

Incomplete Cartels and Antitrust Policy

ISBN 978 90 3610 146 2

Cover design: Crasborn Graphic Designers bno, Valkenburg a.d. Geul

This book is no. **463** of the Tinbergen Institute Research Series, established through cooperation between Thela Thesis and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.

Incomplete Cartels and Antitrust Policy: Incidence and Detection

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. D.C. van den Boom
ten overstaan van een door het college voor promoties
ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op vrijdag 20 november 2009, te 12:00 uur

door

Adriaan Maarten Bos

geboren te Rotterdam

Promotiecommissie

Promotoren: *Prof. dr. A.W.A. Boot*
Prof. dr. M.P. Schinkel

Overige Leden: *Prof. dr. P.J.G. van Cayseele*
Prof. dr. J.E. Harrington
Prof. dr. J. Hinlopen
Prof. dr. J.J.M. Theeuwes
Dr. J. Tuinstra

“I am so not competitive. In fact, I am the least non-competitive, so I win...”

- Family Guy¹

¹See www.familyguyquotes.com.

Contents

List of Figures	vii
Acknowledgments	ix
1 Motivation and Outline	1
1.1 Introduction	1
1.2 Motivation	3
1.3 Methodology	5
1.4 Outline	9
2 The Economics of Incomplete Cartels	11
2.1 Introduction	11
2.2 Incomplete Cartels in Practice	14
2.3 Foundations of Cartel Theory	19
2.3.1 The Incentive Constraint	21
2.3.2 The Participation Constraint	22
2.3.3 Why are Cartels Bad?	23
2.4 On the Profitability of Incomplete Cartels	25
2.4.1 Collusive price leadership	27
2.4.2 Differentiated goods	29
2.4.3 Quantity competition	30
2.4.4 Comparison	33
2.5 On the Sustainability of Incomplete Cartels	33
2.5.1 Collusive price leadership	34
2.5.2 Differentiated goods	34

2.5.3	Quantity competition	35
2.5.4	Comparison	36
2.6	The Participation Puzzle	37
2.6.1	Collusive price leadership	38
2.6.2	Differentiated goods	39
2.6.3	Quantity competition	40
2.6.4	Comparison	42
2.7	Coalition Formation with Positive Externalities	42
2.7.1	Simultaneous cartel formation	43
2.7.2	Sequential cartel formation	45
2.8	Incomplete Cartels and Firm Heterogeneity	47
2.9	Incomplete Bidding Rings	49
2.10	Discussion	53
3	A Theory of Incomplete Cartels with Heterogeneous Firms	55
3.1	Introduction	55
3.2	A Model of Collusion with Asymmetric Capacity Constraints	58
3.2.1	Static Nash Equilibrium	59
3.2.2	Infinitely Repeated Game	61
3.3	Optimal Cartel Size	64
3.3.1	Costless Collusion	65
3.3.2	Costly Collusion	67
3.4	Incentives to Collude	70
3.5	The Most Profitable Cartel	75
3.6	Incomplete Cartels and Mergers	77
3.6.1	Merger Incentives	78
3.6.2	Coordinated Effects of Mergers	80
3.7	Concluding Remarks	89
4	Cartel Detection and Antitrust Law Enforcement	93
4.1	Introduction	93
4.2	Goal and Scope of Cartel Detection	95
4.2.1	Why do we need Cartel Detection?	96
4.2.2	Why do we need Economics to Detect Cartels?	97
4.3	Economic Methods of Cartel Detection	99
4.3.1	Market Structure	100
4.3.2	Cartel Conduct	105
4.3.3	Market Performance	110
4.4	Potential Pitfalls in Cartel Detection	111
4.4.1	No Result, is a Result	112
4.4.2	Descriptive Flaws, Empirical Limitations and Theoretical Complications	112
4.4.3	On the Problem of Defining a Workable Benchmark	114
4.5	Detecting Incomplete Cartels	114
4.5.1	A Variance Screen for Collusion: an Example (1)	115
4.6	The Need for Industry Specific Detection Tests	118

4.6.1	A Variance Screen for Collusion: an Example (2)	120
4.7	Discussion	121
5	Tracing the Base: A Topographic Test for Collusive Basing-Point Pricing	123
5.1	Introduction	123
5.2	Basing-Point Pricing	126
5.3	A Model of Basing-Point Pricing	131
5.3.1	Competitive Basing-Point Pricing	134
5.3.2	Collusive Basing-Point Pricing	135
5.4	Detecting Collusive Basing-Point Pricing	137
5.4.1	Variance Screens	138
5.4.2	Bid-distance Correlation	139
5.5	Tracing the Base	141
5.5.1	Base Recovery	142
5.6	A Likelihood Measure of Collusion	144
5.7	Concluding Remarks	149
6	Summary and Conclusions	153
6.1	Introduction	153
6.2	Main Findings	153
6.3	Implications for Antitrust Policy	156
6.4	Future Research	158
6.4.1	Cartel Formation with Heterogeneous Firms	158
6.4.2	Disentangling Overt Collusion and Tacit Collusion	159
A	BaseLocator[®]	163
A.1	Steps to Trace the Base	163
A.2	Kernel of the Software	165
	References	171
	Samenvatting (Summary in Dutch)	185

List of Figures

2.1	Incentives to collude illustrated.	20
2.2	Collusive price leadership equilibrium.	28
4.1	Price pattern of the citric acid market 1987-1997. Source: Connor (2001).	108
4.2	Price pattern of the lysine market 1992-1995. Source: Connor (2001).	108
4.3	Canonical cartel price path. Source: Harrington (2006).	109
5.1	Bird's eye view of basing-point pricing in a regionally isolated market.	127
5.2	Cloud-shaped blocking zone in a regionally isolated market.	137
5.3	Bird's eye view of base locations in the collusive base zone resulting in a price variance that mimics competition.	140
5.4	A least-squares point estimator of the base location.	144
5.5	Robustness of the LoC-measure as an enforcement tool for two examples of collusive basing-point pricing.	148

Acknowledgments

For a few years now, I have been practicing the art of surfing. In thinking about my time as a Ph.D. student, it struck me that there is an interesting parallel between pursuing a Ph.D. and surfing. In fact, one can view the process of writing a Ph.D. thesis as surfing in the “ocean of ideas”. In both cases, catching a wave requires balance, patience, the ability to be at the right place at the right time, technique and a bit of luck. Waves come in many different forms. Some look good from a distance, but prove useless when they get up close to you. Some are great, but impossible to deal with. Unfortunately, you often realize this too late due to a lack of experience. Occasionally, somebody else is catching a good wave right in front of you. Moreover, once you are on a promising wave, you must go from the bottom to the top and back down again in order to have a successful ride. Finally, you need to have the skills and courage to change direction along the way whenever the circumstances require it. During my time as a Ph.D. student, I have had plenty of rides, both successful and less successful, which culminated in this Ph.D. thesis.

In writing my Ph.D. thesis, I received support from many people, some of whom deserve special mention. First of all, I would like to express my gratitude to my supervisors Arnoud Boot and Maarten Pieter Schinkel from whom I have learned a lot in many ways. I thank Arnoud for emphasizing the importance of focus and timing, which was much needed. I thank Maarten Pieter, who was my daily supervisor, for his encouragement, advice and patience during all stages of the research process. In particular, he succeeded in convincing me that my inclination to select monster waves only is not very fruitful. It took me a while to realize that starting with the smaller ones is often already challenging enough. More generally, I thank the Amsterdam Center for Law and Economics for providing me with excellent surfing conditions.

