also provides five years of limited exclusivity for companies that secure an NDA approval from the FDA.

¹¹See Yorke (1984).

¹²For example, prior to the Act, if a company filed for a patent for a new drug in 1980, but took 12 years to secure FDA approval for the drug's use, there would be only 8 years of exclusivity left. After the Hatch-Waxman Act, this same company would enjoy the 8 remaining years of exclusivity under the patent plus an additional five years, for a total of 13 years.

7.3.4. Exclusive Government Agency Licenses

The government itself often undertakes research in developing new drugs. In such instances, government research can be transferred to commercial companies only under certain licensing requirements. Generally, a governmental agency cannot grant an exclusive license if such a grant would tend to substantially lessen competition anywhere in the country in any relevant area of commerce. The Federal Technology Transfer Act creates a limited exception to these requirements when the license is granted under a CRADA. The CRADA typically contains the terms and conditions under which a private party can obtain clinical data and other fruits of government research in exchange for commercializing the government-developed technology and other consideration.

7.4. History of Paclitaxel

Paclitaxel is a complex plant alkaloid originally derived from the bark of the mature Pacific Yew tree, *Taxus brevifolia*, a slow-growing and long-lived tree sparsely distributed in the old growth forests of the Pacific Northwest. In the late 1960s, the National Cancer Institute (NCI)¹³ commenced a screening program involving 35,000 natural materials. Paclitaxel was discovered through this process. There was much controversy in the early years of development, as environmentalists sought to protect the yew.

Paclitaxel works by disrupting cancer cell growth and shrinking tumors. It has proven effective in late-stage ovarian and breast and lung cancer. It also is being used off-label to treat other cancers. Over 200,000 people die each year from ovarian, breast and lung cancer.¹⁴ Thus, paclitaxel has potentially great social and economic significance.

The U.S. government spent \$30 million to develop economically feasible methods to extract paclitaxel from Yew tree bark and create a clinically effective formulation that could be used to treat humans. In the early 1980s, NCI began to conduct clinical trials designed to test the efficacy of using paclitaxel for ovarian cancer treatments. Phase I testing, to introduce paclitaxel to humans, began in 1983. NCI started Phase II testing to examine paclitaxel's effectiveness in 1985. NCI began Phase III testing, the expanded, controlled clinical tests under expected medical conditions, in 1990. At about this time, NCI sought a commercial partner to bring paclitaxel through the FDA approval process, and to manufacture and distribute the drug. NCI sought a commercial partner, in part, because it did not have the resources to bring a paclitaxel

¹³NCI is an institute of the National Institutes of Health (NIH), which is in turn, a U.S. Government research institute that conducts and supports research for determining the cause, diagnosis, prevention and treatment of cancer.

¹⁴See Day and Frisvold (1993).

drug to market. To effectively accomplish this would require “forward linkages to pharmaceutical production and marketing and backward linkages into the forest products sector.”¹⁵ NCI had neither. Thus, to effectuate commercial paclitaxel production and marketing, NCI published, on August 1, 1989, its notice seeking a partner to enter into a Cooperative Research and Development Agreement (CRADA). A CRADA is used for joint research projects between private industry and public research institutions.¹⁶ In 1991, BMS and NCI entered into a CRADA to develop a paclitaxel drug to treat refractory ovarian cancer. Under the terms of the CRADA, BMS received the exclusive rights to all U.S. NIH funded research done under the CRADA. BMS agreed to provide NIH with 17 kilos of Taxol and use its best efforts to commercialize Taxol. Subsequent agreements with the USDA and Department of Interior gave BMS the exclusive right to harvest Yew trees on Forest Service lands.

In 1992, the FDA approved BMS’ New Drug Application (NDA) 20–262 seeking approval to market Taxol for treating refractory ovarian cancer. All the data produced by BMS was apparently based on NIH sponsored research and clinical trials. With the FDA’s approval, BMS also received a five-year period of non-patent exclusivity to market Taxol.

On December 3, 1996, NCI and BMS entered into an agreement to extend the CRADA until December 1997 and exclusive license agreements. These agreements gave BMS the exclusive rights to three NCI inventions involving Taxol, including a patent for a 96-hour infusion therapy using a paclitaxel drug to treat breast cancer.

BMS has been extremely successful in marketing Taxol. In 1999, BMS held nearly 100 percent of the market share drugs used to treat ovarian cancer, the indication for which Taxol was first approved. BMS has also made significant inroads into other cancer treatments, where it holds 48 percent of the market share for drugs used to treat breast cancer, and 96 percent of the market share for drugs to treat lung cancer. Thus, by obtaining under the CRADA the government research data, BMS was able to file an NDA for ovarian cancer. Subsequent FDA approval provided them with limited exclusivity by statute (which I discuss in greater detail below). BMS has subsequently secured a monopoly position with regards to cancer treatments for breast cancer, lung cancer and Kaposi’s sarcoma through off-label sales of Taxol.¹⁷ Undeniably, BMS did invest substantial time and money in getting FDA approval for the drug and marketing the drug. However, in exchange for bringing the drug through the FDA approval process and to market, BMS received limited exclu-

¹⁵See Day and Frisvold (1993).

¹⁶CRADAs are authorized by the Federal Technology Transfer Act of 1986.

¹⁷Off-label uses of Taxol and other anti-cancer drugs are well documented. Some 70 to 80% of all approved oncology therapies are used outside an indication.

sivity for Taxol to treat ovarian cancer, exclusive rights to harvest Yew trees on federal land, and exclusive rights to a huge array of knowledge generated from NCI's investment of public funds. BMS has leveraged these advantages into significant monopoly power in the paclitaxel market for treating cancers other than the one for which it first received an FDA indication.

7.5. Market Power Measures

7.5.1. Herfindahl-Hirschman Index (HHI)

The HHI is a widely accepted method to measure market concentration. It involves first determining the relevant geographic and product markets, market participants, and each participant's market share. The HHI is calculated by squaring the market share possessed by each participant firm selling a particular good in a specific, well-defined geographic market and then summing the squares across all firms in the market.

At one extreme is an industry with a single firm selling all the output. The HHI for a pure monopoly is calculated by taking the market share of 100 percent and squaring it. Thus, a pure monopoly yields an HHI of 10,000. At the other extreme is a pure competitive industry where each participant firm's market share is about zero.¹⁸ Accordingly, the possible HHI results range from 0 to 10,000.

A great deal can be learned by calculating a specific firm's market share. When the market share is close to zero, the firm would not be able to manipulate price to increase profits because it would be too small. When market share is close to 100 percent, the firm would generally be able to set or, at a minimum, greatly influence prices in the market. This is generally accomplished by using its market size to control output. A dominant supplier might be able to reduce output to drive up prices and maximize its profit without fear of encouraging competitive entry or response.

This interest in overall competitiveness (or lack thereof) is also related to concern over the prospect for collusion. In general, an industry with many small suppliers would not likely have the ability to collude by controlling supply or entry in order to drive up price. However, collusion can be relatively easy in an industry with a small number of firms, each with a significant market share. And, even a small firm might collude with a dominant, quantity-controlling, price-setting segment of the industry. Accordingly, the market's overall concentration is an important factor used to determine the potential for pricing above marginal cost, or the competitive market standard.

¹⁸A market with N participants of equal size has an HHI of $\sum_{i=1}^N (1/N)^2 = N/N^2 = 1/N$. As N gets large (pure competition), the HHI approaches zero.

By itself the HHI does not determine whether or not anti-competitive behavior is actually taking place by a particular firm or a group of firms in a particular market. Instead, the statistic is used to establish ranges, which in some cases suggest the need for additional analysis. Generally, an HHI of 0 to 1,000 or less is used to conclude that a market is competitive.¹⁹ An HHI of 1,000 to 1,800 is usually thought to represent a workably competitive market. Regulators and anti-trust enforcers often presume that an HHI over 1,800 indicates the potential for a firm to exercise market power.²⁰ The HHI should be coupled with analyses (such as mitigation or residual regulation) to ensure that anti-competitive behavior does not take place. Other factors, therefore, may be considered to establish that a firm lacks market power even if HHIs appear relatively high.

There are significant practical problems with market share and market concentration measures. In practice, defining the relevant geographic and product markets can be somewhat ambiguous and is nearly always controversial. Additionally, not all markets have a homogeneous product. Nevertheless, competition can be fierce. Consider clothing. Defining the product as blouses will necessarily yield different firm share and market concentration measurements than when the product is defined alternatively as ladies' apparel, clothing, department stores, etc. Some analysts may include mere substitutes such as natural gas and fuel oil, in the same market. Other analysts may exclude them. A certain degree of arbitrariness and subjectivity inevitably enters into the firm share and market concentration calculations.

Another complication arises in those industries in which firms might simultaneously compete in multiple markets. The traditional HHI measures market power for firms selling a product in a single market. This analysis does not help determine an industry's competitiveness if the various markets do not function as a single market. The traditional HHI also does not help determine a single firm's market power across the various markets in which it competes.

The Federal Energy Regulatory Commission (FERC) Staff²¹ has suggested variations on the traditional HHI analysis for industries where firms compete in

¹⁹"The Agency divides the spectrum of market concentration as measured by the HHI into three regions that can be broadly characterized as unconcentrated (HHI below 1000), moderately concentrated (HHI between 1,000 and 1,800), and highly concentrated (HHI above 1,800)." *Horizontal Merger Guidelines*. See also, O'Neill, *et al.* (1991).

²⁰It is generally acknowledged that the 1,800 threshold recommended in the Merger Guidelines is used as a level of concentration that triggers intensive scrutiny. The Merger Guidelines explain that, although there is a presumption of market power, it may be overcome by showing that factors set forth in the Guidelines make it unlikely that the merger will create or enhance market power or facilitate its exercise, in light of market concentration and market shares.

²¹"On Developing a Framework for Assessing Competition in Natural Gas Transportation," Appendix C—A Further Discussion of Market Power and Competition Measures of Firms Within Connected Multi-Market Industries"; July 1989.

several markets simultaneously. The pharmaceutical industry fits this scenario. These analytic extensions are the Weighted HHI, the Weighted Market Share and the Market Share Weighted HHI. The following describes these three concepts.

q_{ik} Sales made by firm i in market k
 $i = 1, 2, \dots, I \quad k = 1, 2, \dots, K$

S_{ik} Market share of firm i in market k within the K markets
 $= q_{ik} / \left(\sum_i q_{ik} \right)$

α_{ik} Market share of firm i across the K markets
 $= q_{ik} / \left(\sum_k q_{ik} \right)$

HHI_k Herfindahl in market k
 $= \sum_i S_{ik}^2$

WMS_i Weighted market share for firm i
 $= \sum_k \alpha_{ik} S_{ik}$

$WHHI_i$ Weighted HHI for firm i
 $= \sum_k \alpha_{ik} HHI_k$

$MSWH_i$ Market share weighted HHI
 $= \left(\sum_k \alpha_{ik} S_{ik} HHI_k \right)^{1/2}$

7.5.2. Weighted HHI

This measure can be used to estimate the overall concentration of a set of markets. It can be used to analyze the entire industry, or particular parts (*e.g.*, the concentration of markets served by any particular firm). For example, when a pharmaceutical firm (i) competes across K product markets, the weighted HHI can be used to determine the overall concentration of the combined market in which the firm competes. The Weighted HHI is measured by summing the HHIs for each market in which the i th firm operates, weighting by the percent of the i th firm's sales made in each product market. This statistic describes the overall competitive conditions faced by the i th firm across the K markets in which it operates.

7.5.3. Weighted Market Share Statistic

This measure can be used to estimate the overall degree to which any individual company tends to dominate its markets. The Weighted Market Share

HHI, which is used in multi-market conditions, determines the weighted market share for the i th firm across the K markets in which it operates. This statistic is calculated by summing the i th firm's market share in market k , weighted by the i th firm's share of overall sales in sub-market k across all K markets. This statistic explains the i th firm's overall market power across the K markets.

7.5.4. Market Share Weighted HHI

This measure has been proposed by FERC Staff to estimate the overall degree to which a firm might benefit from its specific market share combined with the overall market concentration in the specific markets it serves. The Market Share Weighted HHI measures the potential for anti-competitive behavior confronting the i th firm across multiple markets by investigating the firm's share of overall market concentration in each of the markets in which the firm competes. The FERC Staff recommends combining the Weighted HHI and Weighted Market Share concepts to measure the market power confronting or exerted by the i th firm operating across the K markets.

The weighted average of market specific HHIs multiplied by the firm's market share in each sub-market is taken across each of the K markets in which the i th firm competes. The square root of this calculation is taken to return the estimate to a range similar to the traditional HHI calculation. This approach produces results similar to market power measures under the polar cases of pure monopoly and perfect competition. Implicitly, it also establishes a market power measure for the intermediate stages of competition.

Under some conditions, it can be shown that the FERC staff's proposed multimarket measure is related to the weighted average mark-up in the markets in which a specific firm operates. While the HHI acts as a screen to ascertain whether further inquiry is necessary, the weighted average mark-up calculation can be used in a similar fashion to determine whether there is any market power abuse across markets. Analysis of the multi-market HHIs demonstrates the central role of the firm's price elasticity of demand in determining the potential for supra-competitive price markups above marginal cost.

Demand elasticities, if they can be ascertained, will provide a more exact measure of any rationale for such mark-ups. Thus, the HHI and the multi-market HHI, while important as screening tools, and useful as measures of supply substitutability, can be complemented with a more detailed analysis if demand elasticities can be determined.

7.6. The Relevant Product Market for Paclitaxel Drugs

The extent to which a firm has market power for a particular product in a specified market area is determined by demand substitutability, supply substitutability, and firm entry. Demand substitutability refers to buyers' ability to

substitute away from the marketed good in the relevant market and replace it with another product. Supply substitutability and entry both generally refer to a firm's ability to supply a product in a relevant market even if it currently does not supply that market. Supply substitutability refers to actions taken by existing firms not currently selling in a relevant market.²² Entry refers to a form of supply substitutability that requires significant investment and distribution. However, the U.S. Department of Justice and FTC merger guidelines have indicated that supply substitutability that takes longer than one year to occur is not actually a form of entry. Therefore, such firms are not considered as competitors for evaluating market power.

7.6.1. Antitrust Markets

For antitrust purposes, a market is defined as a "product or group of products, and a geographic area in which it is produced or sold such that a hypothetical, profit-maximizing firm, not subject to price regulation, that was the only present and future producer or seller of those products in that area likely would impose at least a small, but significant and non-transitory increase in price, assuming the terms of sale of all other products are held constant."²³ This definition reflects the separation between demand substitutability and supply substitutability. The relevant market for antitrust purposes may be only one of several overlapping and intersecting markets. Thus, the relevant market is "the smallest group of products and geographic area that constitutes a market."²⁴ The main condition constituting a market is that a hypothetical monopolist could successfully raise price by a "small but significant and non-transitory" amount above current and likely future price levels.²⁵

In evaluating the extent to which a price increase would be possible for a monopolist, it is therefore necessary to consider both demand and supply substitution possibilities.

7.6.2. Market Delineation—Demand Substitutability

Several fairly general factors may be relevant in delineating a product market:

1. Evidence reflecting buyers' perceptions that the products are or are not substitutes. This is particularly relevant if those buyers have shifted purchases

²²Firms that can easily enter a relevant market are considered to be competitors in the market and are assigned market shares in HHI calculations even if they currently have no sales in the relevant market. The examination of demand substitutability revealed by demand elasticity and cross-elasticity as well as the examination of supply substitutability allow a careful delineation of a relevant market.

²³Horizontal Merger Guidelines (1992).

²⁴See Werden (1981).

²⁵Horizontal Merger Guidelines (1992).

between the products in response to changes in relative price or other competitive variables;

2. Similarities or differences in the products' price movement over time;
3. Similarities or differences in the products' customary usage, design, physical competition, and other technical characteristics; and,
4. Evidence of sellers' perceptions that the products are or are not substitutes, particularly a change in a company's business plan that has been based on those perceptions.

7.6.3. Market Delineation—Supply Substitutability

The degree to which the market constrains a firm's ability to act monopolistically depends in part on short-term entry. To analyze the potential for competitive supply response, I consider the concentration of the relevant markets, as measured by the Herfindahl-Hirschman Index (HHI).

The HHI measures the degree of competition and the potential for a competitive response to monopolistic behavior via supply substitution. Also, as discussed above, the elasticity of demand measures the substitutability of products in consumption as prices are increased. The two concepts are related: assuming that the interaction of firms is Cournot competition in market shares, constant marginal costs, and constant elasticity of demand, the mark-up above marginal cost in an industry (Lerner's index) will be equal to the HHI divided by the market demand elasticity. Concentrated markets or markets with small elasticity of demand (few substitutes) will have predictably higher mark-ups above incremental cost. Conversely, unconcentrated markets or markets with large elasticity of demand (many substitutes) will have low mark-ups and therefore approximately competitive outcomes.