As a Ph.D. student, I was fortunate to have the opportunity to train at different surf spots. Part of the research has been conducted at the Department of Economics of the

Johns Hopkins University in Baltimore. I thank this Department for its hospitality. I am particularly indebted to Joe Harrington with whom I had many interesting discussions about cartel theory, antitrust policy and, of course, soccer. I have benefited greatly from his expertise and I thank him for his useful comments on my research and, in particular, for his invaluable input to Chapter 3. Special thanks also go to Eelko Ubels for his contribution to Chapter 5. His programming techniques proved increasingly important in understanding the complexities of competitive and collusive basing-point pricing. I have very much enjoyed our uncountable discussions that often went beyond the question of how to trace bases. Furthermore, the quality of the thesis improved significantly due to comments I received from Patrick van Cayseele, Giuseppe Dari-Mattiacci, Jeroen Hinlopen, Jan Tuinstra and Jeroen van de Ven.

Additionally, I wish to express my gratitude to Igor van Loo who provided mental support, very good food and a place to sleep in Amsterdam during the final phase of this thesis. I also thank Martijn Han, who allowed me to stay at his place in Amsterdam during the summer of 2008, and my other “fellow surfers” Mario Bersem, Marie Goppelsröder, Andrea Günster, Floris Heukelom, Marko Lankhorst, Matej Marinc, Jacob Rüggeberg, Francesco Russo and Josephine van Zeben for their valuable input to this thesis and for the many memorable moments during my Ph.D. years. On a more personal level, I wish to thank my father Jaap Bos, my mother Else Bos and my brother Ruben Bos for their unconditional support throughout my life.

Finally, I thank my girlfriend and love of my life Anita van den Berg who surfed with me since the start and with whom I hope to share many epic waves in the future.

Iwan Bos, Maastricht, September 20, 2009.

1

Motivation and Outline

“In economic life, competition is not a goal: it is a means of organizing economic activity to achieve a goal.” – George Stigler¹

1.1 Introduction

Free market competition is one of the cornerstones of a capitalist society. In a free economy, individuals are left free to pursue their own profits and this is widely believed to enhance economic progress. Adam Smith was among the first to express the view that individual market players need not have the objective to promote social welfare.² Higher welfare standards for society at large will be the unintended consequence of competition between rivals who are primarily motivated by their own well-being. The idea that a free enterprise system is beneficial for society is founded on the premise that markets are indeed competitive. However, competition, by its very nature, erodes the individual gains of competitors. It is precisely for this reason that firms quite naturally strive for obtaining a position uncontested by competition.³ A competitive order is therefore persistently threatened by undertakings which, driven by self-interest, attempt to reduce competitive pressure.

One of the most direct ways in which firms can restrict competition is to engage in a cartel. The word “cartel” is a diminutive of the Latin term *charta*, which can be loosely

¹See Stigler (1983), p. 5.

²Smith (1994). The first edition of this famous book, which is originally titled ‘An Inquiry into the Nature and Causes of the Wealth of Nations’, has been published in London, U.K. in 1776.

³See Neumann (2001).

translated as a card, letter or paper.⁴ Stocking and Watkins (1948) provide a complete definition of what nowadays is understood by a cartel. They define a cartel as “an arrangement among, or on behalf of, producers engaged in the same line of business, with the design or effect of limiting or eliminating competition among them”. To what extent a cartel is effective depends in large part on the number of undertakings that participate in the agreement. According to Liefmann (1977), a cartel naturally aims at excluding *as far as possible* competition within its range of activity. The limit case is then a cartel that succeeds in embracing all enterprises. Indeed, the conventional wisdom is that the perfect (or full) cartel is one that eliminates all competition in the market.

Many known cartels, however, did not include all firms in the relevant market and consequently were operating in the presence of one or more independent outsiders, which formed a so-called competitive fringe. This dissertation is about such *incomplete cartels*.⁵ An incomplete cartel is defined as a cartel with less than one hundred percent market share. In other words, a cartel is incomplete when it does not control all industry supply. A cartel in the North Atlantic shipping industry, for example, controlled approximately 75% of the market.⁶ In the carbonless paper industry, the joint market share of cartel members was estimated to be 85-90%.⁷ Perhaps the most famous cartel, the Organization of the Petroleum Exporting Countries (OPEC), is not all-inclusive. For instance, major players like the United States and Russia are not a member of OPEC. There are many other examples of cartels that did not encompass all market players.

This thesis consists of two main parts. The first part analyzes the nature of incomplete cartels. The second part is about cartel detection and explores ways in which incomplete cartels can be detected. The dissertation is organized around four main research questions.

- What explains optimal cartel size to be less than all-inclusive?
- What (type of) firms take part in an incomplete cartel and what (type of) firms remain independent outsiders?
- What is the relationship between industry structure and optimal cartel size?
- How can economics be used to detect (incomplete) cartels?

Each of these questions is briefly introduced in the next section.

⁴See the American Heritage Dictionary of the English Language (1969).

⁵Incomplete cartels are sometimes referred to as *partial cartels*. In this dissertation, both terms are used interchangeably.

⁶See Escrihuela-Villar (2003).

⁷See Levenstein *et al.* (2003).

1.2 Motivation

The first main aim of this dissertation is to provide a rationale for the existence of incomplete cartels. Standard economic theory of industrial collusion predicts that the most profitable cartel arrangement is one in which all firms in the market together operate as a multiplant monopolist. The question therefore arises: why did many cartels not monopolize the entire market? From an economic theoretical point of view, this question can be reformulated as: What explains optimal cartel size to be less than all-inclusive? That is to say, under what conditions is an incomplete cartel more profitable than an all-inclusive cartel? A possible explanation is that the full cartel is not viable, because one or more firms lack the incentive to abide by the cartel contract. Alternatively, reaching an agreement between all market players might prove to be too challenging, e.g., when there is a substantial difference in unit production cost. Also, a cartel may consciously exclude one or more firms so as not to attract too much attention from direct purchasers or an antitrust authority.

The next main research topic concerns the composition of a cartel. Given a less than all-inclusive cartel, the question to be addressed is: Who is in and who is out? To put it differently, what are the traits of firms that join the cartel? To be able to analyze this question from an economic theoretical perspective requires a setting in which firms differ in at least one respect. For example, some firms might have a more efficient production process or a better access to valuable information sources. Also, a subset of sellers might be located strategically in the market. Alternatively, we might conjecture that the incentive to collude is affected by the position of a firm in the industry. For instance, larger companies may be more inclined to join a cartel or *vice versa*.

Moreover, in the literature it is well-established that some industries are more prone to collusion than others.⁸ In a related vein, we might conjecture that particular industry structures are more conducive to the formation of incomplete cartels as opposed to full cartels. In other words, what is the relationship between industry structure and optimal cartel size? This question is interesting in its own right, but it also potentially yields some important insights that are helpful in antitrust law enforcement.

To safeguard competition in the market place, most capitalist societies have adopted a set of antitrust laws and set up antitrust agents that are given the task to enforce these ‘rules of competition’. The precise content of antitrust laws varies per jurisdiction, but cartel agreements are typically declared illegal. At the end of the nineteenth century there was a growing concern in the United States about the vast increasing economic power of large corporations, which undermined the functioning of dynamic competitive markets and weakened the ability of governmental institutions to prevent excessive business practices. This resulted in the enactment of the Sherman Act in 1890. With the Sherman *Anti-Trust* Act the government had officially established the first measure to prohibit trusts and it gave authority to the federal government to institute proceedings against these practices in order to dissolve them. The first article of the Sherman Act reads:

⁸See, for example, Chapter 4 of Motta (2004).

“Every contract, combination in the form of trust or otherwise, or conspiracy, in restraint of trade or commerce among the several States, or with foreign nations, is declared to be illegal. Every person who shall make any contract or engage in any combination or conspiracy hereby declared to be illegal shall be deemed guilty of a felony, and, on conviction thereof, shall be punished by fine not exceeding \$10,000,000 if a corporation, or, if any other person, \$350,000, or by imprisonment not exceeding three years, or by both said punishments, in the discretion of the court.”