7.6.4. The Relevant Product Market for Taxane Drugs

In defining the relevant product market for taxanes and related anti-cancer drugs, it is helpful to consider the manner in which anti-cancer drugs are allocated.²⁶ One way to divide the anti-cancer drug market is by therapeutic indication or use. As discussed above, paclitaxel drugs are principally used against several major cancers, including lung cancer, breast cancer, ovarian cancer, and Kaposi's sarcoma. The oncology literature suggests that anti-cancer drug therapy for a patient with one type of cancer is not necessarily substitutable for the anti-cancer drug therapy for a patient with a different type of cancer.

²⁶The taxane market consists of antineoplastic drugs in the taxoid family. Currently, this includes only docetaxel (Taxotere) and paclitaxel (Taxol). See Nicolaou, *et al.* (1996).

Therefore, those treatments and drugs that are appropriate for lung cancer may be inappropriate for treating ovarian cancer.

Chemotherapy regimens broadly include cytotoxic chemotherapies and hormonal chemotherapies. Cytotoxic chemotherapies provide a toxic action to cells. Hormonal chemotherapies target a tumor's hormonal sensitivity. One such example is tamoxifen, which is used against estrogen sensitive tumors in breast cancer. Within the cytotoxics, various chemotherapies include antimetabolites, anthracyclines, alkylating agents, mitotic inhibitors, and miscellaneous antineoplastics.

The manner in which each drug class acts on cancer differs. For instance, taxanes are classified by *1999 Drug Facts and Comparisons* as miscellaneous antineoplastics. Taxanes affect cell division by a specific mechanism. Conversely, mitotic inhibitors will affect cell division by different actions or mechanisms. Thus, although mitotic inhibitors and other antineoplastics affect cell division, each does so using a different mechanism.²⁷

Certainly, there can be substitution between chemotherapy drugs in clinical application. However, the substitution is limited by the type of cancer, the chemotherapy regimen, and the type of chemotherapy chosen. For instance, those molecules appropriate to treating a recurrent or refractory form of breast cancer are likely to be a subset distinct from the molecules that can treat all cancer forms of any degree of severity. Relevant product markets must reflect this differentiation. Thus, the relevant product market should reflect only those molecules appropriate to treat a specified type of cancer and therapy.

Arguably, a relevant product market could reflect only those molecules appropriate for treating breast cancer for third line or salvage therapy. Within such a market segment, the first and second line drugs have been exhausted, patients have developed resistance to first and second line therapies, and only a subset of molecules are available for treating these recurrent or refractory breast cancers. Similarly, those molecules appropriate for treating breast can-

²⁷ *Alkylating agents* are compounds with the ability to replace hydrogen atoms of certain organic compounds with alkyl groups. The alkylation of nucleic acids such as DNA produces breaks in the DNA molecules or cross-linking of DNA strands which interfere with DNA replication. The classic alkylating agents include cyclophosphamide. Cisplatin and carboplatin are considered alkylating agents although their mechanism of action is more complex.

Antibiotics (including Doxorubicin (Adriamycin)) produce their tumoricidal effects by binding to DNA usually by interposition between base pairs. *Anthracyclines* form a subset of the antibiotics. Plant Alkaloids include both the *antimitotic* and *miscellaneous antineoplastic* class of molecules. These agents bind to the microtubular proteins found in dividing cells. Since these microtubules are essential for cell division, the binding leads to mitotic arrest (antimitotic action). The antineoplastic class of agents have different cytotoxic effects on the cells. Paclitaxel is unique among the antineoplastics because it promotes the assembly of microtubules that inhibit the reorganization required during cell division.

Antimetabolites (including Methotrexate, and 5-fluorouracil) interfere with normal cell functions by interacting with or replacing normal cellular enzymes. *Hormonal* agents function by interacting and binding to specific cell receptors. The target of hormonal agents is endocrine manipulation. For further discussion see Haskell (1995).

cer are often different from the molecules appropriate for treating lymphoma or prostate cancer. Here, the distinction lies in the type of chemotherapy or in the chemotherapy's action on the cancer.

I would expect the HHI that ignores the chemotherapy regimen to be a conservative estimate of the HHIs that do reflect the chemotherapy regimen, (*i.e.*, by differentiating between first line and second line treatments). In the analysis which follows, the HHIs I calculate do not differentiate between chemotherapy regimen. However, this does not change the fact that these sub-markets are relevant for antitrust purposes, and do exist within the types of cancer being treated. Importantly, the HHIs I calculated are lower than those I would expect for relevant sub-markets.

7.6.5. HHIs—Market Concentration Results

The HHI analysis is based on data taken from the IMS *Price/Quantity Report on Oncology Drugs 1999*, the *1999 Drug Facts and Comparisons*²⁸ and *Cancer Chemotherapy Pocket Guide*.²⁹ I took the annual sales data for each drug from the IMS study. As discussed above, I am primarily interested in breast, lung, and ovarian cancers as they are the major targets of the taxane family. The IMS sales data do not specify the type of cancer the drug was used to treat. Thus, in order to separately identify a "Breast Cancer" market from an "Ovarian Cancer" market I allocated annual sales dollars by cancer types. The allocation is based on a survey of physicians conducted annually by IMS known as the U.S. Appearance data. Doctors are asked to record their uses of all drugs and for what applications the drug was prescribed. Combination therapies (multiple drugs) are also recorded. The survey is a random probability sample and is used to produce national estimates. Tables 7.1.1 through 7.1.3 show the type of chemotherapy agent, the relevant molecule, the drug names for that molecule, the sales in 1999 by company and the various percentages of sales allocated to breast, lung, ovarian, and other cancers. The latter percentages were derived from the IMS U.S. Appearance Data. I recorded all drugs indicated for treating neoplasms of the lung, breast or ovary. I excluded any drug whose sales were sufficient to be included in the U.S. Appearance Data but for which the indicated treatment was other than breast, lung or ovarian.

In some cases, I assumed that the percentage recorded in the U.S. Appearance Data for a specific drug was representative of all drugs for the "family" molecule. For instance, I assumed that the surveyed sales of VePesid penetration were characteristic of all etoposides. In cases when a specific drug showed

²⁸ *1999 Drug Facts and Comparisons* was used principally for the assignment of chemotherapy types as well as to determine off-label uses.

²⁹ *Cancer Chemotherapy Pocket Guide*, Ignoffo (1998) was used to determine indicators for anti-cancer drugs and to identify the relationships between the trade and molecular names of drugs.

more specific information, I did not apply this assumption. For example, I extended the relevant penetration percentages for Fluorouracil to the fluorouracil family of molecules, with the exception of the topical Efudex which separately revealed zero penetration for the treatment of breast, lung, or ovarian cancer. The total sales represented in Tables 7.1.1 through 7.1.3 are \$2.969 billion. The sales allocated to the breast, lung, and ovarian cancer market, are \$1.4 billion.

The breast cancer, lung cancer, and ovarian cancer drug markets have estimated annual sales of \$2.65 billion, \$125.98 million, and \$642.67 million respectively according to *Med Ad News* (May, 1991). Their estimate for the combined segment is therefore \$3.42 billion in annual sales. However, I am aware that some drugs with small sales amounts (too small to appear in the survey of physicians) are not included in my analysis. Since the HHI analysis relies on shares, this omission is not significant.

The HHI results are presented in Tables 7.2.1 through 7.2.6. The analysis is presented for the top six firms based on 1999 sales in the sub-markets for breast, lung, and ovarian cancer drugs. The second column of these tables demonstrates the pharmaceutical company's total sales in 1999 within the market segments. The third column demonstrates the firm's market share across the relevant sub-market. For instance, in this column I observe that BMS' sales for treating breast cancer represent 18.9 percent of its total chemotherapy drug sales used to treat breast, lung, or ovarian cancer. Additionally, 20.4 percent of BMS' sales are for treating lung cancer, 38.2 percent are for treating ovarian cancer, and 22.5 percent are for treating other cancers. These rounded percentages total 100 percent.

Within a category such as the breast cancer market, I first identified all drugs indicated for treating breast cancer by their appearances in the IMS survey. I then identified all drug manufacturers that supply either branded or generic drugs to this sub-market. To calculate a firm's market share within a market (the column labeled "Market Share Within Market"), I first totaled all manufacturers' sales in that sub-market and then determined the specific firm's market share in the sub-market. The last column identifies the firm's market share within a sub-market. For example, in the breast cancer market, I estimated that BMS provides 47.8 percent of all drugs for treating breast cancer. Recall that I previously noted that BMS' drug sales for breast cancer treatment represent 18.9 percent of its supply of drugs. Thus, while BMS supplies roughly half of all drugs in this sub-market, only 20 percent of its drug sales were used to treat breast cancer.

I used the sales information for all manufacturers in the various sub-market segments to calculate market shares within markets. I then estimated the HHI for each market segment. Within the breast cancer market, for instance, I estimate the HHI to be over 3,700. Importantly, BMS' 48 percent market share indicates that the breast cancer market is very concentrated, and that BMS is a

Table 7.1.1. Alkylating Agents, Anthracyclines, Antibiotics

Alkylating Agents							
Molecule	Drug	1999 Dollars (000s)	Company	Breast	Lung	Ovarian	Other
CARBOPLATIN	PARAPLATIN	\$ 390,590	BMO	2.0%	39.0%	34.0%	25.0%
CHLORAMBUCIL	LEUKERAN	\$ 8,512	GWC	0.0%	0.0%	0.0%	100.0%
CISPLATIN	PLATINOL	\$ 14	BMO	1.0%	21.0%	30.0%	48.0%
	PLATINOL AQ	\$ 107,500	BMO	1.0%	21.0%	30.0%	48.0%
CYCLOPHOSPHAMIDE	CYTOXAN	\$ 34,124	BMO	50.0%	3.0%	4.0%	43.0%
	NEOSAR	\$ 2,330	PHU	50.0%	3.0%	4.0%	43.0%

Anthracyclines							
Molecule	Drug	1999 Dollars (000s)	Company	Breast	Lung	Ovarian	Other
DOXORUBICIN	ADRIAMYCIN	\$ 4	PHU	45.0%	3.0%	2.0%	50.0%
	ADRIAMYCIN PFS	\$ 12,876	PHU	45.0%	3.0%	2.0%	50.0%
	ADRIAMYCIN RDF	\$ 3,528	PHU	45.0%	3.0%	2.0%	50.0%
	DOXIL	\$ 46,776	ALZ	45.0%	3.0%	2.0%	50.0%
	DOXORUBICIN	\$ 10,620	GNA	45.0%	3.0%	2.0%	50.0%
	DOXORUBICIN HCL	\$ 16,464	AP+, BDF, FUJ	45.0%	3.0%	2.0%	50.0%
	RUBEX	\$ 406	BMO	45.0%	3.0%	2.0%	50.0%

Antibiotics							
Molecule	Drug	1999 Dollars (000s)	Company	Breast	Lung	Ovarian	Other
ETOPOSIDE	ETOPOPHOS	\$ 2,762	BMO	3.0%	52.0%	4.0%	41.0%
	ETOPOSIDE	\$ 14,058	A-P, AP+, BDF, C/T, GNA, MRP, SEI, SUP	3.0%	52.0%	4.0%	41.0%
	TOPOSAR	\$ 48	PHU	3.0%	52.0%	4.0%	41.0%
	VEPESID	\$ 34,252	BMO	3.0%	52.0%	4.0%	41.0%

Table 7.1.2. Anti-Metabolites, Antineoplastics

Anti-Metabolites							
1999 Dollars							
Molecule	Drug	(000s)	Company	Breast	Lung	Ovarian	Other
FLUOROURACIL	ADRUCIL	\$ 3,448	PHU	22.0%	2.0%	0.0%	76.0%
	EFUDEX	\$ 33,894	ICN	0.0%	0.0%	0.0%	100.0%
	FLUOROPLEX	\$ 2,642	ALL	22.0%	2.0%	0.0%	76.0%
	FLUOROURACIL	\$ 878	FUJ, ICN, SOK	22.0%	2.0%	0.0%	76.0%
METHOTREXATE	FOLEX PFS	–	PHU	0.0%	0.0%	0.0%	100.0%
	METHOTREXATE	\$ 53,394	AGN, BRR, E/L, G.G, IMX, MHL, MJR, MNS MYN, PRL, QLT, RKG ROX, SEI, SUP, INR WTS, Z/G	0.0%	0.0%	0.0%	100.0%
	METHOTREXATE SOD	\$ 9,200	BDF, C/T, IMX, MNS	33.0%	2.0%	0.0%	65.0%
Antineoplastics							
1999 Dollars							
Molecule	Drug	(000s)	Company	Breast	Lung	Ovarian	Other
HYDROXYUREA	HYDREA	\$ 10,882	BMO	0.0%	0.0%	0.0%	100.0%
	HYDROXYUREA	\$ 17,364	BRR, DPI, P.H, QLT ROX, UNR	0.0%	0.0%	0.0%	100.0%
LEVAMISOLE	ERGAMISOL	\$ 3,130	JAN	0.0%	0.0%	0.0%	100.0%
PACLITAXEL	TAXOL	\$ 208	BMO	29.0%	11.0%	45.0%	15.0%
	TAXOL SEMI-SYN	\$ 809,436	BMO	29.0%	11.0%	45.0%	15.0%
Mitotic Inhibitors							
1999 Dollars							
Molecule	Drug	(000s)	Company	Breast	Lung	Ovarian	Other
VINBLASTINE	VELBAN	\$ 208	LLY	12.0%	9.0%	0.0%	79.0%
	VINBLASTINE SULF	\$ 2,004	AP5, BDF, C/T FDG, SEI, SES	12.0%	9.0%	0.0%	79.0%
VINCRISTINE	ONCOVIN	\$ 1,366	LLY	8.0%	5.0%	0.0%	87.0%
	VINCASAR PFS	\$ 1,404	PHU	8.0%	5.0%	0.0%	87.0%
	VINCRISTINE	\$ 2	SES	8.0%	5.0%	0.0%	87.0%
	VINCRISTINE SULF	\$ 1,442	FDG, SEI	8.0%	5.0%	0.0%	87.0%

Table 7.1.3. Hormones

Hormones							
Molecule	Drug	1999 Dollars (000s)	Company	Breast	Lung	Ovarian	Other
FLUTAMIDE	EULEXIN	\$ 55,022	SHC	0.0%	1.0%	0.0%	99.0%
GOSERELIN	ZOLADEX	\$ 188,448	ZNC	1.0%	0.0%	0.0%	99.0%
LEUPROLIDE	LEUPROLIDE ACET	\$ 3,214	BDF, VIE	0.0%	0.0%	0.0%	100.0%
	LUPRON	\$ 26,050	TAP	0.0%	0.0%	0.0%	100.0%
	LUPRON DEPOT	\$ 249,838	TAP	0.0%	0.0%	0.0%	100.0%
	LUPRON DEPOT-3 MO.	\$ 256,636	TAP	0.0%	0.0%	0.0%	100.0%
	LUPRON DEPOT-4 MO.	\$ 201,338	TAP	0.0%	0.0%	0.0%	100.0%
MEGESTROL	MEGACE	\$ 4,066	BMO	40.0%	5.0%	1.0%	54.0%
	MEGESTROL	\$ 2	AGN, WCT	40.0%	5.0%	1.0%	54.0%
	MEGESTROL ACETATE	\$ 13,602	BJJ, BRR, DSE, GNC MHL, MJR, MNS, MTC P.H, PME, PRL, QLT RKG, ROX, RSM, SEI SUP, TEV, UNR, WTS Z/G	40.0%	5.0%	1.0%	54.0%
TAMOXIFEN	NOLVADEX	\$ 66,558	ZNC	76.0%	0.0%	1.0%	23.0%
	TAMOXIFEN CITRATE	\$ 268,690	BRR, MNS, RKG	74.0%	0.0%	0.0%	26.0%
TOTAL:		\$ 2,969		\$ 568 \$ 1,400	\$ 297	\$ 536	\$ 1,569

major source of that concentration. Also, Barr labs supplies over 36.6 percent of the market share, providing another major source of the market concentration.

7.6.6. Market Share Weighted HHI

As I discussed above, I also calculate a market share weighted HHI that reflects a sub-market's concentration, a manufacturer's share in that market, and the relative weight to give the sub-market from the manufacturer's perspective. This measure is larger whenever (i) the HHI for the market is larger, (ii) the share of a market is larger in that sub-market relative to other sellers or (iii) when a firm has a greater proportion of its sales in that sub-market. To calculate this market share weighted HHI, I multiply a manufacturer's market share within the sub-market and the HHI, and then take a fraction representing the importance of that sub-market to the manufacturer compared to the other sub-markets in which it supplies product.

Thus, to calculate the market share weighted HHI for the combined breast, lung, and ovarian cancer market, I first take BMS' 18.9 percent of sales in the breast cancer market (among breast, lung, ovarian, and other cancers in this sub-market). I then multiply this percentage by the sub-market's HHI of 3,744. The product is then multiplied by BMS' market share within the market (47.8 percent) to arrive at the first component of the market share weighted HHI. Similar calculations are then performed for the other market segments (lung cancer, ovarian cancer, and other cancers). Next, I sum the results and use the square root of the resulting sum to derive the overall market share weighted HHI in this sub-market. Using this technique, I calculate the overall market share weighted HHI to be 7,787.

Tables 7.2.1 through 7.2.6 demonstrate the degree of concentration enjoyed by BMS in this market segment. Not only do BMS' market shares within the sub-markets demonstrate that these markets are concentrated, they also demonstrate that the concentration primarily arises due to BMS' dominance within the market segments. In fact, the markets all uniformly concentrated. The HHI for the breast cancer market is 3,744; the HHI for the lung cancer market is 9,279; the HHI for the ovarian cancer market is 9,915; and the HHI for the other cancer market is 2,899. The weighted market share across markets is 71.3 percent, the weighted HHI is 7,039, and the market share weighted HHI is 7,787. These statistics indicate large market power in the relevant sub-market and even larger market power in the relevant sub-market for each cancer type.

Table 7.2.1. Concentration—Bristol-Myers Oncology (1999 Dollars)

Cancer Type	Sales	MS Across Markets	Mkt HHI	MS w/in Market
Breast	263,665	18.91%	3,744	47.86%
Lung	284,455	20.40%	9,279	96.31%
Ovarian	532,289	38.18%	9,915	99.57%
Other	313,831	22.51%	2,899	20.36%
All	1,394,240	100.00%	3,089	47.71%
Summary Statistics:				
Wtd Market Share		71.3%		
Weighted HHI		7,039		
Mkt Share Wtd HHI		7,787		

Table 7.2.2. Concentration—TAP Pharmaceuticals (1999 Dollars)

Cancer Type	Sales	MS Across Markets	Mkt HHI	MS w/in Market
Breast	–	0.00%	3,744	0.00%
Lung	–	0.00%	9,279	0.00%
Ovarian	–	0.00%	9,915	0.00%
Other	733,862	100.00%	2,899	47.60%
All	733,862	100.00%	3,089	25.11%
Summary Statistics:				
Wtd Market Share		47.6%		
Weighted HHI		2,899		
Mkt Share Wtd HHI		3,715		

7.7. Conclusion

The pharmaceutical industry is an extremely complex, high stakes industry. The R&D costs and costs associated with securing FDA approval for a new drug are enormous. Conversely, the profits that can be made by introducing a new blockbuster drug can be equally huge. Statutes providing exclusivity to pharmaceutical companies attempt to create a balance between the companies' incentives in bringing a new drug to market and the public's ability to get the benefit of competitively provided generic substitutes to branded drugs. The result is a complex system of statutory exclusivity, including patents, CRADAs, and special protections accorded to orphan drugs.

Table 7.2.3. Concentration—Zeneca Pharmaceuticals (1999 Dollars)

Cancer Type	Sales	MS Across Markets	Mkt HHI	MS w/in Market
Breast	57,447	22.53%	3,744	10.43%
Lung	–	0.00%	9,279	0.00%
Ovarian	666	0.26%	9,915	0.12%
Other	196,893	77.21%	2,899	12.77%
All	255,006	100.00%	3,089	8.73%
Summary Statistics:				
Wtd Market Share		12.2%		
Weighted HHI		3,108		
Mkt Share Wtd HHI		1,934		

Table 7.2.4. Concentration—Barr Labs (1999 Dollars)

Cancer Type	Sales	MS Across Markets	Mkt HHI	MS w/in Market
Breast	201,612	69.69%	3,744	36.60%
Lung	354	0.12%	9,279	0.12%
Ovarian	71	0.02%	9,915	0.01%
Other	87,241	30.16%	2,899	5.66%
All	289,278	100.00%	3,089	9.90%
Summary Statistics:				
Wtd Market Share		27.2%		
Weighted HHI		3,498		
Mkt Share Wtd HHI		3,170		

It is within this context that BMS developed a relationship with NCI that resulted in a CRADA between NCI and BMS. Pursuant to that CRADA, BMS received exclusive access to NIH data, allowing BMS to obtain FDA approval for Taxol, initially for refractory ovarian cancer, and five years of Hatch-Waxman exclusivity using the research and clinical trials that NCI had completed at public expense.

However, the pharmaceutical industry is unique in that once a drug has been approved for one indication, it is often prescribed by doctors for other treatments that have not been approved by the FDA. This off-label use allowed BMS to expand its Taxol sales for indications such as breast cancer, lung cancer, and other cancers. While BMS received limited exclusivity in certain sub-

Table 7.2.5. Concentration—Pharmacia & Upjohn (1999 Dollars)

Cancer Type	Sales	MS Across Markets	Mkt HHI	MS w/in Market
Breast	8,563	36.23%	3,744	1.55%
Lung	957	4.05%	9,279	0.32%
Ovarian	375	1.59%	9,915	0.07%
Other	13,742	58.14%	2,899	0.89%
All	23,638	100.00%	3,089	0.81%
Summary Statistics:				
Wtd Market Share		1.1%		
Weighted HHI		3,575		
Mkt Share Wtd HHI		612		

Table 7.2.6. Concentration—Glaxo Wellcome Oncology (1999 Dollars)

Cancer Type	Sales	MS Across Markets	Mkt HHI	MS w/in Market
Breast	134	8.53%	3,744	0.02%
Lung	87	5.53%	9,279	0.03%
Ovarian	–	0.00%	9,915	0.00%
Other	1,353	85.94%	2,899	0.09%
All	1,574	100.00%	3,089	0.05%
Summary Statistics:				
Wtd Market Share		0.1%		
Weighted HHI		3,324		
Mkt Share Wtd HHI		155		

markets and thus is entitled to market power in those sub-markets, it has managed to acquire market power in other therapeutic instances where it had not received such exclusivity. This is reflected by BMS' large market share in those sub-markets and the high concentrations determined in the HHI analysis. Indeed, as I showed in the market share and HHI analyses, the breast, ovarian, and lung cancer markets are all uniformly concentrated with few firms supplying the majority of drugs to this market segment.

The pharmaceutical system was designed to balance specific exclusivity incentives with the benefits that competition brings to consumers and society. However, as a result of the concentration in the paclitaxel market and relevant sub-markets, competition has suffered. Prices for Taxol are higher than they

would have been had generic or other branded paclitaxel drugs been allowed to enter the market. Taxol's high price for both its indicated and off-label uses unquestionably denies some consumers access to the drug and its potential life-saving characteristics.

Chapter 8

THE ALLOCATION OF POLICE SERVICES IN RURAL ALASKA

8.1. Introduction

In this chapter I examine whether the State of Alaska allocates scarce police resources in a race-based manner. My analysis explores this hypothesis by developing an econometric model to explain the provisions of police services in rural Alaska.

There are two basic rules that govern the way in which public resources are allocated. These two rules are equality in inputs and equality in outputs. In other words, resources can be allocated so that each person or place receives an equal number of inputs. For example, imagine a state highway department responsible for plowing snow-covered roads throughout the state. Under equal inputs allocation rules, each road would receive the same number of snowplows, whether they were needed or not. This would lead to an inefficient result, with some roads becoming impassable and others being plowed when they were already clear of snow. Under equal outputs allocation, the snowplows would be allocated based on where the snowfall was heaviest. This is a more efficient result than could be achieved by allocating the snowplows based on equal inputs. It is only by coincidence that equality in inputs will equate to equality in outputs. Generally, it is preferable to use the equal output rule when allocating scarce public services, such as law enforcement. Allocated in this manner, all residents or areas would receive an efficient allocation of scarce resources.

In this chapter, I examine the provisions of police services based on an output measure. The output measure I select is the number of hours Alaska State Troopers spend in 344 places in their territory or general area of responsibility in the State of Alaska. By examining Trooper activity I focus on decisions made at the State level for the numerous villages in Alaska and thus exclude

from analysis the provision of police services in the few large municipalities of Alaska. These places include 216 Alaskan native villages and 128 non-Alaskan native villages. An Alaskan native village is a Census designation for a place that is predominately Alaskan native. The heart of my statistical analysis is comparing these two groups of places.

To make a comparison of Alaskan native and non-native villages I examined several factors that might explain why Troopers expended the hours that they did. To accomplish this, I first collected socio-demographic data for each place, including the percentage of the population that was male, the percentage of the population below the poverty level, the population's racial percentages, per capita income level, the age distribution of the inhabitants, and population density. I also collected geographic data including the distance from each Trooper post to each place and whether the place had year-round or summer only road access. Lastly, I collected information on whether the place examined had some form of self-provision of law enforcement.

My analysis and models reveal that no one factor can completely explain the variance in Trooper hours spent in any particular place. However, several factors are statistically significant. In particular, income differentials, age differentials, distance from a Trooper post, bans on alcohol, and whether a community has its own law enforcement combine to explain a significant percentage of the variance. For example, a place with a lower per capita income generally receives more Trooper hours. Similarly, places located nearer to Trooper posts receive more Trooper hours. Places that have their own law enforcement receive fewer Trooper hours.

However, my analysis found that race was not a statistically significant factor in determining the number of hours a Trooper would spend in a particular place. To the contrary, the factors I examined indicate that Troopers are allocated rationally, in an efficient manner that attempts to balance many factors, especially scarce resource and wide disparities in access and distance to outlying places. However, race is not one of those factors.

In Section 8.2, I review the history that has led to the current scheme of law enforcement in Alaska. In Section 8.3, I discuss allocation theory and how it is applicable to free government services such as police services. I also discuss the empirical literature and review alternative econometric models of public service allocations. In Section 8.4, I discuss my econometric analysis and discuss the factors that I find are statistically significant in explaining the manner in which police services are allocated in Alaska. In Section 8.5, I present my conclusions.

8.2. Historical Background

When Alaska was granted statehood in 1959, the Department of Public Safety endeavored to set up a state agency to take over functions previously

performed by the federal government. These functions included state police protection for the new citizens of Alaska.

The geographic area that must be covered by the Alaska State Troopers (AST or Troopers) in Alaska is enormous. Many factors are taken into consideration in deploying relatively scarce law enforcement resources throughout Alaska's huge rural area. Among these factors are geographic issues, funding issues, availability of other law enforcement, transportation and communication availability, and the ability to be flexible and mobile enough to service other towns and areas as needed. All places¹ and outlying areas, that are not served by their own municipal police departments are served by the Troopers.

After Alaska was granted statehood and the State took over policing activity from the federal government, Troopers traveled to rural villages and cities. However, the relatively small number of Troopers coupled with the large area to be patrolled made it difficult to provide a local police presence.

There are currently 237 Troopers assigned to 34 posts across Alaska, each located within one of five patrol detachments. There are an additional 91 Troopers in the division of Fish and Wildlife Protection, also in 34 posts across Alaska, some of which are different from the 34 posts to which the 237 regular Troopers are assigned. All Troopers have the same basic level of training, although the Troopers assigned to Fish and Wildlife Protection have additional training in fish and wildlife matters.

As I noted above, the Troopers cover all areas in Alaska that are not covered by municipal police departments. Troopers travel from their assigned posts and visit the off-road villages they serve, both Native and Non-Native, that are located within their post. As the name implies, on-road cities are those cities that are directly linked and accessible by road. Off-road places are typically smaller and more rural than their on-road counterparts. As their designation implies, these off-road places are not directly linked to the Alaskan highway system and are not directly accessible by road. The lack of roads makes it impossible for the Troopers to regularly patrol the off-road rural villages and cities in the same manner that on-road more urban areas can be patrolled. The great distances and unpredictable and harsh weather conditions present additional problems with patrolling that no other state in the Union faces. Troopers travel to meet with the community leaders and assist the Village Public Safety Officers (VPSOs) and the Village Police Officers (VPOs). They also travel to these off-road cities and villages to handle emergencies or when they are

¹I define a "place" using the basic Census designations. Places include Census Designated Places (CDPs), Alaska Native Village Statistical Areas (ANVSA), cities and villages. These definitions are not meant to be mutually exclusive. Hence, a place may be both a city and an ANVSA. Places are generally larger than the smallest units tracked by the U.S. Census (i.e., blocks or block groups) and correspond roughly to the notion of towns or cities.

called. The Troopers' performance is evaluated, in part, on how often they make contact with the places in their region.

The VPSO program, where Troopers work closely with and assist VPSOs, had its genesis in the 1960s. Between the period 1967-1977, over 900 Native village residents have received police and correctional officer training. Currently, there is funding for 84 VPSO positions across the state. This funding is provided by the Alaska Legislature. The VPSO officers are selected and hired by the regional non-profit Native corporations. The VPSO officers are trained by the Troopers.²

The Troopers work closely with the VPSOs and VPOs to provide police protection to these far-flung off-road villages and towns. Because the Troopers cannot be in more than one place at one time, and because these villages and towns are difficult to reach, the Troopers must depend upon and work hand-in-hand with the VPSOs and VPOs to provide police protection to these off-road towns and villages, both Native and Non-Native.

8.3. Theory

Economists and political scientists have investigated how free government services, such as police and fire protection, are allocated among areas of a city or a state. In a classic study, Shoup (1989) reviewed the research that had been conducted in this area and discussed the various rules that pertained to distributing a "free" government service among city areas. Among the rules Shoup identified is a rule he called equality in distribution. There are two versions of this rule. In the first, individuals are allocated equal service inputs. In the second, individuals are allocated equal service outputs.

Equality in distribution may require equality in either inputs or outputs. However, equality in inputs is generally not equivalent to equality in outputs. For example, equality in inputs specifies that each individual receives the same level of input. However, due to differences in environment from area to area, the outputs may differ. Alternatively, equality of outputs will generally require inequality in inputs. It is only through coincidence that equality in inputs will equate to equality in outputs.

For example, a state highway department will face various snow conditions across the state. Equality of inputs would require that each highway be plowed equally. However, because snowfall on various highways is likely to be different, equal inputs (i.e., equal plowing) would leave some roads impassable, while others would be plowed even after they are clean. Similarly, equal inputs for garbage collection might specify equal inputs per mile of street, but would

²Small municipalities can also hire and pay for VPOs. The Department of Public Safety supports this program through training and advice. In 1999, \$650,000 went to VPO training. Additionally, tribes can appoint tribal police officers, but the scope of their authority is uncertain.

disregard the density of homes on a street. This would leave trash piling up at some homes while others that produced less trash would see trash picked up more frequently than needed.

Equality of inputs in terms of police protection would similarly ignore differences in crime rates or differences in population density and blindly provide equal police service inputs.³ Whether inputs or outputs of police services are equalized may depend on many factors. To further complicate matters, the distinction between inputs and outputs is sometimes blurred. Police service inputs include such concepts as number of police or budgets. However, for preventive services such as police and fire protection, the equal input rule is not particularly appealing. Shoup, using fire protection services as an example, reasoned that an input per resident that was just enough to prevent virtually all fire damage in an expensive neighborhood might be insufficient to prevent fire damage in a poorer neighborhood characterized by densely packed, more combustible structures. Generally, preventative services, such as police protection, are better allocated based on equality of outputs. Police service outputs include concepts such as hours of public service received per capita or realized crime levels.

The second rule identified by Shoup distinguishes various cases of inequality in the distribution of services. Shoup noted that maximizing services for society as a whole may require allocating differing input levels to different areas if the incremental value of such effort is different. Here, Shoup noted three possible reasons that services might be distributed unequally among areas: to mitigate inequality in income; to meet unequal needs or preferences; or to satisfy complex social welfare functions. Other allocation methodologies include: maximizing total output, equal work load, racial discrimination, income-class discrimination, power-elite favoritism, ecological influences, bureaucracy divisions, complaints-requests, taxes paid, and, daytime versus nighttime allocation.

Clearly, there are many possible rules that can be used to allocate services. Cicchetti (1971) analyzed allocation rules for recreational facilities. Cicchetti compared allocating facilities based on the rule that maximizes usage versus allocating resources to provide equal access. More recently, Behrman and Craig (1987) considered allocation rules based on a governmental welfare function. They concluded that the equity versus productivity tradeoff may be considerable. Importantly, Behrman and Craig concluded that examining inputs is problematic since residents are often concerned with the actual service level outputs provided by government. Behrman and Craig noted that inputs pro-

³Application of Shoup's analysis to police protection service is discussed in Shoup (1964).

vided by government generally do not provide clear insights regarding distributed outcomes.

In order to characterize how efficiently services are provided, economists need a normative rule for allocating services. In other words, economics is silent on the social welfare of the individuals involved. Economics is similarly silent on the production possibility frontier⁴ that determines the level of potential services provided for a fixed available budget or resource. Economists are potentially able to identify a Pareto improving situation where a single individual's welfare is increased without diminished service or utility to anyone else. All Pareto improving movements are economically efficient and increase social welfare. They may, however, be inconsistent with some specific allocation rules (such as equality of inputs or outputs).