Although there existed a strong consensus about the need for antitrust legislation it was not immediately clear how exactly these rules of competition had to be developed and implemented. One of the reasons is that the act was not well documented and very generally formulated. For instance, the term “restraints of trade” was not defined. As a result, most of the ideas written down in the act could not be applied directly and, consequently, needed further explanation, which was to be given by the Courts in concrete antitrust cases. This rather extensive discretion in effect paved the way for, what later became known as, “the per se rule versus rule of reason debate”. The central issue in this debate was if some business conduct could be held illegal per se, without any further inquiry of its effects on competition.

The main result of the antitrust development in courts is that around 1970 certain business practices were judged under a per se rule and some cases were dealt with by applying a reasonableness test. Today, hard-core cartel arrangements are viewed illegal per se. Simply put, this means that naked cartels are never thought ‘reasonable’ business practice. To prosecute this anticompetitive conduct it suffices to convince the judge that suspect parties indeed made an explicit agreement about strategies that directly have impact on the pricing mechanism.

In contrast to developments in the United States around the year 1900, European governments expressed quite a different attitude towards cartels. It was not uncommon to tolerate explicit cartel agreements, although such contracts were often not enforceable by courts. A great many cartels were openly sponsored by governments. This changed at the beginning of the European Union when there was a growing concern about free trade between Member States. Cartel practices have been declared incompatible with the common market in the Treaty of Rome of 1957. Article 85(1) of the original treaty (nowadays Article 81(1)) reads,

“The following shall be prohibited as incompatible with the common market: all agreements between undertakings, decisions by associations of undertakings and concerted practices which may affect trade between Member States and which have as their object or effect the prevention, restriction or distortion of competition within the common market, and in particular those which:

- (a) directly and indirectly fix purchase or selling prices or any other trading conditions;
- (b) limit or control production, markets, technical development, or investment;

- (c) share markets or sources of supply;
- (d) apply dissimilar conditions to equivalent transactions with other trading parties, thereby placing them at a competitive disadvantage;
- (e) make the conclusion of contracts subject to acceptance by the other parties of supplementary obligations which, by their nature or according to commercial usage, have no connection with the subject of such contracts.”

In Europe, like in the United States, the agreements between firms that are considered most harmful are those that have a direct effect on the price mechanism.

In order to ensure compliance, an antitrust authority must know if and where antitrust laws are violated. There currently exist three ways in which antitrust agencies find out about cartel activity. These are,

- Complaint procedures
- Leniency programs
- Cartel detection

The complaint procedure allows customers and competitors of the cartel to file a complaint with the antitrust authority. Also, (former) employees may decide to ‘blow the whistle’. Simultaneously, authorities learn about cartels due to the leniency program. This program offers firms that take part in a cartel agreement an opportunity to self-report in exchange for partial or full immunity of the fine depending on the position a firm takes in a sequence of self-reporters. Typically, the first to self-report receives the highest discount. One of the main advantages compared to the complaint procedure is that the information that self-reporters bring to the authority is likely to be much more detailed, which makes conviction, *ceteris paribus*, easier. In addition, the antitrust agency may start to search for cartel activity upon its own initiative. Note that the first two enforcement instruments are passive in the sense that the antitrust authority moves second, while the third instrument is pro-active.

The second part of this dissertation is about the scope and limits of *economic* methods of cartel detection. In particular, the focus is on detection methods that can be applied to screen markets for incomplete cartels. Understanding what type of industries are prone to the formation of incomplete cartels may shed some light on whether or not a cartel detection test is likely to be effective. For example, a detection technique may require the competitive fringe as benchmark and, as a result, will fail to identify collusion when the cartel is all-inclusive. Furthermore, the detection of incomplete cartels is of special interest, because the behavior of outsiders presumably differs from that of cartel members. In principle, economic theory can help to identify such differences.

1.3 Methodology

This dissertation is almost exclusively theoretical. Occasionally, use is made of descriptive cartel studies and concrete cartel cases, but these are primarily meant for

illustrative purposes. Like with any other theory, the results derived are based on a limited set of implicit and explicit assumptions. Consequently, conclusions are not guaranteed to hold if any of the assumptions are violated. The theories that are discussed and developed make extensive use of concepts that are commonly applied in the (classic) industrial economics literature. In particular, firms are thought of as single decision entities with profit maximization as their sole objective. In other words, undertakings are taken to be a black box and potential organizational issues, although very interesting in their own right, are beyond the scope of this thesis.

Following the industrial economics literature, the technical parts of this thesis uses game theory. Game theory is a discipline that takes a mathematical approach to study situations of conflict and was originally developed to analyze parlor games more rigorously. However, there appears to be many similarities between competition in the market place and ordinary games like chess, checkers, monopoly, *etcetera*. For example, in both situations there is a limited number of rivals that aims to achieve a particular result, while taking into account the rules of the game. Simply put, a cartel can be viewed a coalition among cheating players who violate the rules of competition to their own advantage.

The games that are analyzed in this thesis are noncooperative games. The main difference between noncooperative games and cooperative games is that with the latter players can form binding agreements. As we have described above, cartel arrangements are typically illegal and, as a result, a cartel agreement between firms is not binding, at least not from a legal perspective. It is therefore natural to take a noncooperative game approach to study cartels. An important solution of these type of games is the Nash equilibrium. A particular outcome of a game is a Nash equilibrium when each player maximizes his payoff given the strategies chosen by the other players. Thus, given the strategies adopted by all rivals, none of the players has an incentive to change its own strategy. Hence, a cartel is a Nash equilibrium of a game when all cartel members *individually* prefer to abide by the cartel contract and all outsiders *independently* stick to their strategies. As a result, all solutions to the games that we analyze are self-enforcing.

Game theory is well-suited to analyze the core topics of this thesis, but taking this approach has some important implications. Throughout this dissertation it is assumed that firms are perfectly logical players that are solely interested in maximizing their own profits. Therefore, ‘winning the game’ in this context does not mean “to beat all your opponents”, but it refers to achieving as much personal gain as possible. The goal of firms is to maximize their own profits and, in pursuing this objective, they are supposed to take into account the strategies adopted by competitors. Firms are therefore assumed to behave purely in their self-interest and, for instance, will not refuse to take part in a cartel agreement on moral grounds, e.g., because it is against the law. That is to say, if the strategy “colluding” dominates “competing” a firm will take part in the cartel.

The assumption of perfect rationality is arguably unrealistic. However, the degree of realism in the assumptions made is only of modest importance. By definition, every model is a simplified representation of reality, but this does not mean that models are useless. For example, a city map undoubtedly leaves out many interesting details. Yet, the map may be sufficiently realistic to find your way in town and if that is your

goal, the map is useful. What matters, therefore, is if the model is *sufficiently* realistic in the sense that it yields insights that are useful. In this respect, the rationality assumption is probably far less restrictive for business firms than for individuals. Firms, unlike individual persons, take decisions that are the result of interactions between individuals working within the firm. Arguably, this will neutralize a large part of irrational or emotional driven decisions of individuals. This is not to say that this approach allows us to perfectly predict firm behavior, but it should give us sufficient confidence that the rationality assumption is workable.⁹

Collusive market behavior has been of interest to economists ever since the start of industrial organization as a distinct discipline in economics. The theory of industrial organization developed almost hand in hand with the theory of imperfect competition, which is rooted in the 1930s.¹⁰ The theory of imperfect competition and, in particular, the theory of oligopoly was generally felt to describe market competition more accurately than the classic theories of perfect competition and monopoly.¹¹ Like many real-world markets, collusion is not easily explained with the traditional models of perfect competition and monopoly. The theory of monopoly trivially excludes the possibility of collusion, while the theory of perfect competition supposes all market players to be *price takers*. Clearly, it is difficult to see how firms can fix prices if the price decision is assumed exogenous. Hence, to study cartels properly requires a setting in which sellers have some market power.

Traditionally, cartels are believed to be potentially viable only in markets with only a few sellers. As a result, collusion is commonly studied in oligopolistic models. At the heart of the theory of oligopoly lies the hypothesis that in industries consisting of only a limited number of sellers, firms will realize their mutual interdependence. Hence, a strategic action by one firm not only has an impact on its own profit level, but will also affect the profits earned by others. With only a small number of undertakings this mutual recognition is thought to be conducive to a coordination of actions, which in the extreme case may lead to monopolistic market performance. Or as Chamberlin (1933) has put it,

“If each seeks his maximum profit rationally and intelligently, he will realize that when there are only two or a few sellers his own move has a considerable effect upon his competitors, and that this makes it idle to suppose that they will accept without retaliation the losses he forces upon them. Since the result of a cut by any one is inevitably to decrease his own profits, no one will cut, and although the sellers are entirely independent, the equilibrium result is the same as though there were a monopolistic agreement between them.”¹²

⁹For an in-depth and extensive discussion of the rationality assumption in economics the reader is referred to Smith (2008).

¹⁰Theories of Imperfect Competition have as their subject markets with more than one seller that are not perfectly competitive. Theories of Industrial Organization are concerned with the strategic behavior of firms and how this affects the workings of markets and *vice versa*.

¹¹A particular example of theories of imperfect competition are theories of oligopoly, which focus on markets with only a limited number of sellers.

¹²See Chamberlin (1933), p. 48.

The *monopolistic agreement* Chamberlin refers to is an implicit contract, which cannot be enforced by legal means. However, due to the illegality of cartel arrangements, explicit cartel contracts cannot be enforced by law either. This implies that any successful cartel must be self-enforcing and it is widely believed that firms will have a great many difficulties in forming an effective cartel. At least since Stigler (1964), economists are well aware that one of the most prominent threats to a cartel is the incentive of firms to chisel on the arrangement. To take account of the incentive to cheat, the vast majority of theories on collusion uses a dynamic approach to study cartels. Indeed, modern theory on collusion is based on so-called supergames. In a supergame the one-shot or stage game is played multiple times. A popular solution concept in these type of games is the Subgame Perfect Nash Equilibrium (SPNE). A SPNE requires players' strategies to constitute a Nash equilibrium in every subgame of the supergame.

The theory of incomplete cartels that is developed in this thesis also takes a supergame approach. In a supergame, market players are assumed to believe that interaction will take place for multiple periods and to sustain some level of collusion it is typically required that firms believe the game has no end date or that the end date is unknown. Friedman (1971) was among the first to show that, when interaction occurs for an infinite number of periods, firms can sustain some level of collusion if they are sufficiently patient.¹³ In these type of settings, collusion can be sustained when firms adopt some credible punishment strategy. One of the main shortcomings of this approach is that it typically fails to distinguish between tacit and overt collusion. It therefore remains unclear how sellers coordinate actions to select a particular equilibrium.¹⁴

The oligopoly models that are analyzed in the first part of the thesis are known as 'representative consumer models' or aggregate demand models. These type of models do not consider the behavior of individual customers, but simply assume there exists some (aggregate) demand for a certain product. In the second part of this dissertation, we take a different route and develop a spatial model of imperfect competition. In a spatial (or location) model, firms as well as customers are characterized by their location. Hence, products of firms are typically imperfect substitutes due to their geographical dimension. The literature that takes this approach dates back at least as far as Hotelling (1929). It is quite generally assumed that differentiation exists in a

¹³In fact, any individual rational outcome can arise in an infinitely repeated game given that players are sufficiently patient. In the literature this is sometimes referred to as the 'Folk Theorem'. See, for instance, Tirole (1988).

¹⁴Basically, there exist two strands of literature that attempt to solve the coordination problem. On the one hand, there is the literature on 'cheap talk', i.e., communication that does not directly affect the pay-offs, which deals with the question whether or not communication can influence equilibrium outcomes. For an overview, the reader is referred to Farrell and Rabin (1996). On the other hand, there is a literature that asks how revealing private information may affect the equilibrium outcome. See, for instance, Athey and Bagwell (2001). See also Compte (1998) and Kandori and Matsushima (1998) who study communications in settings in which sellers receive private but imperfect signals about past play.

bounded one-dimensional world, i.e., sellers compete on a line or circle.¹⁵ By contrast, we develop a two-dimensional spatial model.

1.4 Outline

As mentioned, this dissertation consists of two parts, both of which contain two chapters. The first part (chapter 2 and 3) analyzes the nature of incomplete cartels. The second part (chapter 4 and 5) focuses on economic methods of cartel detection, with a special emphasis on the detection of incomplete cartels.

Chapter 2 discusses the economics of incomplete cartels. It provides a survey of the relevant economic literature and some questions are distilled to which the literature provides no satisfactory answer. The major part of the chapter deals with three main issues concerning cartels that are not all-inclusive and these are discussed in reference to five oligopoly models. First, we examine under what conditions incomplete cartels are profitable. This question is of interest, because the competitive fringe could potentially undercut the cartel price and attract a significant number of customers, which might render an incomplete cartel unprofitable. Second, we ask when incomplete cartels are sustainable. In particular, we ask whether or not collusion is more likely to be sustainable when more firms are included in the conspiracy. The third issue concerns the incentives to take part in the cartel or to remain a fringe member instead. In the remainder of the chapter, we briefly discuss some cartel formation games with externalities and survey some theoretical contributions that study incomplete cartels with heterogeneous firms. Finally, we consider incomplete collusion in auctions. The chapter concludes with a discussion, which paves the way for the analysis in Chapter 3 by listing omissions and potential extensions of the existing literature.

Chapter 3 builds on the previous chapter and develops a novel theory of incomplete cartels. The main goal of this chapter is to provide an answer to the first three research questions as formulated above. To that end, we develop a price setting supergame in which firms differ in terms of production capacity, which is taken as a proxy for firm size. In this setting, we first explore what is the optimal cartel size. We find that the optimal cartel size is all-inclusive when colluding is costless, but less than all-inclusive when colluding is costly and the smallest firm in the industry is sufficiently small. Then, we explore what type of firms have a stronger incentive to collude. It is shown that larger firms are more inclined to join a cartel. In particular, we show that sufficiently small firms have no incentive to join any cartel. Moreover, a cartel comprising the largest producers is proven to be a subgame perfect equilibrium of the game. In addition, we examine whether or not firms have an incentive to form the most profitable cartel and find that the answer is in the positive when its smallest member is sufficiently large. It is noteworthy that these results find considerable support in examples of real-world incomplete cartels. Finally, we discuss changes in the size distribution of firms, for instance, due to a merger between two or more companies. We

¹⁵Hotelling (1929) analyzes competition in a ‘linear city’. Competition in a ‘circular city’ has been studied by, for example, Salop (1979).

show that firms have an incentive to merge only if they are in the cartel or become part of the cartel post-merger. Particularly, our results suggest that the most severe coordinated effects may come from mergers involving moderate-sized firms.