Several empirical studies in the economics literature have examined how police protection resources are allocated. For example, Beaton (1974) used New Jersey cities of differing sizes between 1960 and 1970 to analyze police service expenditures per capita. His explanatory factors included changes in population, population density, crime rates, industrial composition, age distribution, and tax base measures. In a related study, Weicher (1971) also analyzed police protection per capita in Chicago (measured by number of policemen). His explanatory variables included income measures, retail sales per capita, and the population in Chicago police districts.⁵

The literature also contains several studies that discuss allocation rules. Based on the Behrman and Craig critique, I expect the allocation of outputs, or realized effort, to be more probative of potential race discrimination in the provision of police services. In part, this is because inputs, measured either as the number of Troopers or by annual budgets, are allocated at the post level rather than at the village level. Hence, the input allocations are both physically and temporally far removed from the villages that comprise larger percentages of native Alaskans. Further, most theoretical and empirical literature has consistently considered inputs and outputs measured on a per capita basis. Similarly, this is the approach I adopted.

8.4. Allocation Analysis

As Shoup (1989) discussed, many allocative rules are possible. There are two key questions here. First, what rule is operating in the state of Alaska to

⁴A production possibility frontier is defined as the set of output combinations that can be secured from a fixed set of inputs. For instance, with fixed annual budgets Troopers may be able to allocate their time across the places in Alaska in a variety of ways varying from very little in any given place to more nearly equal outcomes. The set of possible outcomes for Trooper hours is defined by a production possibility frontier. For further discussion see Shoup (1964).

⁵Earlier survey and statistical studies on the allocation of police service were summarized in Shoup, *et al.* (1989, pp. 114–116).

govern the Trooper allocation? Second, what factors determine this allocation? In other words, do matters of equity or efficiency explain the manner in which police services are allocated or does racial discrimination explain the actual allocation pattern?

To answer these questions, I constructed a historical dataset that is based on measuring output. I define output as the number of hours actually spent by Troopers in each of the 344 places (including CDPs⁶ and ANVSAs⁷) in the State of Alaska for which the state troopers have acknowledged responsibility. I selected these places based on those considered in (Table 4.4—Population Estimates and Alphabetical Listing of Alaska Communities, 1998–1990) of the Alaska Department of Labor’s publication “Alaska Population Overview: 1998 Estimates,” and augmented this information with places from the 1990 U.S. Census. I excluded from my analysis those places that appeared in “Alaska Population Overview: 1998 Estimates” but did not appear in the 1990 Census summary tape file. Table 8.1 provides for a complete list of the places in my analysis.

Table 8.1. Places Used in Statistical Analysis

Place	ANVSA	Own Law Enforcement	Road Access (Year Round)
Adak Station CDP	0	0	0
Akhiok city	1	0	0
Akiachak city	1	0	0
Akiak city	1	0	0
Akutan city	1	0	0
Alakanuk city	1	0	0
Alatna	1	0	0
Alcan CDP	0	0	1
Aleknagik city	1	0	0
Alexander	1	0	0
Allakaket city	1	0	0

⁶Census Designated Places (CDPs) are delineated for the decennial census as the statistical counterparts of incorporated places. CDPs comprise densely settled concentrations of population that are identifiable by name, but are not legally incorporated places. Their boundaries, which usually coincide with visible features or the boundary of an adjacent incorporated place, have no legal status, nor do these places have officials elected to serve traditional municipal functions. CDP boundaries may change with changes in the settlement patterns; a CDP with the same name in previous censuses does not necessarily have the same boundaries. To qualify as a CDP in the State of Alaska for the 1990 census, an unincorporated community must have 25 or more persons if outside an Urbanized Area, and 2,500 or more persons if inside an Urbanized Area delineated for the 1980 Census or a subsequent special census. Census of Population and Housing (1990).

⁷Alaska Native village statistical areas (ANVSAs) constitute tribes, bands, clans, groups, villages, communities, or associations in Alaska that are recognized pursuant to the Alaska Native Claims Settlement Act of 1972, Public Law 92-203. For the 1990 census, the Census Bureau cooperated with officials of each participating Alaska Native Regional Corporation (ANRC) to delineate boundaries that encompass the settled area associated with each ANVSA.

Table 8.1. Places Used in Statistical Analysis

Place	ANVSA	Own Law Enforcement	Road Access (Year Round)
Ambler city	1	0	0
Amchitka CDP	0	0	0
Anaktuvuk Pass city	1	0	0
Anchor Point CDP	0	0	1
Anchorage city	0	1	1
Anderson city	0	0	1
Angoon city	1	1	0
Aniak city	1	0	0
Annette CDP	0	0	0
Anvik city	1	0	0
Arctic Village CDP	1	0	0
Atka city	1	0	0
Atmautluak city	1	0	0
Atkasuk city (Atkasook)	1	0	0
Barrow city	1	0	0
Beaver CDP	1	0	0
Belkofski	1	0	0
Bethel city	1	1	0
Bettles city	0	0	0
Big Delta CDP	0	0	1
Big Lake CDP	0	0	1
Bill Moore's	1	0	0
Birch Creek CDP	1	0	0
Brevig Mission city	1	0	0
Buckland city	1	0	0
Butte CDP	0	0	1
Cantwell CDP	1	0	1
Canyon Villiage	1	0	0
Central CDP	0	0	0
Chalkyitsik CDP	1	0	0
Chase CDP	0	0	0
Chefornak city	0	0	0
Chenega CDP	1	0	0
Chevak city	1	0	0
Chickaloon CDP	0	0	0
Chignik city	1	0	0
Chignik Lagoon CDP	1	0	0
Chignik Lake CDP	1	0	0
Chilkoot	1	0	1
Chiniak CDP	0	0	0
Chistochina CDP	1	0	1
Chitina CDP	1	0	1
Chuathbaluk city	1	0	0
Chulloonawick	1	0	0
Circle CDP	1	0	0

Table 8.1. Places Used in Statistical Analysis

Place	ANVSA	Own Law Enforcement	Road Access (Year Round)
Circle Hot Springs CDP	0	0	0
Clam Gulch CDP	0	0	1
Clarks Point city	1	0	0
Coffman Cove city	0	0	1
Cohoe CDP	0	0	0
Cold Bay city	0	0	0
College CDP	0	0	1
Cooper Landing CDP	0	0	1
Copper Center CDP	1	0	1
Copperville CDP	1	0	1
Cordova city	1	1	0
Council	1	0	0
Covenant Life CDP	0	0	1
Craig city	1	1	1
Crooked Creek CDP	1	0	0
Crown Point CDP	0	0	1
Cube Cove CDP	0	0	0
Deadhorse CDP	0	0	1
Deering city	1	0	0
Delta Junction city	0	0	1
Dillingham city	1	1	0
Diomedea city (Inalik)	1	0	0
Dora Bay CDP	0	0	1
Dot Lake	1	0	1
Dot Lake CDP	0	0	1
Dry Creek CDP	0	0	1
Eagle city	0	0	0
Eagle Village CDP	1	0	0
Edna Bay CDP	0	0	0
Eek city	1	0	0
Egegik city	1	0	0
Eielson AFB CDP	0	0	1
Eklutna	1	0	1
Ekuk	1	0	0
Ekwok city	1	0	0
Elfin Cove CDP	0	0	0
Elim city	1	0	0
Emmonak city	1	0	0
English Bay CDP	1	0	0
Ester CDP	0	0	1
Evansville	1	0	0
Evansville CDP	0	0	0
Eyak CDP	1	0	0
Fairbanks city	0	1	1
False Pass city	1	0	0

Table 8.1. Places Used in Statistical Analysis

Place	ANVSA	Own Law Enforcement	Road Access (Year Round)
Ferry CDP	0	0	1
Fort Greely CDP	0	0	1
Fort Yukon city	1	1	0
Fox CDP	0	0	1
Fox River CDP	0	0	1
Freshwater Bay CDP	0	0	0
Fritz Creek CDP	0	0	1
Gakona	1	0	1
Galena city	1	1	0
Gambell city	1	1	0
Game Creek CDP	0	0	0
Georgetown	1	0	0
Glennallen CDP	0	0	1
Golovin city	1	0	0
Goodnews Bay city	1	0	0
Grayling city	1	0	0
Grouse Creek Group	1	0	0
Gulkana CDP	1	0	1
Gustavus CDP	0	0	0
Haines city	1	1	1
Halibut Cove CDP	0	0	0
Hamilton	1	0	0
Happy Valley CDP	0	0	1
Harding Lake CDP	0	0	1
Healy CDP	0	0	1
Healy Lake CDP	1	0	0
Hobart Bay CDP	0	0	0
Hollis CDP	0	0	1
Holy Cross city	1	0	0
Homer city	0	1	1
Hoonah city	1	1	0
Hooper Bay city	1	0	0
Hope CDP	0	0	1
Houston city	0	0	1
Hughes city	1	0	0
Huslia city	1	0	0
Hydaburg city	1	0	1
Hyder CDP	0	0	1
Igiugig CDP	1	0	0
Iliamna CDP	1	0	0
Ivanof Bay CDP	1	0	0
Jakolof Bay CDP	0	0	0
Juneau city	0	1	1
Kachemak city	0	0	1
Kake city	1	1	0

Table 8.1. Places Used in Statistical Analysis

Place	ANVSA	Own Law Enforcement	Road Access (Year Round)
Kaktovik city	1	0	0
Kalifornsky CDP	0	0	1
Kaltag city	1	0	0
Karluk CDP	1	0	0
Kasaan city	1	0	1
Kasigluck city	1	0	0
Kasilof CDP	0	0	1
Kenai city	0	1	1
Kenny Lake CDP	0	0	1
Ketchikan city	0	1	1
Kiana city	1	1	0
King Cove city	1	1	0
King Salmon CDP	1	0	0
Kipnuk CDP	1	0	0
Kivalina city	1	0	0
Klawock city	1	1	1
Klukwan CDP (Chilkat)	1	0	1
Knik CDP	1	0	1
Kobuk city	1	0	0
Kodiak city	0	0	1
Kodiak Station CDP	0	1	1
Kokahonak CDP	1	0	0
Koliganek CDP	1	0	0
Kongiganak CDP	1	0	0
Kotlik city	1	0	0
Kotzebue city	1	1	0
Koyuk city	1	0	0
Koyukuk city	1	0	0
Kupreanof city	0	0	0
Kwethluck city	1	0	0
Kwigillingok CDP	1	0	0
Labouchere Bay CDP	0	0	1
Lake Minchumina CDP	1	0	0
Larsen Bay city	1	0	0
Lazy Mountain CDP	0	0	1
Levelock CDP	1	0	0
Lignite CDP	0	0	1
Lime Village CDP	1	0	0
Long Island CDP	0	0	0
Lower Kalskag city	1	0	0
Lutak CDP	0	0	0
Manley Hot Springs CDP	1	0	0
Manokotak city	1	0	0
Marshall city	1	0	0
Mary's Igloo	1	0	0

Table 8.1. Places Used in Statistical Analysis

Place	ANVSA	Own Law Enforcement	Road Access (Year Round)
McCarthy CDP	0	0	0
McGrath city	1	0	0
McKinley Park CDP	0	0	1
Meadow Lakes CDP	0	0	1
Medfra	1	0	0
Mekoryuk city	1	0	0
Mendeltna CDP	0	0	1
Mentasta Lake CDP	1	0	1
Metlakatla CDP	0	1	0
Meyers Chuck CDP	0	0	0
Minto CDP	1	0	0
Moose Creek CDP	0	0	1
Moose Pass CDP	0	0	1
Mosquito Lake CDP	0	0	1
Mountain Villiage city	1	0	0
Naknek CDP	1	0	0
Napaimute	1	0	0
Napakiak city	1	0	0
Napaskiak city	1	0	0
Naukati Bay CDP	0	0	0
Nelson Lagoon CDP	1	0	0
Nenana city	1	1	1
New Stuyahok city	1	0	0
Newhalen city	1	0	0
Newtok city	1	0	0
Nightmute city	1	0	0
Nikiski CDP	0	0	1
Nikolaevsk CDP	0	0	1
Nikolai city	1	0	0
Nikolski CDP	1	0	0
Ninilchik CDP	0	0	0
Noatak CDP	1	0	0
Nome city	0	1	0
Nondalton city	1	0	0
Noorvik city	1	0	0
North Pole city	0	1	1
Northway CDP	0	0	1
Northway Junction CDP	0	0	1
Northway Village CDP	1	0	1
Nuiqsut city	1	0	0
Nulato city	1	0	0
Nunapitchuk city	1	0	0
Ohogamiut	1	0	0
Old Harbor city	1	0	0
Oscarville CDP	1	0	0

Table 8.1. Places Used in Statistical Analysis

Place	ANVSA	Own Law Enforcement	Road Access (Year Round)
Ouzinkie city	1	0	0
Paimiut	1	0	0
Palmer city	0	1	1
Paxson CDP	0	0	1
Pedro Bay CDP	1	0	0
Pelican city	1	0	0
Perryville CDP	1	0	0
Petersburg city	0	1	0
Pilot Point city	1	0	0
Pilot Station city	1	0	0
Pitkas Point CDP	1	0	0
Platinum city	1	0	0
Pleasant Valley CDP	0	0	1
Point Baker CDP	0	0	0
Point Hope city	1	0	0
Point Lay CDP	1	0	0
Polk Inlet CDP	0	0	0
Port Alexander city	0	0	0
Port Alice CDP	0	0	0
Port Alsworth CDP	0	0	0
Port Clarence CDP	0	0	0
Port Graham CDP	1	0	0
Port Heiden city	1	0	0
Port Lions city	1	0	0
Port Protection CDP	0	0	0
Portage Creek	1	0	0
Primrose CDP	0	0	1
Prudhoe Bay CDP	0	0	1
Quinhagak city	1	1	0
Rampart CDP	1	0	0
Red Devil CDP	1	0	0
Ridgeway CDP	0	0	0
Rowan Bay CDP	0	0	0
Ruby city	1	0	0
Russian Mission city	1	0	0
Salamatof CDP	0	0	0
Salcha CDP	0	0	1
Sand Point city	1	1	0
Savoonga city	1	0	0
Saxman city	1	0	0
Scammon Bay city	1	0	0
Selawik city	1	0	0
Seldovia city	0	1	0
Seward city	0	1	1
Shageluk city	1	0	0

Table 8.1. Places Used in Statistical Analysis

Place	ANVSA	Own Law Enforcement	Road Access (Year Round)
Shaktoolik city	1	0	0
Sheldon Point city	1	0	0
Shishmaref city	1	0	0
Shungnak city	1	0	0
Sitka city	0	1	0
Skagway city	0	1	1
Skwentna CDP	0	1	0
Slana CDP	1	0	0
Sleetmute CDP	1	0	0
Soldotna city	0	1	1
Solomon	1	0	0
South Naknek CDP	1	0	0
St. George city	1	0	0
St. John Harbor CDP	0	0	0
St. Mary's city	1	0	0
St. Michael city	1	0	0
St. Paul city	1	0	0
Stebbins city	1	0	0
Sterling CDP	0	0	1
Stevens Village CDP	1	0	0
Stony River CDP	1	0	0
Sutton CDP	0	0	1
Takotna CDP	1	0	0
Talkeetna CDP	0	0	1
Tanacross CDP	1	0	0
Tanana city	1	1	0
Tatitlek CDP	1	0	0
Tazlina	1	0	0
Telida	1	0	0
Teller city	1	0	0
Tenakee Springs city	1	0	0
Tetlin CDP	1	0	0
Thorne Bay city	0	0	1
Togiak city	1	0	0
Tok CDP	1	0	1
Toksook Bay city	1	0	0
Tonsina CDP	0	0	1
Trapper Creek CDP	0	0	1
Tuluksak city	1	0	0
Tuntutuliak CDP	1	0	0
Tununak city	1	0	0
Twin Hills CDP	1	0	0
Two Rivers CDP	0	0	1
Tyonek CDP	1	0	0
Ugashik	1	0	0

Table 8.1. Places Used in Statistical Analysis

Place	ANVSA	Own Law Enforcement	Road Access (Year Round)
Ukivok	1	0	0
Unalakleet city	1	0	0
Unalaska city	1	1	0
Upper Kalskag city	1	0	0
Valdez city	0	1	1
Venetie CDP	1	0	0
Wainwright city	1	0	0
Wales city	1	0	0
Wasilla city	0	1	1
Whale Pass CDP	0	0	0
White Mountain city	1	0	0
Whitstone Log Camp CDP	0	0	0
Whittier city	0	1	1
Willow CDP	0	0	1
Wiseman	1	0	1
Womens Bay CDP	0	0	1
Wrangell city	0	1	0
Yakutat city/Borough	1	0	0
TOTAL	216	41	101

Based on the discussion above, I consider the output allocations (i.e., hours spent by the Troopers in specific places) rather than the input allocations (i.e., budgets). The “hours” variable in my analysis is a combination of Alaska State Trooper and Fish and Wildlife Protection Services time. The State of Alaska provided an account of time for Troopers and Fish and Wildlife Protection Services (FWPS) from 1994-1998 as reported in their Officer Activity Reporting System (OARS).⁸ The OARS system collects data on a per case basis. Each case is attached to a “beat,” OARS’ smallest geographic unit. The State of Alaska provided Trooper and FWPS hours at the beat level, and matched their beats to my list of places. There were 17 beats that did not map exactly to one particular place. In those situations, I allocated Trooper and FWPS time by comparing the population of the places and weighting them proportionately.