In the second part of this dissertation, we redirect attention to cartel detection and antitrust law enforcement. The main aim of Chapter 4 is to provide an overview of economic methods of cartel detection and to explore its potential. As a start, the goal and scope of cartel detection are discussed. In particular, we make the case that economics is likely to play an increasingly important role in cartel detection and that cartel detection itself will become a key instrument in antitrust enforcement. Then, we survey the economic literature on cartel detection and list some potential pitfalls in the use of these methods. Several detection methods work on the premise that the cartel does not encompass all firms. These techniques typically attempt to discriminate between cartel and fringe behavior so as to establish the existence of collusive conduct. It is shown that these methods might fail to delineate competition from collusion, because it can be advantageous for fringe firms to closely follow the cartel policy. Finally, we illustrate that an economic method of detection is vulnerable when it fails to take into account the idiosyncrasies of the industry to which it is applied. As a result, we argue that substantial progress can be made through the design of economic detection techniques that are tailored to a certain (type of) industry. An example of such a detection tool is presented in Chapter 5.

In Chapter 5 a novel detection method is designed that can be used to screen markets in which firms apply a so-called basing-point system. Historically, these type of industries are well-known to be prone to collusion. Examples include markets for lumber, iron, steel and cement. The detection test requires information on customer project locations and transaction data, i.e., prices and quantities. This information is shown to be sufficient to recover base-point locations from which the delivered prices were calculated. Base-point locations are useful to determine the likelihood of collusion. In a theoretical framework, we establish that in equilibrium all firms use a mill location as basing-point, whereas a collusive base is always located sufficiently far from the production centers. In particular, basing-points situated relatively close together and significantly far from mill locations are indicative of collusion. The likelihood of collusion is captured with a measure that takes a value between zero and one, with higher values corresponding to a higher likelihood of collusion. A software is developed in order to be able to deal with large data sets as well as with noise in the data. Finally, it is noteworthy that basing-point pricing is especially well-suited to facilitate collusion among a subset of firms that are located relatively close together. There exist some real-world examples of incomplete cartels that applied single basing-point pricing to protect local markets against distant competitors.

In the final chapter, Chapter 6, we summarize and discuss the main findings of this thesis. We further discuss implications for antitrust policy and outline some avenues for future research.

2

The Economics of Incomplete Cartels

“Where is the rest of me?” – Kings Row¹

2.1 Introduction

This chapter provides an overview and discussion of economic theories of incomplete cartels. There exists an abundance of literature on industrial collusion, but theoretical analyses of incomplete cartels are relatively scarce and infrequent. A possible explanation for the modest interest in incomplete cartels is that assuming an all-inclusive cartel typically greatly simplifies the analysis and taking account of cartel size is often not required to study the research question at hand properly. Also, incomplete cartels have never been a major topic of debate. Rather, in many studies, the issue of cartel size is touched upon only marginally and contributions that are explicitly concerned with incomplete cartels have been published irregularly. The main objective of this chapter is to survey and discuss this literature and to highlight some key issues with an eye on future research.

To begin, we might ask: What could be reasons for a cartel not to encompass all firms in the industry? The literature offers at least two main explanations, which both provide an interesting parallel with the most prominent causes of cartel failure. First, the full cartel could be unachievable for reasons that have to do with the internal organization of the cartel and second, cartel size is partly determined by factors external to the cartel.

¹Drake McHugh (Ronald Reagan) in the movie Kings Row (1942).

Prospective cartel members face an incentive problem and a coordination problem.² The incentive problem concerns the potential incentives of cartel members not to abide by the collusive arrangement, while the coordination problem is about reaching consensus on the content of the cartel contract. Real-world examples of cartels reveal that the incentive problem is non-trivial, but perhaps not as important as the coordination problem.³ Indeed, many studies suggest that bargaining problems and not secret cheating was actually the main cause of cartel failure.⁴ In fact, the failure to reach consensus may even temporarily lead to more severe market competition.⁵ These internal organization issues form a possible explanation for the existence of incomplete cartels. For instance, there might be too many producers in the market to reach an agreement among all parties.⁶ Arguably, it is easier to reach consensus in smaller groups, all else unchanged. Hence, when establishing the all-inclusive cartel appears to be impossible, there might still be an opportunity to install an effective cartel arrangement among a subset of sellers. Also, the cartel may initially be all-inclusive, but one or more members may have decided to secretly cheat on the agreement. Opportunistic behavior by cartel participants could result in an incomplete cartel when the remaining firms still find it in their interest to adhere to the collusive arrangement.

Whether or not a cartel is effective not only depends on these organizational matters, but also on the market environment in which the cartel operates. Indeed, the ability of a cartel to anticipate (a change in) market conditions is pivotal to its success. The effectiveness of a cartel arrangement will partly depend on factors like government interventions, technological advances and market entry. Thus, even when a cartel is able to solve organizational issues, it may not be successful due to external causes.⁷ Incomplete cartels may emerge because of structural changes in the market. For example, the cartel may initially encompass all firms, but its collusive gains could attract additional production.⁸ If new suppliers neither join the conspiracy nor lead to the

²See, for instance, Whinston (2006).

³The electrical-equipment conspiracy among twenty-nine U.S. manufacturers in the 1950s is a prominent example of a cartel in which participants persistently cheated. For a formal analysis of cheating on collusive agreements see, for instance, Slade (1990).

⁴See Levenstein and Suslow (2004), which compares many cartel studies and concludes that “About one quarter of the cartel episodes ended because of bargaining problems. Bargaining issues affected virtually every industry studied.” An experimental study conducted by Goppelsroeder (2008) also suggests that coordination problems can be significant.

⁵See Levenstein (1996) who uses the term ‘bargaining price wars’ to describe the price wars that sometimes follow conflicts in the cartel bargaining phase. Bargaining price wars are reported to have occurred for example in the Bromine and Tea industries. See Gupta (1997).

⁶In some instances it may be more accurate to talk about the number of decision makers instead of the number of firms. See Cyert *et al.* (1995).

⁷The nineteenth century cartel among U.S. salt producers is illustrative in this respect. Despite a very sophisticated organizational structure it was only modestly successful, because there were not sufficient barriers to entry. See Levenstein (1995). A similar conclusion is drawn by Clay and Troesken (2002) who analyze collusion in the market for distilled alcohol in the late nineteenth century. Also, as noted by Pindyck (1979), even if a cartel is able to solve organizational issues, it may not have sufficient market power to rise price well above the competitive level. For example, market power is limited if the demand curve is highly elastic.

⁸For instance, the successful mercury cartel in the 1950s and 1960s induced entry. See MacKie-Mason and Pindyck (1987).

demise of the cartel, then market entry results in a cartel with less than one hundred percent market share.

A first major issue to consider is then under what conditions incomplete cartels are viable. An incomplete cartel is viable when it enables participants to earn substantially more profits than they would have earned in absence of collusion.⁹ A second core issue is whether or not firms have an incentive to form a particular cartel, which we may loosely label ‘the participation problem’. This is especially relevant in a discussion of incomplete cartels, because then only a subset of firms should find it in their interest to collude. The participation problem further concerns questions about how many and which firms have an incentive to join the conspiracy.

In this chapter, these and related issues are analyzed from an empirical and a theoretical point of view. We first examine quite a few real-world examples of incomplete cartels and distill four ‘stylized facts’. The examples listed suggest that (i) cartels are often not all-inclusive, (ii) incomplete cartels tend to have a dominant position in the market, (iii) the market share of a cartel tends to decline over time, and (iv) incomplete cartels often comprise the larger firms in the industry.

We also address these issues in reference to five basic oligopoly models. The settings that we consider are:

- Simultaneous Bertrand competition with homogeneous products;
- Collusive price leadership;
- Simultaneous Bertrand competition with product differentiation;
- Simultaneous Cournot competition with homogeneous products;
- Collusive quantity leadership.

The choice for these models is not arbitrary. Except for the first, the vast majority of theoretical studies of incomplete cartels uses one of these settings. In fact, we show that no viable incomplete cartel exists in the first model. Nevertheless, it is very useful in illustrating the main concepts and it helps in clarifying the analyses of incomplete cartels in the other settings. The literature reveals that (i) viable incomplete cartels exist in all settings, but the first, provided that cartel members are sufficiently patient, (ii) it is unclear whether or not firms have an incentive to engage in an incomplete cartel in price setting games, (iii) In the simultaneous Cournot model, firms have an incentive to form an incomplete cartel with approximately 80-90% market share, and (iv) In a setting of collusive quantity leadership, firms will form an incomplete cartel with roughly 50% market share.

In the remainder of the chapter, we briefly discuss coalition formation with positive externalities, theories of incomplete cartels with heterogeneous firms and incomplete bidding rings. The five oligopoly models are useful to explore the incentives of firms to form an incomplete cartel. However, more often than not, the actual cartel formation

⁹In principle, a profitable cartel may have negative earnings. For example, prices may fall even in the presence of a cartel when demand is also declining.

process is not considered. There is a separate strand of literature that deals with the actual cartel formation process. In addition, theoretical contributions on incomplete cartels typically assume identical firms. Studies that assume heterogeneous firms allow to address the question of what type of firms have a stronger incentive to join a cartel. Finally, quite a few incomplete cartels have been discovered in auctions. We briefly survey the main theoretical contributions on incomplete bidding rings.

This chapter proceeds as follows. In Section 2, we present some real-world examples of incomplete cartels and distill some ‘stylized facts’. Section 3 lays out the basics of cartel theory and the standard model that is presented is used as benchmark in subsequent sections. Section 4 and 5 discusses the profitability and sustainability of incomplete cartels respectively. The sustainability of incomplete cartels is studied in the context of an infinitely repeated game. In Section 6 and 7 we consider the participation problem. In contrast to analyses that are concerned with the incentive problem of incomplete cartels, the participation problem is typically studied in static models. In the five oligopoly settings, Section 6 explores the incentives of firms to participate in an incomplete cartel. Section 7 discusses cartel formation games with externalities. Theories of incomplete cartels with firm heterogeneity are surveyed in Section 8. Section 9 discusses economic research on incomplete bidding rings. Section 10 concludes with a brief discussion and some unresolved issues regarding theories of incomplete cartels are pointed out.

2.2 Incomplete Cartels in Practice

Due to the secret nature of many cartels, cartel research necessarily relies quite heavily on, as Ronald Coase once phrased it, “blackboard economics”.¹⁰ This is not to say that we should limit ourselves to a theoretical discourse. In order to arrive at a better understanding of incomplete cartels, we believe it is instructive to first examine some real-world examples. It must be noted that cartel data are only scarcely available. One particular problem is that, more often than not, no information exists about the (combined) market share of cartel members so that the inclusivity of a cartel is unknown. Still, some illustrative evidence of incomplete cartels can be found in both descriptive cartel studies and antitrust cases. We briefly discuss our main findings in this section.

First, however, a word of caution is in order. Most of the evidence presented is, directly or indirectly, derived from antitrust cases. It is not unthinkable and even quite probable that this sample is biased. For instance, some industries have been under closer scrutiny for political reasons, which is likely to have increased the probability of discovery in these sectors substantially. We therefore do not know to what extent the results provide a reliable picture of incomplete cartels. This may lead some to conclude that the use of cartel data is misleading. Be that as it may, it will arguably be equally misleading to completely ignore this empirical knowledge. After all, antitrust cases are the prime source of information on cartels that operated “beyond the boundaries

¹⁰Coase (1991), p. 5.

of a blackboard”. Moreover, we can take some comfort in the fact that the empirical evidence that was found appears to point in the same direction.

Cartels have a very long history. They date back at least as far as one thousand years before Christ. In ancient times, like today, cartels did not necessarily encompass all sellers. For example, the Phoenicians, which are among the first merchants recorded in history, maintained their market power by establishing cartel arrangements in which Greek rivals took no part. Piotrowski (1933, p. 87) writes:

“The towns of Phoenicia, in particular Carthage, maintained during whole centuries their monopolistic position owing to the cartel contracts with the neighbors protecting their markets against Greek competition. These agreements had at the time excluded competition between the contracting parties themselves by strict division of the markets, exactly as is done by the district cartels to-day.”

Around the year 1900, quite a few historical cartel studies were published that report on less than all-inclusive cartels.¹¹ A recurring theme in these works is “the purpose of a cartel”. Most authors seem to agree that ‘full monopoly’ has been the ideal for many cartels, but, at the same time, it is concluded that the overwhelming majority of cartels had never succeeded to attain this ideal.¹² However, it was generally felt that the monopoly position was not required for a cartel to be effective. Controlling or monopolizing the market is not equivalent to having a ‘full monopoly’, that is. These studies suggest that cartels do not necessarily have to encompass all firms in the industry as long as the coalition has sufficient power to control the market and raise industry prices well above competitive rates. Broadly speaking, cartels seem not to be too worried about outside production as long as the competitive pressure is limited.

Yet, one particular problem for less than all-inclusive cartels is that outsiders often have an incentive to expand their output levels. For example, Genesove and Mullin (2001) reports with respect to the sugar cartel in the 1930s that the sugar market was bigger than the national market and that foreign suppliers increased their exports to the U.S. once the national sugar cartel was installed. Fringe members that are minor players initially may therefore become formidable competitors over time and ultimately result in a cartel breakdown. This problem is particularly severe when outsiders are unaware of the cartel arrangement. When the competitive fringe silently cooperates with the cartel, it is not unthinkable that fringe investments remain limited. If, however, the incentive to expand production capacity is too strong this ultimately could lead to the termination of the cartel.

The global incomplete cartels *Vitamin C* and *Citric acid* are illustrative in this respect. Both cartels were confronted with Chinese non-participants that expanded

¹¹ A detailed description of all these works is beyond the scope of this thesis. For a detailed discussion and references the reader is referred to Piotrowski (1933) and Liefmann (1977).

¹² See, for example, Grunzel (1902).

their production and these producers captured an ever bigger share of the market.¹³ Ultimately, the Chinese rivals undermined the stability of both cartels. For example, the case description of *Citric acid* reads,

“During the second period, from mid-1993 until the ending of the cartel in May 1995, it became increasingly difficult for the participating companies to sustain the price levels, in no small measure due to a dramatic increase of citric acid imports from China, particularly into the European market. Accusations of cheating on the agreement, especially against Jungbunzlauer, became rife and the level of trust between cartel members deteriorated.”¹⁴

and,

“The falling market share of the cartel members was also a matter of concern. From 1991 to 1993 the cartel members’ world market share in terms of total sales had fallen from around 70% to less than 60% and continued to fall to 52% in 1994. This continuous decline meant that the size of the ‘pie’ being shared out between the companies in the cartel was steadily decreasing, a factor that led to increased tension between them.”¹⁵

There were instances in which the cartel faced severe competitive pressure from the start. An illustrative example is a case called *District heating pipes*. This Danish incomplete cartel dominated the European market for district heating systems, but saw itself confronted with one major competitor, Powerpipe in Sweden. According to the European Commission, the cartel systematically attempted to drive Powerpipe out of the market after it refused to join the cartel. In particular, a collective boycott was organized after Powerpipe had won a major project in Germany. The idea behind the boycott was to prevent Powerpipe from getting essential supplies. Yet, this strategy appeared unsuccessful and as a consequence the cartel lost some valuable projects to Powerpipe.

“Unlike some other smaller producers, Powerpipe not only rejected pressure to join the club: it incurred the wrath of the cartel by systematically underbidding the favourite [the member designated by the cartel to win a particular project] and winning a series of major projects in Germany.”¹⁶

¹³See, for example, de Roos (2001), who applies a structural dynamic model to the *Vitamin C* cartel. In particular, his analysis suggests that a cartel will persist only if fringe competitors remain small.