⁸The measure I used in my statistical analysis includes some time provided by the Fish and Wildlife Services. The preponderance of the time (99.4 percent) is, however, attributable to the Troopers. Nearly identical statistical results were obtained using a measure of hours that eliminated Fish and Wildlife activity.

8.4.1. The Dataset

The dataset I compiled consists of five years of observations for Trooper hours for 344 places in Alaska. The Census designates 216 of the 344 places as ANVSAs. The remaining 128 places are non-ANVSAs.⁹

To help explain the hours in different locations, I collected various socio-demographic, geographic, and place specific information. From U.S. Census sources for each place, I collected factors for: the percent of the male population at the place; the percent of the population below the poverty level; the age distribution of inhabitants; the percentage of the population that is American Indian, Eskimo or Aleut; the percent of the population that is black; the percent of the population that is white; and the per capita income level.

For the socio-demographic factors, I used the 1990 U.S. Census' measures of population by place, age, gender, race, 1989 per capita income, and poverty status. While the Census updates the population figures for non-decennial years (decennial years occur each decade—1980, 1990, 2000), I used the population figures from 1990 to have a consistent basis to derive the explanatory variables.

The population variable in my analysis is equivalent to the 1990 U.S. Census report of total persons by place. Since the Census reports population numbers by age in over 30 categories, I combined several categories so the breakdown of ages matched the 12 age categories in the Census poverty statistics. These groupings indicate the number of individuals less than five years old, five years old, six to eleven years old, and so forth. Using the midpoints of the age distributions, I formed a weighted average for each place. Similarly, the Census provides the population numbers by males, blacks, Natives (Aleuts, Eskimos, and Native Americans), Asians (Asians and Pacific Islanders), and "others." I used these population figures from the 1990 U.S. Census to calculate the percentage male, the percentage white, and so forth.

⁹The multivariate regression analysis reported below relies on observations from 308 of the 344 places included in the dataset. In the 36 places that are omitted from the regression analysis, 29 have zero Trooper hours, 14 have zero population between 1994 and 1998, and 34 have either zero Trooper hours or zero population. Two additional observations are omitted from my analysis because the population in 1990 was missing and consequently the race percentages could not be calculated. Places with zero per capita income or zero population cannot be used in the logarithm transformed dependent variable and are therefore eliminated in the regression analysis. Observations in which the race percentages could not be calculated were similarly unusable in the regression analysis. The 36 places not included in the regression analysis represent less than 0.4 percent of total Trooper hours (1,671/431,195) represented in the full dataset of 344 places. These omitted areas similarly represent less than 1.3 percent of the population (33,932/2,638,260) between 1994 and 1998 represented in the full dataset.

The analysis of variance results reported below are calculated after eliminating the three largest Alaska cities; they are not comparable to the other included places without a treatment effect. The appropriate treatment effect is included in the regression analysis which consequently is based on three additional observations.

I also define a percentage of the population as “young” people and a percentage of the population as “old” people using 11 and 65 years as the respective cutoff ages for these two groups.¹⁰ The 1990 U.S. Census provided the number of individuals below the poverty level by age group (the average poverty threshold for a family of four persons was \$12,674 in 1989¹¹). These were combined to determine an aggregated measure of percent below the poverty level at each place. Finally, the per capita income variable in my analysis is the Census report of 1989 per capita income and is measured in dollars (Census of Population and Housing, 1990).

I next examined several geographic variables. First, I measured the latitude and longitude of the place and compared this to the latitude and longitude of the Trooper post assigned to that location. I then calculated the actual distance between the post and place based on line of sight distance. To calculate the distance between a place and Trooper post, I relied on the Troopers’ designation of the latitude and longitude of their posts, and the 1990 U.S. Census specification of the latitude and longitude of each places’ “internal point.”¹² I also calculated the line of sight distance to the nearest post (as compared to the assigned post).¹³

I also considered factors for the places’ year round road accessibility. The State of Alaska supplied me with an indication of those places that have road access. They specified road access as having either year round access, or summer months access only. About one hundred places had year round road access, while six were limited to summer time only road access. Table 8.1 identifies those places with year round road access.

I also collected data on the places over time to ascertain what kind of law enforcement program, if any, was in effect. I was able to calculate the number of Full-Time Equivalents (FTE) for local police provision as follows. I received an indication of placement and dates of service for VPSOs from a University of Alaska study. Thus, I was able to determine the service level each place received from VPSOs as a fraction of a year from 1994 through 1998. To

¹⁰The categorization of age into “young” and “old” is purely to illustrate the age profile of the places in the analysis of variance. In the multivariate regression analysis, I use the continuous measure of age described above. Therefore, the regression analysis does not rely on the determinations of “young” and “old.”

¹¹Census of Population and Housing, 1990

¹²As defined by the Census, “an internal point is a set of geographic coordinates (latitude and longitude) that is located within a specified geographic entity. A single internal point by the Census is identified for each entity. For many entities, this point represents the approximate geographic center of that entity. If the shape of the entity causes this point to be located outside the boundaries of the entity, it is relocated so that it is within the entity. If the internal point for a block falls in a water area, it is relocated to a land area within the block.” Census of Population and Housing, 1990).

¹³I have not been able to calculate the actual cost of a trip from a place to an assigned post. To the extent that ease of access is constant and related to mileage, distance measured in this way serves as a proxy for the cost of visiting a place. In any case, I expect the two measures to be highly correlated.

calculate this fraction, I considered the start date and end date of employment at the place and measured the number of full days the VPSO was employed.¹⁴

The State of Alaska provided a list of places considered by the Troopers to have their own law enforcement (typically a municipal police department). AST and FWPS provide services to areas just outside the boundary of places that have their own law enforcement, and report their hours to the nearest beat. Thus, it is possible for a place to have its own law enforcement and to have non-zero AST and FWPS hours. Of the 344 places I considered, 41 have their own law enforcement. Table 8.1 identifies those places the Troopers consider as having their own law enforcement.

To distinguish the most populated places in Alaska from all others, I defined an indicator variable for Anchorage, Juneau, and Fairbanks. These three cities account for 60 percent of the population in the 344 places (measured in 1990). On the other hand, these cities account for roughly 21 percent of the total Trooper hours spent in Alaska places between 1994 and 1998 (145,791/691,453). I expect that the large cities will receive a disproportionately smaller number of per capita hours compared to other places.

In Alaska, local places can adopt an option to ban the sale, importation, or possession of alcohol. Many places have exercised this option and have banned alcohol at least over some spans of time. Crime and social disruption can be affected dramatically when the possession of alcohol is banned.¹⁵ The alcohol variable in my analysis is derived from the Schedule of Local Option Communities report produced by the Alaska Alcoholic Beverage Control Board (Griffin, 2000). I define the alcohol variable to reflect whether a place has banned the sale, importation, or possession of alcohol for any part of each year in my analysis (1994-1998). If a place elected to ban either alcohol sale, importation, or possession during a year, that place would have a value of "one" for the alcohol variable. If a place did not elect to ban either alcohol sale, importation, possession during a year, that place would have a value of "zero" for the alcohol variable.

Finally, I constructed a population density measure. The Census computes population density by dividing the total population of a geographic unit by its land area, measured in square kilometers (Census of Population and Housing, 1990). Thus, I calculated the population density of each place by dividing the total population as reported in the 1990 U.S. Census by its area in square kilometers of land as also reported in the U.S. Census.

¹⁴For example, if a VPSO was present at a place from January 1 through December 31 for a given year, the FTE variable would equal one. Likewise, if two VPSOs were each employed for exactly six months of a given year, the FTE variable would also equal one.

¹⁵For discussion, see e.g., (Landen, *et al.*, 1997).

For the preponderance of the information I collected, separate measurements were not available from year to year. For example, “per capita income” was measured by the U.S. Census in 1989. Also, factors such as the “percent male” or “percent white” are not known at the place level from year-to-year. Since the majority of the explanatory factors do not reflect temporal variation, it is appropriate to time average the data and base the analysis on the average of the factors measured over time.

8.4.2. Analysis of Variance

My first investigation is the degree to which the factors I have identified are different for the places with ANVSA census designation and without ANVSA designation.

First I find that the Troopers’ hours are statistically significantly lower in ANVSAs ($p = 0$). Average annual Trooper hours for five years in ANVSA places are 185 hours and 682 hours in non-ANVSA places. However, the size of the various places as measured by population is generally different as well. The average population in ANVSAs is 345 people. The average population in non-ANVSAs is 1,104 people, about three times greater. The difference is statistically significant ($p = 0$).

As I discussed above, it is appropriate to adjust the hours spent by the Troopers in a place by the place’s population. I created a variable for per capita hours based on the Trooper hours measured in each of the years between 1994 and 1998 and the annual population measured in the corresponding year. Per capita hours are therefore defined annually in each Alaska place. To form the time-average of Trooper hours by place, I weight by the annual population figures in each place and year. Consequently, the hours per capita variable analyzed below is equivalent to the total hours spent in a place by Troopers over the full five years divided by the total population in that place for the same five-year period. The weighted average of per capita hours (measured across places using population weights for each place) indicates that the average number of hours for ANVSA places is 0.468 hours per capita and is 0.572 hours per capita in the non-ANVSA specified places. On a weighted average basis, the ANVSA areas fair only slightly worse than the non-ANVSA areas. The weighted average of hours per capita is arithmetically equivalent to the total Trooper hours for the five-year period (across ANVSA and non-ANVSA places separately) divided by the total population served in these same places during the five-year period.

Using per capita hours, I find that there is no longer a statistical difference in per capita hours for ANVSA and non-ANVSA places. Per capita hours are not statistically different between ANVSA and non-ANVSA places (0.88 hours per capita in ANVSA places versus 1.28 hours per capita in non-ANVSA places; $p = 0.11$). However, population is not the only factor that potentially explains the difference in Trooper hours expended.

The ANVSA places are, on average, equally male as compared with the non-ANVSA places (54.8 percent versus 55.1 percent, $p = 0.79$). ANVSA places are much more likely to be below the poverty line (27.0 percent versus 11.6 percent; $p = 0$) than non-ANVSA places. The ANVSA places have a statistically higher percentage of young people (27.1 percent versus 21.4 percent; $p = 0$) than do the non-ANVSA places. A similar difference does not exist for older people, where ANVSA places had a non-statistically lower significant percentage of older people (9.7 percent versus 8.4 percent; $p = 0.28$) than did non-ANVSA places. As expected, the ANVSA places are more highly comprised of Native Americans (76.5 percent versus 9.0 percent; $p = 0$). The places are somewhat different in black composition (0.20 percent in the ANVSA places versus 1.01 percent in non-ANVSA places; $p = 0.0$). The places are not statistically different in Asian composition ($p = 0.82$). The places are very different in white composition (22.0 percent in the ANVSA places versus 88.2 percent in the non-ANVSA areas; $p = 0$).

The difference in per capita income is quite striking. ANVSA places have an average income per capita of \$10,015. Conversely, non-ANVSA places have a per capita income of \$16,865 ($p = 0$).¹⁶ In sum, there are significant socio-demographic differences between the ANVSA and non-ANVSA places. These factors are further evaluated in my multivariate analysis.

The geographical differences between ANVSA places and non-ANVSA places are also readily apparent. The average distance between ANVSA places and their assigned Trooper posts is 118.2 miles. For non-ANVSA places, the average distance is 69.3 miles ($p = 0.01$). The distance to the closest post is 62.3 miles for ANVSA places versus 28.5 miles for the non-ANVSA places ($p = 0$). The similarity of measured distance does not rule out the empirical possibility that Trooper hours are a function of distance. I discuss the distance issue in greater detail below.

The ANVSA places are somewhat less likely to provide their own law enforcement (10.0 percent versus 16.0 percent; $p = 0.13$) than the non-ANVSA places. Population density is not found to be statistically different between ANVSA and non-ANVSA areas ($p = 0.387$). Regarding road access, the ANVSA areas are only accessible year round in 11.0 percent of ANVSA but are accessible year round in 65.1 percent of the non-ANVSA places ($p = 0$).

The ANVSA areas are much more likely to have a VPSO presence. The VPSO FTE measure is on average 0.38 in ANVSA areas (less than one half FTE per annum) versus an average of 0.01 in non-ANVSA areas ($p = 0$). Therefore, ANVSA areas are more likely to be covered by VPSO programs

¹⁶The per capita income figures for both groups are somewhat low due to the 1989 measurement date. Additionally, these per capita figures use total population counts, and are likely to be low due to presence of non-earners.

and non-ANVSA areas are more likely to be covered by other law enforcement programs. Finally, alcohol prohibitions are much more likely to occur in ANVSA places than in non-ANVSA places (41.6

In summary, the ANVSA and non-ANVSA places vary in many respects. While I have found that on average in ANVSA and non-ANVSA places hours per capita provided by Troopers in these areas were similar, there are several factors that help explain place to place differentials and permit a test of racial discrimination as a potential explanatory factor. These issues are addressed in the next section.

8.4.3. Econometric Model

In Table 8.2, I display the regression model that I developed to explain the differences in the Troopers' hours spent in the various places. The regression analysis simultaneously controls for differences in many factors and generalizes the analysis of variance results I presented above. I use the regression model to test the hypothesis that, after controlling for other important factors, race plays some role in determining Trooper hours spent per capita in places. The first transformation I make is to take the logarithm of the per capita hours. This allows me to better control for heteroscedasticity in the regression modeling. However, similar results were obtained without this transformation.

I find that places in which more males are present receive more Trooper time. This is due to the fact that men are more likely than women to be involved in crimes. The percentage of the population below the poverty also helps explain the Trooper hours expended in different locations and similarly reflects a crime effect. I also find that the age variable is positively related to Trooper hours expended. Hence, areas with older populations are more likely to (and did) receive more attention from the Troopers.¹⁷

Race effects are not apparent in the multivariate regression analysis. For instance, if a Native American race effect were present, one should expect the coefficient on the variable for the percentage of Native Americans to be negative and statistically significant in the regression model. In other words, as the percentage of Native Americans in a place increases, fewer Trooper hours would be allocated to that place in the presence of a racial bias. This is not the case.

The other key factor measuring the possible influence of race on Trooper hour allocation is the ANVSA designation. This indicator does not reveal statistical significance. If a race effect or discrimination effect were present I

¹⁷In other models, I found that younger population groups (age 11 and under) received less attention while older population groups (age 65 and over) received more attention in Trooper time. Therefore, the age effect reflects the fact that elderly individuals require more Trooper assistance. Whether this extra time is crime or safety related is not discernible.

would expect the ANVSA indicator variable to show lower per capita Trooper hours (all other things equal). This was not the case at the 95 percent level of statistical confidence.

Similarly, the other race effects are jointly insignificant. As the omitted category in the model is "white," it is correct to say that blacks, Asians, and Native Americans are treated identically to whites on a statistical basis. Therefore, the multivariate regression analysis fails to demonstrate any race bias.

Areas with lower per capita income have more Trooper hours. However, this result was not statistically well determined. As the place's per capita income falls, more Trooper time is required (other things equal). This is also likely to be a crime effect, with more impoverished areas having more crime. It is worth noting here that crime statistics are not measured at the place level. In fact, crimes are clearly related to Trooper hours, ignoring the issue that some crimes go unreported. Thus, it was not possible to include crime related explanatory variables in the regression model. Further, it would not have been appropriate to do so given the close connection between the dependent variable (Trooper hours) and crime.

The distance from a place to its closest Trooper post has a negative and statistically significant effect on Trooper hours spent in the place.¹⁸ Hence, the further away a place is from any Trooper post, the less likely it will receive attention, all other things equal.¹⁹ I interpret this distance result as indicating that more distant locations will and do receive less attention. This is not because they are racially discriminated against. It is simply the reality that with a fixed budget, Troopers find it more costly in money and time to travel the greater distance to very remote areas. This allocation reflects economic efficiency. If the marginal value of an additional Trooper trip is to be maximized, it is necessary to equate the marginal benefit less marginal cost of additional Trooper trips across all places. Assuming that the marginal benefit of an additional Trooper trip is constant and equal across all places, economic efficiency requires that the marginal cost of additional Trooper trips be equalized across places. Since the marginal cost of a trip to a remote area is greater than the marginal cost of a trip to a nearby area, efficiency requires fewer trips to outlying areas. On the other hand, the marginal benefit of an additional trip will generally not be equal across places because Troopers respond to calls of varying degrees of urgency. For instance, the marginal benefit to Alaskans may be much larger when Troopers respond to homicides as compared to less severe

¹⁸The road access variables (year round and summer only) were not significant determinants of per capita Trooper hours in the multivariate regression analysis and are not displayed in Table 8.2.

¹⁹Regression models that relied on the distance to the assigned post rather than to the closest post found no statistically significant distance effect.

misdeemeanor offenses. In fact, it is likely that Troopers respond to all calls in which the urgency is immediate irrespective of the distance.²⁰

Places with their own law enforcement were statistically less likely to receive Trooper attention. Conversely, those areas without law enforcement were the most likely to receive Trooper attention measured on an annual hours per capita basis.

The presence of a VPSO program as indicated by the full-time employment VPSO variable was not statistically significant in determining the per capita Trooper hours (all other factors equal). Instead, the VPSO program serves to augment Trooper efforts. Population density was not found to be related to Trooper hours spent.

As discussed above, I find that the large cities (Anchorage, Juneau and Fairbanks) receive a disproportionately lower share of Trooper hours as compared to other places. This result, however, was not statistically significant.

As discussed above, the alcohol variable represents bans on the sale or importation of alcohol or bans on its possession. Places imposing bans of either kind on alcohol may discourage alcohol-related crimes or misdemeanors. Consequently, the presence of an alcohol ban is expected to reduce the demand for Trooper hours. This effect is demonstrated in the multivariate regression analysis. The alcohol variable is negatively and statistically significantly related to per capita Trooper hours (all other things equal). Hence, alcohol bans are a form of own law enforcement and reduce Trooper related calls to Alaskan rural places. Conversely, places without alcohol bans apparently require and receive greater Trooper attention (controlling for other factors).

Finally, whether or not a place has the ANVSA designation was not related to per capita hours received from the Troopers, controlling for other influences. In sum, neither race nor place designated status influenced the Troopers' allocation decision.

8.5. Conclusions

There are many factors that influence the number of hours spent by Troopers in places in Alaska. No one factor completely explains the picture. However, collectively, income differentials, age differentials, distance, alcohol bans, and own law enforcement explain a significant percentage of the variance of the Trooper hours spent in the places. Importantly, I conclude that there is no evidence that race has an effect on the number of Trooper hours spent in areas designated by the U.S. Census with ANVSA status). Therefore, as there are no race effects, I conclude that there is no discrimination in allocating scarce

²⁰To test this hypothesis, I respecified my regression model to explain the number of homicides handled by Troopers in Alaskan places. In these regressions, I continue to find no evidence of race bias. I also fail to find an allocation related to distance as might be expected given the severity of the crime.

police resources. Instead, my analysis reveals a rational response to resource allocation that balances many factors efficiently. Clearly, a larger budget would allow the Troopers to spend more time in remote villages. Alternatively, increased VPSO and VPO programs could provide a form of own law enforcement that could reduce the necessity for Trooper visits.

Table 8.2. Alaska Allocation Analysis

Dependent Variable:	Log Per Capita Hours
Constant	-2.44 (-4.04)
Alaska Native Village Statistical Area	0.056 (0.21)
Percent Male	2.43 (2.25)
Percent Below Poverty	0.884 (1.92)
Age	0.023 (1.67)
Percent Native	-0.191 (-0.48)
Percent Black	-9.15 (-2.14)
Percent Asian	-3.98 (-1.68)
Per Capita Income, 1989	$-4.92e - 6$ (-0.42)
Minimum Distance from Place to Post	-0.00393 (-2.91)
Own Law Enforcement	-0.479 (-1.87)
Population Density	-0.00024 (-0.30)
Village Public Safety Officer	0.359 (1.44)
Alcohol Ban	-0.443 (-2.09)
Anchorage, Fairbanks, or Juneau	-0.387 (-0.48)
Observations	308
R-squared	0.16

t-statistics in parenthesis

Chapter 9

FINANCIAL MARKET REACTION TO THE FAST FOOD HAMBURGER HEALTH SCARE OF 1993

9.1. Introduction

In 1993 a bacteria epidemic killed several children in the Pacific Northwest. The bacteria was a virulent strain of *E. coli* and was traced to undercooked hamburgers at Jack-in-the-Box restaurants. The outbreak was the largest and most serious for *E. coli* with a total of 400 confirmed or probable instances. Ultimately the tainted hamburgers forced 125 people to be admitted to a hospital and led to kidney failure in 29 people.

Jack-in-the-Box, whose corporate name was Foodmaker, did not handle the public relations crisis very well in the first few days of the crisis. It took several days before the company addressed the public and removed all meat from its restaurants. However, Jack-in-the-Box did respond with a quality assurance program to ensure the safety of food at each point in the distribution system from meat supplies to the cooking process. The process now mandates that if a supplier's products test positive for harmful *E. coli* on more than one occasion that supplier is dropped. Today, restaurant management tests the cooking system and cooks sample products. There are weekly inspections and each beef patty is checked by a company certified employee before it is removed from the grill.

Public health officials are worried about virulent strains of *E. coli* including *E. coli* 0157:M7 because they may appear anywhere, from the municipal water supply to apple cider, and in rare meat. Outbreaks in meat are particularly difficult to avoid because cattle are a major reservoir of the bacterium. *E. coli* 0157:M7 is more virulent than other forms of *E. coli* because it clings to cell walls in the human bowel and produces a toxin that causes bleeding. Doctors do not have an effective treatment plan for *E. coli* 0157:M7 and are baffled by the spectrum of symptoms. Most commonly a problem is detected in otherwise

healthy children that present with kidney disease including hemolytic-uremia syndrome or thrombocytopenic purpura.

Hamburgers pose the greatest risk because the grinding and mixing of the meat provides a large surface area for the growth of the bacteria. Steaks and ribs are less susceptible due to the heating of the exterior even in "rare" preparations.

As a result of the Jack-in-the-Box outbreak and because of increasing awareness of the health hazards of *E. coli* 0157:M7, the U.S. government adopted new inspection programs for the monitoring and testing of meat and poultry from slaughter to consumption. New standards were adopted for salmonella bacteria; all meat processes were required to test for *E. coli* 0157:M7; each plant must now identify the critical points for inspection and these must be approved by the Department of Agriculture on a can by can basis. Most importantly the Food and Drug Administration has increased the federal minimum cooking temperature for ground beef from 140 degrees to 155 degrees.¹

While Jack-in-the-Box was able to respond fairly rapidly to the public relations nightmare it was not quick enough to control the damage to its reputation. Over the two years following the outbreak, Jack-in-the-Box lost 30 percent of its stock market value and had its debt reclassified to junk bond status by Moody's investor rating services. Foodmaker was reported to have lost at least \$138 million over two years. Fortunately for Foodmaker this financial disaster was subsequently reversed as Foodmaker stood behind its franchises and used a public relations campaign to humanize its fast-food chain by resurrecting the "Jack" icon including an ad in which Jack dynamited company headquarters. At this time Foodmaker's annual sales were approximately \$129 billion with two-thirds of the revenue coming from the Jack-in-the-Box outlets. Reported sales losses were, in fact, much too low.

In this chapter I use historical sales data to estimate a demand model for hamburgers sold in the fast-food market. I use this model to compare Jack-in-the-Box actual and forecasted sales for the two-year period following the *E. coli* outbreak. My conclusion is that Jack-in-the-Box sustained much greater losses and at levels that were consistent with investor expectations that the losses would be permanent.

In section 9.2, I discuss the data sources used in this analysis. In section 9.3, I present my econometric model of hamburger demand and contrast it to two other approaches for forecasting demand, one based on sales revenues of Jack-in-the-Box directly while the second based on a market share approach. The final section presents estimates of lost sales and my conclusions.

¹McDonald's corporation has always cooked its burgers to a temperature of at least 157 degrees while Jack-in-the-Box was much closer to the then-permissible standard of 140 degrees according to federal reports.

9.2. Data Sources

The data for this study were derived from four sources: (1) Jack-in-the-Box sales for several regions from 1987 through 1994 obtained through Restaurant Trends, (2) national income product accounts by year and quarter, (3) consumer price index and unemployment data by region and quarter from the Bureau of Labor Statistics; and (4) earnings, population, and employment data by year and region from the Bureau of Economic Analysis.

9.2.1. Restaurant Trends Data

The Restaurant Trends data represent surveys of Quick Service Restaurants (QSR) in major regions of the United States. Data were obtained for ten cities: Los Angeles, Sacramento, San Diego, San Francisco, St. Louis, San Antonio, Phoenix, Dallas, Houston, and Seattle. (The sales for these cities represent approximately 80 percent of reported domestic and international sales.) The data provide revenues and number of stores on a quarterly basis for the period 1987 through 1994. The econometric analysis is based on 240 observations (ten regions, six years, and four quarters). The estimation period is based on the period from 1987 through 1992 and excludes 1993 and 1994 so that a lost-sales calculation can be made using predicted and actual sales. The QSR reports provide information on the major hamburger chains including Burger King, McDonald's, Wendy's, Arby's, Carl Junior's, Jack-in-the-Box, Hardy's, and Whataburger.

9.2.2. Bureau of Economic Analysis

Data were obtained on the ten primary cities for the years 1987–1996 from the Regional Economic Information System of the Bureau of Economic Analysis (BEA). The BEA data provide measures of population, per-capita income, earnings at retail establishments, earnings at eating and drinking establishments, total employees, and employment at retail establishments.

9.2.3. National Income and Product Accounts

Data were obtained from the National Income Products Accounts, including real gross domestic product, real personal consumption expenditure, real personal consumption expenditure on food, real personal consumption expenditure on services, national population levels, real disposable personal income, and the price index for personal consumption expenditures on food. Data were obtained on a quarterly basis from 1987 through 1996.

9.2.4. Bureau of Labor Statistics

From the Bureau of Labor Statistics, I collected CPI data for all urban consumers and unemployment statistics by region. Data were collected on

a monthly basis and converted to a quarterly basis for the years 1987 through 1996.²

9.3. Econometric Models

In order to model the historical sales at Jack-in-the-Box, I utilize a time-series cross-section structure with quarterly observations. The use of the ten primary regions, six years (1987–1992), and four quarters per year, results in 240 observations. The regression models I specify in this section consider: (1) revenues, (2) number of stores, and (3) revenues per store (average unit volumes). A variable glossary is provided in Table 9.1.

Table 9.1. Variable Glossary

r2–r10	Region dummy variables
q2–q4	Quarter dummy variables
perfran	Percentage of stores that are franchised
trend	time trend
lrpci	log of real per-capita income
lur	log of unemployment rate
lrrev	log of real sales revenue per capita
lrauv	log of real sales revenue per store
lrsto	log of stores per capita

When modeling revenues, I use real sales per capita, *i.e.*, I adjust the quarterly sales for changes in the price level and express the result on a per-capita basis. All adjustments are done using the values of the factors at specific regions and time periods. When modeling average unit volumes, the sales per store are expressed in real terms. When modeling the number of stores, I divide the number of stores by the regional population. The latter variable is the reciprocal of the number of individuals in the population per store and corrects for possible heteroscedasticity.³

²Where specific regional information was not available, the values for nearby regions were substituted.

³The time-series cross-section form of the data allowed generalized least squares estimation. The optimal weighting (estimated ratio of within and between variances) was 0.95 so that a fixed-effects approach was adopted. The models were estimated in log-linear form. The estimation period is from 1987 through 1992 and represents the pre-outbreak period.

9.3.1. Revenue per Capita

To explain the real sales per capita of Jack-in-the-Box stores in a quarter, I use nine region-specific dummy indicator variables (the fixed effects), three quarterly dummies (to control for seasonality), the percentage of franchised stores in the region, a trend term, real per-capita income, and the unemployment rate. Based upon the 240 historical observations, the R -squared is over 97 percent. A significant trend effect was determined implying a 3.5 percent annual increase in real per-capita sales for the period 1987–1993. Increases in real per-capita income increase sales, while increases in unemployment decrease sales. The percentage of franchised stores apparently does not affect the real per-capita sales level. Some seasonality is demonstrated with lower sales in January through March. Finally, the fixed effects were very precisely determined. The estimated regression model is given in Table 9.2.

9.3.2. Average Unit Volume

Table 9.3 considers real average unit volume. Average unit volume is defined as revenue divided by the number of stores. The results are similar to those described in Table 9.2, with approximately 87 percent of the variation explained.

There are some notable differences. In this model, the trend in average unit volumes is slightly negative over the period. Additionally, the percentage of stores that are franchised in an area leads to higher per-store revenue. Regional sales are otherwise higher in regions with a higher percentage of franchises.

9.3.3. Stores per Capita

The qualitative results for per-capita stores were not dissimilar to those obtained in the average unit volume regression (Table 9.4). Not surprisingly, there was no quarterly store effect.

There was a positive trend effect measured at a 4.4 percent per year growth in store placements per capita. Areas with higher real per-capita income or lower unemployment rates also revealed a greater penetration in stores. The R -squared for this model was 96 percent.

9.4. Simulations and Conclusions

The estimated models were used to forecast “but-for” sales in the post-1993 period. The purpose of the “but-for” simulation is to predict sales levels as if the E. coli outbreak had not occurred. The estimated lost sales for the 1993–1994 period inclusive is \$267 million, or approximately 16.6 percent of historical actual Jack-in-the-Box sales.

Actual sales for 1993 and 1994 were \$1.6099 billion, while predicted sales were \$1.8771 billion. Models based on average unit volume and number of

Table 9.2. Revenue per Capita

Dependent Variable: lrrev		
Independent Variable	Estimated Coefficient	t-Statistic
(1)	-18.662	-20.34
r2	0.138	6.91
r3	0.099	5.40
r4	-0.261	-3.22
r5	-0.728	-23.51
r6	-0.506	-7.99
r7	0.209	7.63
r8	-0.050	-2.01
r9	-0.265	-9.28
r10	-0.623	-17.12
q2	0.087	7.74
q3	0.114	9.30
q4	0.149	11.56
perfran	0.007	0.11
trend	0.034	9.48
lrpci	2.073	10.01
lur	-0.128	-4.20
Observations		240
Corrected R-squared		0.978
Mean of Dependent Variable		-5.417

stores provided identical results (when combined) given the logarithmic transformation of the dependent variables.

9.4.1. Market Share Logit Model

To corroborate these estimates I estimated a multinomial logit market share model. The market share model is based on the ratio of the market share of each chain's sales to that of a normalizing alternative. I use sales at McDonald's stores as the normalizing alternative. While this can be done without loss of generality (*i.e.*, predicted market shares would be identical using another normalizing alternative), the interpretation of the coefficients for the explanatory factors is affected by this choice. The dependent variables in these regressions

Table 9.3. Average Unit Volume

Dependent Variable: lrauv		
Independent Variable	Estimated Coefficient	t-Statistic
(1)	-4.765	-6.48
r2	0.109	6.76
r3	0.053	3.58
r4	-0.241	-3.70
r5	-0.069	-2.81
r6	-0.148	-2.91
r7	0.045	2.06
r8	-0.173	-8.56
r9	-0.110	-4.79
r10	-0.135	-4.64
q2	0.090	9.95
q3	0.118	12.01
q4	0.161	15.52
perfran	0.201	3.79
trend	-0.009	-3.32
lrpci	1.244	7.50
lur	-0.126	-5.15
Observations		240
Corrected R-squared		0.883
Mean of Dependent Variable		0.507

are the difference in the log average unit volume of a particular chain and log average unit volume at McDonald's. Explanatory factors include alternative specific constants, a time trend, real per-capita income, and unemployment. The market share model is estimated using four regression equations (one less than the number of alternatives). Here, I assumed that there are five major alternatives in the market share model (Jack-in-the-Box, Burger King, Wendy's, Carl Junior's, and McDonald's).

In order to simulate the market share model, I constructed a variable to measure the total revenue for the five hamburger chains used in the market share equations. To model real per-capita sales at the five hamburger chains, I used a constant term, time trend, the ratio of employment in the retail sector to total

Table 9.4. Stores per Capita

Dependent Variable: lrsto		
Independent Variable	Estimated Coefficient	<i>t</i> -Statistic
(1)	-13.897	-14.69
r2	0.029	1.44
r3	0.046	2.45
r4	-0.020	-0.25
r5	-0.658	-20.62
r6	-0.358	-5.48
r7	0.164	5.80
r8	0.122	4.70
r9	-0.155	-5.28
r10	-0.488	-13.00
q2	-0.002	-0.23
q3	-0.004	-0.32
q4	-0.011	-0.85
perfran	-0.194	-2.85
trend	0.044	11.78
lrpci	0.828	3.88
lur	-0.002	-0.07
Observations		240
Corrected <i>R</i> -squared		0.962
Mean of Dependent Variable		-5.924

population, and the real earnings at retail establishments divided by employment at retail establishments (a measure of per-capita earnings for workers at retail establishments). The estimated regression model showed that real sales per capita for the five hamburger chains increased both as the retail sector grew in real terms and with increases in the payments to those working in the retail segment. The *R*-squared for this regression is approximately 65 percent. Given the similarity in findings to the revenue per-capita models, the regression results are omitted.