¹⁴See case description at (91). *Official Journal of the European Union*, L 239/18, 6.9.2002. Case No COMP/E-1/36 604.

¹⁵See case description at (118). *Official Journal of the European Union*, L 239/18, 6.9.2002. Case No COMP/E-1/36 604.

¹⁶See *Competition Policy Newsletter*, number 1, February, 1999, 27-28.

In general, it has been quite common for cartels that attempted to monopolize the market to take reprisals against unwilling outsiders indirectly by influencing the environment of the competitive fringe. A prime example are the so-called “exclusive trading” clauses. The key element of this method is to either cut off supply (e.g., raw materials), demand (customers) or both.¹⁷ Direct purchasers are forced to buy exclusively from cartel members and raw material suppliers are urged to deliver exclusively to the cartel. Such contracts are typically enforced by means of a boycott. The cartel is simply not selling or buying when sellers or buyers do not respect the clause. In a few instances a modified version is used by asking for a so-called ‘loyalty rebate’. Purchasers that do not respect the contract are charged a higher price. The additional earnings are then divided among those who do respect the contract. Evidently, these methods only work if the cartel is sufficiently powerful.¹⁸ Notice that these type of methods secure stability of the cartel and, moreover, render entry of new competition difficult.¹⁹

A cartel might also decide to take direct reprisals against fringe members. For instance, a cartel may express predatory behavior if it has sufficiently ‘deep pockets’. The central idea is to lower combined prices in the short-run until outside producers are knocked out. They are then given the option either to close down or to join the cartel. Another method is buying-up competing firms. Cartels may create common pools to finance the take-over and rival works are often closed down after buying them up. In some instances the cartel has better access to technology required for producing the products. One simple strategy to limit non-cartel supply is then not to share this technology with fringe members. This is what happened in the *sorbates* cartel and the *graphite electrodes* cartel.²⁰ In light of antitrust enforcement, however, it may be quite risky to take measures against outsiders. For instance, installing a boycott might induce non-cartel members to file a complaint with the antitrust agency and this may significantly increase the probability of discovery.

Occasionally, incomplete cartels are supported by legal institutions and the government. In 1890, for instance, a German bookseller’s cartel was challenged by a bookstore not in the ring. Among other things, this outsider had been subject to a boycott. Nevertheless, the court ruled in favor of the cartel and not only that, it also established a precedent against interference by nonconspirators.²¹ An example of governmental support for cartels are the Japanese export cartels, which were mandated by government decrees to extend to all firms in the industry.²² However, governmental support does not always guarantee effectiveness of the cartel. Pindyck and Rubinfeld (2001)

¹⁷An example is a cartel in seafood processors in the U.S. in the 1980s. The cartel faced external competition by Viking Seafood Co. of Malden. Although not successful, the conspirators responded by trying to cut off Viking’s supply of processed fish blocks, used to make fish sticks.

¹⁸For example, Feldenkirchen (1992) argues that, although German cartels in iron and steel had influence and often used it against opposing outsiders their success was limited.

¹⁹An illustrative example is a railroad cartel in the U.S. at the end of the nineteenth century, which in cooperation with the Standard Oil Company raised rivals’ costs. In particular, the cartel arrangement functioned as a significant barrier to entry. See Granitz and Klein (1996).

²⁰See Harrington (2006).

²¹See Kinghorn and Nielsen (2004).

²²See Dick (2004).

report on an incomplete cartel in the market for milk in the United States. The cartel is called the Northeast Interstate Dairy Compact and is exempt from antitrust laws. The cartel comprises the producers of milk in the Northeastern corner of the U.S. (New England), which together make agreements on the minimum wholesale price. Prices, however, are not much higher than the competitive milk prices, which is due to significant fringe supply from dairy farmers located in the surrounding states.

In addition, there exists some scarce evidence on what type of firms tend to take part in the cartel. The above discussion suggests that larger firms are more prone to collusion than smaller firms. This is confirmed by Asch and Seneca (1975) which analyzes a sample of 101 large manufacturing companies for the period 1958-1967. The sample consists of 51 convicted price-fixers and 50 randomly selected 'non-colluders'. The samples are used to determine what characteristics may distinguish collusive from non-collusive firms. It is important to note that the paper does not contain any information about the inclusivity of the cartels. Yet, the authors repeatedly find that low profit rates induce firms, independent of size, to collude, but that firms with high market shares have a stronger incentive to collude than firms with a relatively low market share.

All the above finds support in Hay and Kelly (1974), which is one of the most well-known empirical cartel studies. They analyzed a limited sample of 65 horizontal cartel cases dealt with by the Antitrust Division of the U.S. Department of Justice between 1963 and 1972. In 20 cases, the inclusivity of the cartel could not be determined. Of the remaining 45 cases, 32 dealt with a cartel that was less than all-inclusive. The incomplete cartels consisted on average of 10-11 participants. However, this average is somewhat misleading in the sense that 26 partial cartels consisted of less than ten firms. Information about the combined market share of cartel members was available in 20 cases.²³ The average market share is approximately equal to 88%.²⁴ The lowest combined market share that is reported equals 65% and the market share of the incomplete cartel exceeded 90% in 14 cases.²⁵

To receive an impression of the size distribution of firms in the industry we may use the widely applied CR4 market concentration ratio in combination with the number of firms in the industry. The CR4 is the percentage of the value of total sales accounted for by the four largest firms in an industry. Hay and Kelly (1974) were able to compute the CR4 in 22 cases in which the cartel is less than all-inclusive. The average CR4 for these 22 cases is approximately 73%. The number of firms in the market was known in only 14 of these 22 cases. The average number of firms is around 12, while the average CR4 is 75%. This result suggests that market shares were unevenly spread in these industries. Hay and Kelly (1974, p. 21) further remarks that,

²³In 27 cases, data were available to determine the combined market share of the cartel, seven of which were estimated to be all-inclusive.

²⁴More precisely, the average is equal to 88.75%. However, in one of these 20 cases the market share was reported as "< 91", which is the reason we rounded the average to the lowest integer.

²⁵For the sake of completeness, the combined market share was between 70 and 79% in 2 cases and between 80% and 89% in 3 cases.

“...it is not necessary for a conspiracy to include all of the firms in the market to exist, or indeed, to be successful...In general, they [non-conspirators] seem to be the smallest competitors. It is difficult to believe that these non-conspirators were unaware of the conspiracy. It might be assumed that they were willing silent accomplices living under the price umbrella provided by the conspirators.”

In summary, the results listed above lend some support to the view that:

- Cartels are often incomplete;
- Incomplete cartels tend to dominate the industry in which they operate;
- Incomplete cartels tend to lose market share over time;
- Incomplete cartels often comprise the larger firms in the industry.

2.3 Foundations of Cartel Theory

In this section, we lay out the basics of cartel theory. The main objective is to introduce two key concepts that play a central role in a theoretical discussion of incomplete cartels: The incentive constraint and the participation constraint. A cartel is viable only when none of its members has an incentive to defect from the agreement. Hence, the incentive constraint is important to assess whether or not a particular cartel contract is feasible. The participation constraint indicates whether or not firms have an incentive to form a particular cartel. In the following, we closely follow the literature by analyzing the incentive problem and participation problem separately, because there exists hardly no contributions that take both problems into account. Moreover, discussing both concepts separately is helpful in identifying the main issues that are involved in theories of incomplete cartels. After we have introduced these two concepts, we briefly discuss why cartels are considered bad.

To begin, we might ask: what drives firms to engage in a cartel? To explore this issue, consider a single market with three identical and perfectly informed sellers that compete in price and produce a homogeneous commodity. Suppose that unit costs are constant and that market demand is a linear decreasing function of price.²⁶ This setting is visualized in the following diagram.