Using the results of the market share equations and the total revenue per-capita equation at the five chains, it was possible to forecast the post-1993 sales period. The sales simulation proceeds under two scenarios. In the first

scenario, I assume that the actual revenues for the five hamburger chains for 1993–1994 were identical to their historical levels. My assumption here is that any loss in sales at Jack-in-the-Box would simply shift to another competitor but that total sales would not be affected. This assumption should be correct provided that the degree of substitution among the major hamburger chains is closer than that between Jack-in-the-Box and non-hamburger substitutes.

Under this zero-sum scenario, in which Jack-in-the-Box sales will go to another hamburger competitor, I calculate that the lost sales for 1993 and 1994 would be \$282 million for the ten regions. This estimate is similar to the estimate of lost sales using the revenue per-capita model.

9.4.2. Conclusions

As Foodmaker's annual sales circa 1993 were \$1.29 billion with two-thirds of the revenue derived from hamburger sales, annual hamburger revenues were approximately 859 million dollars. If investors concluded that the lost earnings were permanent, the stock price would fall in proportion to the decline in earnings due to the E. coli scare. As stock prices lost some 30 percent of their value, an estimate of lost earnings consistent with the stock decline would be \$258 million per annum. The econometric estimates show that lost sales were closer to \$133.5 million per annum (\$267m for two years). This figure is nearly double that reported in the trade news. The decline in stock prices does not necessarily reflect irrational expectations of investors. Investors correctly perceived the news to be much worse than that reported in the press. However, the initial stock price decline appears to have been too large to be explained by the financial consequences of the E. coli outbreak, even if investors believed the effects to be permanent. As revenues subsequently stabilized and returned to normal levels, the financial market revised stock prices accordingly and returned them to levels consistent with long run trends.

References

- Baker, J. E. 1986. *Lung Cancer Incidence Amongst Previous Employees of an Asbestos Mine in Relationship to Crocidolite Exposure and Tobacco Smoking*. Ph. D. Thesis.
- Bass, Frank M. 1969. "A New Product Growth Model for Consumer Durables," *Management Science*, Vol. 15 (January), No. 5, pp. 215–227.
- Beaton, W. Patrick. 1974. "The Determinants of Police Protection Expenditures," *National Tax Journal*, Vol. 27, No. 2, pp. 335–349.
- Becker, Gary S. 1991. "A Note on Restaurant Pricing and Other Examples of Social Influences on Price," *Journal of Political Economy*, Vol. 99, No. 5, pp. 1109–1116.
- Behrman, Jere R. and Steven G. Craig. 1987. "The Distribution of Public Service: an Exploration of Local Governmental Preferences," *American Economic Review*, Vol. 77, pp. 37–49.
- Berry, G., M. L. Newhouse, and P. Antonis. 1985. "Combined Effect of Asbestos and Smoking on Mortality From Lung Cancer and Mesothelioma in Factory Workers," *British Journal of Industrial Medicine*, Vol. 42, pp. 12–18.
- Berry, G., M. L. Newhouse, and M. Turok. 1972. "Combined Effect of Asbestos Exposure and Smoking on Mortality From Lung Cancer in Factory Workers," *Lancet*, Vol. 2, pp. 476–479.
- Blot, W. J., J. M. Harrington, A. Toledo, R. Hoover, C. W. Heath Jr., and J. F. Fraumeni. 1978. "Lung Cancer After Employment in Shipyards During World War II," *New England Journal of Medicine*, Vol. 299, pp. 620–624.
- Blot, W. J., L. E. Morris, R. Stroube, I. Tagnon, and J. F. Fraumeni. 1980. "Lung and Laryngeal Cancers in Relation to Shipyard Employment in Coastal Virginia," *Journal of the National Cancer Institute*, Vol. 565, pp. 571–575.
- Bovenzi, M. G. Stanta, G. Antiga, P. Peruzzo, and F. Cavallieri. 1993. "Occupational Exposure and Lung Cancer Risk in a Coastal Area of Northeastern Italy," *Internal Archives of Occupational and Environmental Health*, Vol. 65, pp. 35–41.
- Bureau of the Census. 1992. "Census of Population and Housing, 1990" Summary Tape File on CD-ROM Technical Documentation/prepared by the Bureau of the Census. Washington: The Bureau, (A-2).
- Bureau of the Census. 1992. "Census of Population and Housing, 1990" Summary Tape File on CD-ROM Technical Documentation/prepared by the Bureau of the Census. Washington: The Bureau, (A-8).
- Bureau of the Census. 1992. "Census of Population and Housing, 1990" Summary Tape File on CD-ROM Technical Documentation/prepared by the Bureau of the Census. Washington: The Bureau, (A-10).

- Bureau of the Census. 1992. "Census of Population and Housing, 1990" Summary Tape File on CD-ROM Technical Documentation/prepared by the Bureau of the Census. Washington: The Bureau, (A-11).
- Bureau of the Census. 1992. "Census of Population and Housing, 1990" Summary Tape File on CD-ROM Technical Documentation/prepared by the Bureau of the Census. Washington: The Bureau, (B-16).
- Bureau of the Census. 1992. "Census of Population and Housing, 1990" Summary Tape File on CD-ROM Technical Documentation/prepared by the Bureau of the Census. Washington: The Bureau, (B-28).
- Cicchetti, Charles J. 1971. "Some Economic Issues in Planning Urban Recreation Facilities," *Land Economics*, Vol. 47, No. 1, pp. 14–23.
- Connor, J. M. and E. B. Peterson. 1992. "Market-Structure Determinants of National Brand Private Label Price Differences of Manufactured Food Products." *Journal of Industrial Economics*, Vol. 40, pp. 157–72.
- Day, Kelly and George B. Frisvold. 1993. "Medical Research and Genetic Resources Management: The Case of Taxol," *Contemporary Policy Issues*, Vol. 11, No. 3, pp. 1–11.
- deKlerk, N.H., A. W. Musk, B. K. Armstrong, and M.S.T. Hobbs. 1991. "Smoking, Exposure to Crocidolite, and the Incidence of Lung Cancer and Asbestosis," *British Journal of Industrial Medicine*, Vol. 48, pp. 412–417.
- Department of Justice and Federal Trade Commission. *Horizontal Merger Guidelines*. 1992. "Horizontal Merger Guidelines" (April 2), revised April 8, 1997.
- DeSerpa, Allan C. 1994. "To Err is Rational: A Theory of Excess Demand for Tickets," *Managerial and Decision Economics*, Vol. 15, pp. 511–518.
- DeSerpa, Allan C. and Roger L. Faith. 1996. "Bruce: The Simple Economics of Mob Goods," *Public Choice*, Vol. 89, pp. 77–91.
- Dosser, Douglas. 1964. "Notes on Carl S. Shoup's 'Standards for Distributing a Free Governmental Service: Crime Prevention,'" *Public Finance*, Vol. 19, No. 4, pp. 393–402.
- Dubin, Jeffrey. 1998. *Studies in Consumer Demand—Econometric Methods Applied to Market Data*. Kluwer Academic Publishers, Boston, MA.
- Eckstein, Zvi, and Kenneth I. Wolpin. 1989. "The Specification and Estimation of Dynamic Stochastic Discrete Choice Models: A Survey," *The Journal of Human Resources*, Vol. 24, No. 4., pp. 562–598.
- Federal Energy Regulatory Commission (FERC). 2000. "Regulation of Short-Term Natural Gas Transportation Services, and Regulation of Interstate Natural Gas Transportation Services," 90 FERC 61, 109 (February 9) Order 637.
- Federal Energy Regulatory Commission (FERC). 1989. "On Developing a Framework for Assessing Competition in Natural Gas Transportation, Appendix C—A Further Discussion of Market Power and Competition Measures of Firms Within Connected Multi-Market Industries," (July).
- Fisher, Franklin M. 1980. "Multiple Regression in Legal Proceedings," *Columbia Law Review*, Vol. 80, pp. 702–36.
- Gardner, M. J., and A. G. Munford. 1980. "The Combined Effect of Two Factors On Disease in a Case-control Study," *Applied Statistics*, Vol. 29, No. 3, pp. 276–281.
- Grabowski, Henry G., and John M. Vernon. 1992. "Brand Loyalty, Entry and Price Competition in Pharmaceuticals After the 1984 Drug Act," *Journal of Law and Economics*, Vol. 35, p. 331.
- Griffin, Douglas B. 2000. "Schedule of Local Option Communities," April 2000: Alcoholic Beverage Control Board. May 1, 2000. <http://www.revenue.state.ak.us/abc/localopt/htm>.
- Hammond, E. C., I. J. Selikoff, and H. Seidman. 1979. "Asbestos Exposure, Cigarette Smoking and Death Rates," *Annals of the New York Academy of Sciences*, Vol. 330, pp. 473-490.

- Hannan, Timothy. 1997. "Market Share Inequality, the Number of Competitors and the HHI: An Examination of Bank Pricing," *Review of Industrial Organization*, Vol. 12, pp. 23–35.
- Haskell, Charles M., ed. 1995. *Cancer Treatment*, 4th Edition, W. B. Saunders Company, Pennsylvania: Philadelphia.
- Ignoffo, Robert J., Carol S. Viele, and Lloyd E. Damon, *Cancer Chemotherapy Pocket Guide*. 1998. Lippincott Williams and Wilkens Publishers.
- Kjuus, H., R. Skjaerven, S. Langard, J. T. Lien, and T. Aamodt. 1986. "A Case-referent Study of Lung Cancer, Occupational Exposures and Smoking. I. Comparison of title-based and Exposure-based Occupational Information," *Scandinavian Journal of Work, Environment & Health*, Vol. 12, pp. 193–202.
- Kjuus, H., R. Skjaerven, S. Langard, J. T. Lien, and T. Aamodt" 1986. "A Case-referent Study of Lung Cancer, Occupational Exposures and Smoking. II. Role of Asbestos Exposure," *Scandinavian Journal of Work, Environment & Health*, Vol. 12, pp. 203–209.
- Kralewski, J. E., L. Pitt, and B. Dowd. 1983. "The Effects of Competition on Prescription Drug Product Substitution," *New England Journal of Medicine*, Vol. 309, No. 4, pp. 213–216.
- Landen, Michael G., Michael Beller, Elizabeth Funk, Michael Propst, John Middaugh, and Ronald L. Moolenaar. 1997. "Alcohol-Related Injury Death and Alcohol Availability in Remote Alaska," *Journal of the American Medical Association*, Vol. 278, pp. 1755–58.
- Leffler, Keith B. 1981. "Persuasion or Information? The Economics of Prescription Drug Advertising," *Journal of Law and Economics*, Vol. 24, No. 1, p. 45.
- Leibenstein, H. 1950. "Bandwagon, Snob, and Veblen Effects in the Theory of Consumers' Demand," *Quarterly Journal of Economics*, Vol. 64, No. 183, p. 207.
- Liddell, F.D.K., D. C. Thomas, G. W. Gibbs, and J. C. McDonald. 1984. "Fibre Exposure and Mortality from Pneumoconiosis, Respiratory and Abdominal Malignancies in Chrysotile Production in Quebec, 1926-75," *Annals of the Academy of Medicine*, Vol. 13, pp. 340–344.
- Manildi, Ralph. 1988. *Journal of Corporate Computing*, (March/April), pp. 48–50.
- Martischinig, K. M., D. J. Newell, W. C. Barnsley, W. K. Cowan, E. L. Femman, and E. Oliver. 1977. "Unsuspected Exposure to Asbestos and Bronchogenic Carcinoma," *British Medical Journal*, Vol. 1, pp. 746–749.
- McDonald, J. C., F. D. K. Liddell, G. W. Gibbs, G. E. Eyssen, and A. D. McDonald. 1980. "Dust Exposure and Mortality in Chrysotile Mining, 1910-75," *British Journal of Industrial Medicine*, Vol. 37, pp. 11–24.
- Mirecki, Ted. 1987. "Speed infusion (Part II)," *PC Tech Journal*, Vol. 5 (April), No. 4, pp. 66–74.
- Nicolaou, K. C., Rodney K. Guy, and Pierre Potier. 1996. "Taxoids: New Weapons Against Cancer," *Scientific American*. (June), pp. 84–89.
- Noll, Roger G. 1974. "Attendance and Price Setting," in Roger G. Noll, ed., *Government and the Sports Business*, Washington, DC: The Brookings Institution, Vol. 115, p. 57.
- Norton, John, and Frank Bass. 1987. "A Diffusion Theory Model of Adapting and Substituting for Successive Generators of High-Technology Products," *Management Science*, Vol. 33 (September).
- O'Neill, Richard, *et al.* 1991. "On Developing a Framework for Assessing Competition in Natural Gas Transportation; Appendix B; Common Framework for the Discussion of Structure and Performance," FERC mimeo.
- Pastorino, U., F. Bearino, A. Gervasio, V. Pesenti, E. Riboli, and P. Crosignani. 1984. "Proportion of Lung Cancers Due to Occupational Exposure," *International Journal of Cancer*, Vol. 33, pp. 231–237.
- Rao, C. R. 1973. *Linear Statistical Inference and its Applications*, Wiley, New York.
- Rhoades, Stephen A. 1995. "Market Share Inequality, the HHI, and Other Measures of the Firm-composition of a Market," *Review of Industrial Organization*, Vol. 10, pp. 657–674.

- Rothman, K. J. 1976. "The Estimation of Synergy or Antagonism," *American Journal of Epidemiology*, Vol. 103, No. 6, pp. 506–511.
- Salkever, D. 1976. "The Use of Dummy Variables to Compute Predictions, Prediction Errors, and Confidence Interval," *Journal of Econometrics*, Vol. 4, pp. 393–397.
- Saracci, R. 1977. "Asbestos and Lung Cancer: An Analysis of the Epidemiological Evidence on the Asbestos-Smoking Interaction," *Indian Journal of Cancer*, Vol. 20, pp. 323–331.
- Scherer, F. M. 1993. "Pricing, Profits and Technological Progress in the Pharmaceutical Industry," *Journal of Economic Perspectives*, Vol. 7 (Summer), No. 3, pp. 97–115.
- Selikoff, I. J., E. C. Hammond, and J. Churg. 1968. "Asbestos Exposure, Smoking, and Neoplasia," *Journal of the American Medical Association*, Vol. 204, pp. 104–110.
- Selikoff, I. J., H. Seidman, and E. C. Hammond. 1980. "Mortality Effects of Cigarette Smoking Among Amosite Asbestos Factory Workers," *Journal of the National Cancer Institute*, Vol. 65, pp. 507–513.
- Shoup, Carl S. 1964. "Standards for Distributing a Free Governmental Service: Crime Prevention," *Public Finance*, Vol. 19, No. 4, pp. 383–392.
- Shoup, Carl S. 1989. "Rules for Distributing a Free Government Service Among Areas of a City," *National Tax Journal*, Vol. 42 (June), No. 2, pp. 103–121.
- Smith, Gordon V. 1997. *Trademark Valuation*. John Wiley & Sons, Inc., New York: New York.
- Steenland, K. and M. Thun. 1986. "Interaction Between Tobacco Smoking and Occupational Exposures in the Causation of Lung Cancer," *Journal of Occupational Medicine*, Vol. 28, pp. 110–118.
- Swift, E. M. 2000. "Hey Fans: Sit On It!," *Sports Illustrated*, Vol. 15 (May), p. 73.
- Temin, Peter. 1980. *Taking Your Medicine: Drug Regulation in the United States*, Harvard University Press, Massachusetts: Cambridge.
- Theil, H. 1961. *Economic Forecasts and Policy*. Amsterdam: North Holland.
- Vainio, H. and P. Boffetta. 1994. "Mechanisms of the Combined Effect of Asbestos and Smoking in the Etiology of Lung Cancer," *Scandinavian Journal of Work, Environment & Health*, Vol. 20, pp. 235–242.
- Vena, J. E., T. E. Byers, D. Cookfair, and M. Swanson. 1985. "Occupation and Lung Cancer Risk. An Analysis of Histologic Subtypes," *Cancer*, Vol. 56, pp. 910–917.
- Viscusi, W. Kip, John M. Vernon, and Joseph E. Harrington, Jr. 1998. *Economics of Regulation and Antitrust*, 2nd Ed. The MIT Press, Massachusetts: Cambridge.
- Walker, Sam. 2000. "Solving the Mystery of the Miami Heat and Its Empty Seats," *Wall Street Journal*, Vol. 21 (April), p. A1.
- Weicher, John C. 1971. "The Allocation of Police Protection in Income Class," *Urban Studies*, Vol. 157, pp. 207–220.
- Welki, Andrew M., and Thomas J. Zlatoper. 1994. "U.S. Professional Football: The Demand for Game-Day Attendance in 1991," *Managerial and Decision Economics*, Vol. 15, pp. 489–495.
- Welki, Andrew M., and Thomas J. Zlatoper. 1999. "U.S. Professional Football Game-Day Attendance," *American Economics Journal*, Vol. 27, No. 3, pp. 285–298.
- Werden, Gregory. 1981. "The Use and Misuse of Shipments Data in Defining Geographic Markets," *The Antitrust Bulletin*, Vol. 26 (Winter), No. 721.
- Woolf, B. 1955. "On Estimating the Relation Between Blood Group and Disease," *Annals of Human Genetics*, Vol. 19, pp. 251–253.
- Yorke, Brian A. 1984. "Pharmaceutical Patent Protection," *Medicinal Research Reviews*, Vol. 4, No. 1, pp. 25–46.

Index

- ACS, 118–121
additive, xix, 87, 89, 90, 99, 101, 106, 107, 109–112, 114, 115, 119, 121–125, 129–132, 138–140
additivity, xix, 89, 105, 112, 116, 129, 132, 138, 140
admail, xviii, 51–54, 58, 59, 61–64
advertise, 37
aftermarket, 80
aggregation, 20
Alaska, xx, 167–170, 172, 173, 181–185, 191
alcohol, 184, 187, 189, 191, 205
alkalation, 156
alkaloid, 148
alkylating, 156, 159
amosite, 117, 120, 122, 206
amphibole, 117
ANOVA, 139, 140
antineoplastic, 155, 156
antitrust, 146, 154, 157, 206
Antonis, 122
ANVSA, 173–181, 185–189
Armstrong, 129
asbestos, xix, 111, 117–132, 138–140, 203, 205, 206
asbestosis, 89
asian, 186, 191
asymptotic, 99, 116
attendance, xviii, 31, 33, 39, 43, 47, 49, 206

bacteria, xx, 193
Baker, 121, 123, 127
bandwagon, xviii, 34, 49
basis, xvii, 1–3, 9–11, 20–24, 30, 32, 54, 106, 147, 182, 195, 196
Bass, xviii, 34, 74, 76
BEA, 195
Bearino, 124, 125
Beaton, 172
Becker, xvii, xviii, 31, 34, 35, 37, 39, 48, 49

beef, 193, 194
Behrman, 171, 172
Berry, 121, 122, 132, 138
bias, 20
black, 37, 43, 47, 48, 186, 191
blockbuster, 163
Blot, 126–128
Boffetta, 122–125, 129, 131
Bonferroni, 89, 99, 107
Bouquai, 69
Bovenzi, 126
bowel, 193
brand, xx, 69, 70, 76, 142, 144, 145, 204
breast, 141, 148, 149, 155–158, 162–165
bronchogenic, 205
Byers, 129

CAF, 69
Canada, 51, 52, 54
Canadian, 52–54
cancer, xix, xx, 90, 118, 119, 121–129, 132, 140–142, 148, 149, 155–158, 162, 165, 203–206
carboplatin, 156
carcinogens, xix, 118
cases, 67, 68, 89, 90, 96, 97, 117, 122–129, 131, 139, 140, 142, 151, 153, 157, 171
CDP, 169, 173
cellular, 129, 156
chemotherapy, xx, 146, 156–158, 205
chrysotile, 117, 122, 205
Cicchetti, 171
cigar, 118, 120, 126
cigarette, 125, 126, 204, 206
cisplatin, 156
cohort, xix, 89, 90, 110, 111, 113–115, 118–120, 122, 129, 138–140
collude, 150
collusion, 150
comparable, 122, 182

- competition, xvii, xx, 1, 3, 12, 22–26, 32, 80, 141, 145, 147, 148, 151, 155, 165, 204, 205
 computer, xviii, 65, 66, 69–72, 74, 76, 80, 81, 83, 84
 concentration, xx, 11, 28, 142, 150–153, 155, 162, 165
 Connor, 9
 controls, 12, 90, 97, 122–129, 131, 139, 187
 coprocessor, 66–70
 covariance, 58, 93, 98, 100
 CRADA, 148, 149, 164
 Craig, 171, 172
 crime, 171, 172, 184, 187, 188, 204, 206
 crocidolite, 122, 125, 203
 cytotoxic, 156
 cytotoxic, 156

 daughterboard, 67, 70
 Day, 148, 149
 deKlerk, 125, 126, 131
 demand, xvii, xviii, xx, 2, 10, 14, 21, 23, 31–35, 37, 39, 43, 47–49, 51–54, 57, 59, 68, 74, 80, 142, 145, 153–155, 189, 194, 204, 206, 207
 density, 171, 184, 186, 189, 191
 deregulation, 1
 DeSerpa, 31, 34, 35, 37, 39, 48, 49
 diagnosis, 126
 diffusion, 205
 discount, xvii, 1, 9, 10, 12, 14, 21–26, 30, 33
 discounting, 1, 2, 9–12, 14, 20–22, 24, 25, 27, 30
 discrimination, 172, 173, 187, 189
 disease, xix, 87–89, 108, 117, 146, 194, 204
 distance, xx, 11, 12, 21, 22, 25–27, 168, 183, 186, 188, 189, 191
 DNA, 156
 docetaxel, 155
 Dowd, 144
 drug, xx, 141–147, 149, 155–158, 163, 164, 166, 194, 204–206
 Dubin, 9, 28, 51–53, 207
 durable, 74, 207
 dust, 117, 118, 122, 123

 Eckstein, 80
 ecological, 171
 econometric, xviii, xx, 2, 10, 21, 31, 39, 51, 167, 168, 194, 195, 201, 207
 economics, 80, 172, 204–207
 EIA, 71
 elastic, xvii, 23, 30, 53
 elasticity, 52, 53, 153–155
 endocrine, 156
 enforcement, 168, 169, 173–181, 183, 184, 186, 187, 189–191
 environment, 1, 170, 205, 206

 epidemic, xx, 193
 epidemiology, 89
 Eskimo, 182
 estrogen, 156
 experiments, 146
 exposure, xix, 87–91, 95, 96, 111, 117–132, 138–140, 203–205

 Faith, 34
 FDA, 142, 143, 146–150, 163, 164
 FERC, 3, 10, 153, 204, 205
 fiber, 117, 123
 Foodmaker, 194
 football, xvii, xviii, 31–33, 35, 37, 39, 43, 49, 206
 forecast, xviii, 51, 54, 57, 59, 197, 200
 Frisvold, 148, 149
 FTC, 154

 game, 32, 33, 39, 43, 47, 49
 Gardner, 89, 104
 gas, xvii, 1–3, 10–12, 14, 20–22, 26, 28, 117, 151, 204, 205
 gastrointestinal, 117
 gender, 121, 138, 139
 generic, 65, 83, 142–145, 147, 158, 163, 166
 Grabowski, 145

 hamburger, 194, 195, 199–201
 Hammond, 118–120, 129, 138
 Hannan, 9
 hardware, 67
 Harrington, 126, 127
 Haskell, 156
 helpful, 49, 155
 Herfindahl, xx, 22, 142, 152, 155
 heteroscedasticity, 187
 HHL, 9, 27, 28, 150–155, 157, 158, 162, 165
 hispanic, 43, 47, 48
 Hobbs, 129
 homicides, 188, 189
 hormonal, 156
 hospital, 128, 193

 Ignoffo, 157
 IMS, 157, 158
 income, 43, 82, 168, 172, 182, 183, 186, 188, 189, 195, 197, 206
 independence, 108, 112
 industrial, 172, 203–205
 inelastic, 53, 74
 inelasticity, 53
 inequality, 170, 171
 inflation, 54
 infringement, 83
 injury, 205
 insulation, 117–119

- interaction, xvii, 31, 37, 87, 113, 123, 125, 131, 155
 ISA, 69
- kidney, 193, 194
 Kjuus, 128
 Koch, 1–3, 9–14, 20–28, 30
- laryngeal, 203
 larynx, 117
 Leffler, 144
 license, 31, 65, 148, 149
 licensor, 82, 83
 Liddell, 122, 123, 138
 lifecycle, 74, 76
 likelihood, 12, 33, 34, 89, 90, 101, 105–107, 116, 125, 132, 138
 logistic, 87, 89, 126
 logit, 87, 88, 90, 125, 127–129, 198
 lottery, 35
 Louisiana, 2, 3
 lung, xix, 90, 117–119, 121–129, 132, 140–142, 148, 149, 155–158, 162–165, 203–206
 lymphoma, 157
- macroeconomic, 52
 mail, xviii, 51, 52, 59
 male, 118–122, 128, 182, 186, 191
 malignancies, 205
 Manildi, 68
 manufacturing, 83, 117
 market, xvii, xx, 1–3, 9, 11–13, 20–28, 30, 31, 33–35, 37, 67, 68, 71, 74, 76, 82, 83, 141–146, 149–158, 162–165, 194, 198–201, 204, 207
 Martischnig, 124
 McDonald, 122, 123
 meat, 193, 194
 medical, 124, 148, 205, 206
 memory, 66, 74, 76
 mesothelioma, 203
 microprocessor, 68, 69, 72, 76
 microtubular, 156
 microtubules, 156
 mines, 123
 Mirecki, 68
 mitotic, 156, 160
 MMBTU, 2
 module, 67
 molecular, 157
 monopoly, 31, 142, 145, 149, 150, 153
 morbidity, 89, 110
 Morris, 127, 128
 mortality, 89, 110, 122, 127, 203, 205
 motherboard, 67–72, 76, 80, 82, 83
 multinomial, 92, 101, 107, 198
- multiplicative, xix, 87–91, 94, 96, 98, 99, 104, 106, 107, 110, 112–115, 118, 119, 121–126, 128–132, 138–140
 multiplicativity, xix, 89, 105, 116, 129
 Munford, 89, 104
 Musk, 125, 126
- NBA, 37
 NCI, 148, 149, 164
 NDA, 147, 149
 neoplasms, 157
 Newell, 124
 Newhouse, 121, 122
 NFL, xvii, xviii, 31–33, 39, 47, 49
 Nicolaou, 155
 NIH, 149, 164
 Noll, 39
 nucleic, 156
- O'Neill, 151
 OARS, 181
 officer, 142, 170, 181, 191
 oncology, 149, 155, 157, 163, 165
 option, xviii, 32, 33, 49, 71, 72, 80, 81, 184, 204
 orphan, 145–147, 163
 ovarian, xx, 141, 142, 148–150, 155–158, 162–165
 overclocking, 66, 67
- paclitaxel, 141, 142, 148–150, 155, 156, 165, 166
 Pareto, 172
 partnerships, 145
 Pastorino, 124, 125
 patent, xviii, xx, 65, 66, 68, 69, 82, 83, 86, 141–147, 149, 206
 patents, 65, 66, 147
 patrol, 169
 Pearson, 112
 Pentium, 65, 66, 68, 69, 72, 76
 pharmaceutical, xix, xx, 141–147, 149, 152, 158, 163–165, 206
 pharmacists, 144
 physicians, 157
 pipeline, xvii, 1–3, 11, 12, 14, 22, 25–27
 plastics, 117
 playoff, 32, 33
 pleural, 89
 police, xx, 167–173, 183, 184, 190, 203, 206
 poultry, 194
 poverty, 43, 168, 182, 183, 186, 187, 191
 prediction, 59, 206
 prescription, 205
 pricing, 3, 9, 10, 20, 21, 25, 27, 33, 80, 150, 203
 probability, 12, 87–92, 106, 108, 113, 130, 157
 processor, 67, 68, 72, 76, 80, 81
 productivity, 171
 profits, 82, 83, 145, 146, 150, 163, 206

- promotion, 144
 prostate, 142, 157
 protection, xx, 141, 145, 146, 169–172, 181, 203, 206
 PSL, 32

 QSR, 195

 race, xx, 48, 127, 168, 172, 182, 187–189
 receipt, 3, 9, 10, 20–22, 26, 27
 recreation, 204
 recreational, 171
 reflects, 3, 23, 33, 130, 154, 162, 187, 188
 refractory, 149, 156, 164
 regression, xviii, 11, 20–22, 24, 26, 28, 30, 39, 47, 51–54, 57, 58, 126, 139, 140, 182, 183, 187–189, 196, 197, 199, 200, 204
 regulation, 1, 3, 204, 206
 reservation, 10, 32, 35
 respiratory, 205
 restaurant, 34, 37, 193, 195, 203
 retail, xviii, 51–54, 58, 172, 195, 199, 200
 retrospective, 89, 90
 Rhoades, 9
 risk, xix, 81, 87–91, 95, 97, 99, 109, 111, 117, 118, 121, 124, 129, 130, 132, 144, 194, 203
 Rothman, 100, 107–110, 132, 140
 royalty, xviii, 65, 66, 72, 80, 82, 83
 rural, xx, 167, 169, 189

 Salkever, 57
 salmonella, 194
 salvage, 81, 156
 Saracci, 121, 124
 sarcoma, 149
 Scherer, 144
 seasonality, 197
 seat, 31, 43, 47, 48
 SEER, 142
 Seidman, 119, 120, 129, 138
 Selikoff, xix, 118–121, 129, 138, 139
 sellouts, 31, 39, 49
 share, 11–13, 21, 22, 24, 26, 28, 68, 145, 149–153, 158, 162–165, 189, 194, 198–200, 205
 Sherman, 146
 shipbuilders, 127
 shipper, 3, 22, 26–28
 shipyard, 127, 203
 Shoup, 170–172
 simulation, 197, 200
 simultaneous, 9
 Skjaerven, 128
 SLN, 11, 12, 26, 28
 Smith, 82, 83

 smoker, 122
 smokers, 118–125, 127, 128, 131, 138, 140
 smoking, xix, 111, 117–119, 121–132, 139, 140, 203, 204, 206
 social, xvii, 31, 34, 37, 148, 171, 172, 184, 203
 socket, 66–72, 80–83
 sports, 43, 205, 206
 stadium, 34, 37, 39, 47
 Stanta, 126
 statistic, xx, 9, 28, 58, 59, 92, 99, 106, 107, 112, 130–132, 138, 140, 142, 151–153
 statutes, 163
 Steenland, 121, 122, 124, 131
 strategy, 140, 145
 stratification, 121
 supplier, 24, 25, 150, 193
 survey, 11, 14, 26, 27, 119, 157, 158, 172
 Swift, 37
 synergism, 87, 88, 123
 synergistic, xix, 87, 88, 117, 118, 139, 140
 synergy, xvii, xix, 89, 107, 108, 118, 119, 124, 130–132, 138–140, 206

 Taiwan, 71
 Taiwanese, 71
 taxane, 155, 157
 taxanes, 155, 156
 taxoid, 155
 Taxol, 141, 142, 149, 150, 164, 165
 Taxus, 148
 technology, xviii, 66–71, 74, 76, 147–149, 207
 Temin, 144
 Theil, 58, 59
 therapies, 149, 157
 Thomas, 123
 thrombocytopenic, 194
 Thun, 121, 122, 124, 131
 ticket, xviii, 31–33, 37, 39, 43, 47–49
 tobacco, xix, 117, 126, 140, 203, 206
 tobit, 21, 49
 trademark, 66, 82, 206
 treatment, 142, 148, 157, 158, 182, 193
 tribal, 170
 trooper, xx, 167, 168, 172, 173, 181–190
 tumoricidal, 156
 Turok, 121

 updates, 182
 upgrade, xviii, 66–71, 74, 76, 80–83
 upgradeable, 69, 70, 76
 upgrades, 66–69, 80
 USDA, 149

 Vainio, 122–125, 129, 131
 Vena, 129
 Vernon, 145
 Viscusi, 142–144
 VPO, 170, 190

INDEX

211

VPSO, 170, 184, 186, 189, 190

Wald, 89, 99, 101, 107, 115, 116, 130, 132, 138,
140

wallboard, 117

warranties, 69

wealth, 81

Weicher, 172

Welki, 39

Werden, 154

wildlife, 169, 181

Wolpin, 80

women, 121, 141, 187

Woolf, 89, 95, 96, 98

Yew, 148–150

Yorke, 147

Zlatoper, 39

About the Author

Jeffrey A. Dubin received his A.B. in Economics from the University of California, Berkeley in 1978 and a Ph.D. in Economics from the Massachusetts Institute of Technology in 1982. In 1982, Dr. Dubin joined the faculty of the California Institute of Technology where he is currently an Associate Professor of Economics.

Dr. Dubin was Senior Advisor at Putnam, Hayes, and Bartlett from 1989 through 1992. He was Senior Economist at Arthur Andersen Economic Consulting from 1992 through 1993 and Director of Statistics and Econometric Analysis at Arthur Andersen from 1993 through 1996. In 1996, he became a co-founding partner of the Pacific Economics Group. Dr. Dubin received the Econometric Society Frisch Medal in 1986. His first book was "Consumer Durable Choice and the Demand for Electricity," published by Elsevier North-Holland in 1985. His second book was "Studies in Consumer Demand—Econometric Methods Applied to Market Data" published by Kluwer Academic Publishers in 1998.