In Figure 2.1, price and cost are depicted as the dependent variable and the x -axis represents total market output Q . Demand and corresponding marginal revenue curve are respectively denoted by D and MR . In this setting, the average and marginal cost curve coincide, which is indicated by the solid horizontal line $AC = MC$.

Competition between the three rivals exerts a downward pressure on prices and margins. Consequently, the competitive Nash equilibrium is such that every firm sets a price equal to marginal cost and none of the competitors is making economic profits.

²⁶For simplicity and without loss of generality, we assume there are no fixed costs.

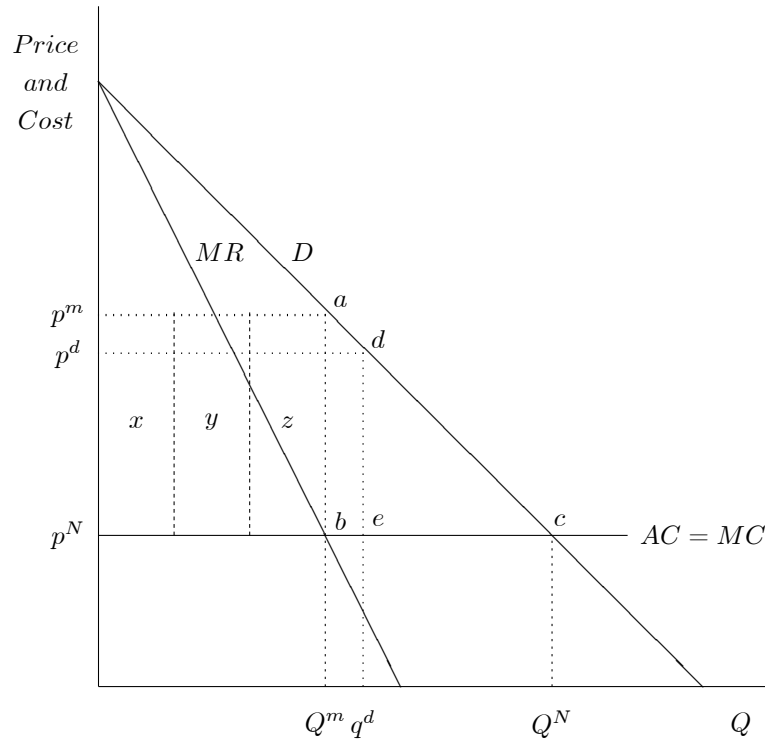


Figure 2.1 Incentives to collude illustrated.

The competitive market price is indicated by p^N . Demand at p^N is denoted by Q^N and we may safely assume that every seller produces one third of the competitive output. By comparison, a single supplier would optimally produce Q^m , at the point where MR intersects with MC , and sell these products at the monopoly price p^m . As a result, a profit-maximizing monopolist would make economic profits. The monopoly profits are indicated by the area $p^m abp^N$ in Figure 2.1.

In this setting, sellers can improve their situation by forming a cartel. The most profitable cartel arrangement is one in which firms mimic a monopolist. That is to say, cartel profits are maximal when the three sellers operate as if they are a single multiplant undertaking. To achieve this market outcome it is required that firms reduce their production levels in such a way that cumulative outputs are equal to Q^m . The three undertakings might therefore fix the price at p^m and/or install output quota so that total cartel outputs equal Q^m . Such an all-inclusive hard-core cartel arrangement will yield total cartel profits equal to monopoly profits indicated by the area $p^m abp^N$. The profits per cartel member are given by areas x, y and z in Figure 2.1 under the assumption that firms receive an equal share of total profits. Observe that cartel profits are higher than competitive profits for all firms, which illustrates the persistent incentive to collude.

2.3.1 The Incentive Constraint

Based on the analysis so far, it may appear that cartels are likely to be the rule rather than the exception. After all, every firm is better off than in competition so that forming a cartel agreement should not be too difficult. Yet, even in this simple framework, a cartel arrangement is all but trivial. The reason is that cartel participants have a strong incentive not to abide by the agreement. To see this, consider the perfect cartel in Figure 2.1. At the cartel price p^m , the three firms all produce $1/3Q^m$. Note that at this point marginal revenue exceeds marginal cost and, as a result, all participants have an incentive to increase their production level.

To illustrate, suppose one member lowers its price slightly to p^d . This allows the chiseling firm to increase its production level to q^d , given that the other members still adhere to the agreement. This deviating strategy would yield profits equal to the area $p^d dep^N$, which is clearly larger than any of the areas x, y, z . It is noteworthy that the incentive to defect is positively correlated with the number of participants. The reason is that maximum industry profit, i.e., the square $p^m abp^N$, is independent of the number of firms, which implies that every member receives a smaller share of cartel profits the larger is the number of participants. All-inclusive cartels are therefore, *ceteris paribus*, more likely to emerge in concentrated industries.

A major challenge for firms is then to form a cartel that is sustainable and often extra measures are required to make a cartel agreement effective. For a cartel to be sustainable, none of the cartel participants should find it profitable to change its strategy given the strategies adopted by the other sellers. However, as illustrated in Figure 2.1, every member typically has an incentive to lower its price and increase its production under the assumption that fellow members abide by the cartel contract. The problem of collusion is therefore very much akin to the prisoner's dilemma. That is to say, any collusive strategy is strictly dominated by some alternative strategy available to firms. Consequently, the cartel is not viable in the static model discussed above.

Yet, cooperation might be sustainable if the players believe that the game goes on forever, i.e., if the one-shot game is repeated for an infinite number of periods. Collusive outcomes can be commonly achieved by means of punishment strategies. The essential difference with the static setting is that in the repeated version of the game cheating might come with a price. Even though cheating may still be profitable in the short-run, it could cause the cartel to dissolve for a number of periods when fellow members retaliate, which implies a loss. Clearly, deviating from the agreement is unattractive when future losses outweigh short term gains.

One particular punishment strategy that is often applied in theoretical studies of collusion is the so-called grim-trigger strategy. Simply put, this strategy prescribes that in the event of cheating all firms compete in all periods following the period of defection. The seminal work is due to Friedman (1971) who showed that, given that firms adopt 'trigger-strategies', there always exists a discount factor for which (full) collusion is sustainable. Formally, a collusive outcome can be supported as an equilibrium outcome of the game if the following condition holds,

$$\delta \geq \delta^* = \frac{\pi^d - \pi^c}{\pi^d - \pi^N}. \quad (2.1)$$

Competitive and collusive profits per firm are respectively given by π^N and π^c . The one-period maximum profit that can be obtained by a chiseling firm is denoted π^d and $\delta \in (0, 1)$ is the actual common discount factor used to value future income. Cartel participants will find it profitable to adhere to the cartel agreement when the actual discount factor δ (weakly) exceeds the critical discount factor δ^* . In the literature this condition is referred to as the incentive compatibility constraint (ICC).

It is important to realize that in the homogeneous goods Bertrand game introduced above an incomplete cartel might be sustainable even though all incomplete coalitions are unprofitable in this particular setting. That is, firms are indifferent between collusion and competition as they earn zero profits in both situations. By contrast, profitable cartels are sustainable in the context of an infinitely repeated game only under the assumption that cartel members are sufficiently patient.

2.3.2 The Participation Constraint

If the incentive constraint is satisfied for all cartel members, then none of the participants has an incentive to deviate from the cartel agreement. The incentive constraint therefore guarantees the stability of a cartel arrangement. However, it tells us nothing about the incentives of firms to take part in a cartel. For a particular coalition to form, it must be individually rational for every (potential) member to join that cartel. Likewise, no firm outside the coalition should be willing to join the coalition. This notion is captured by the ‘internal stability condition’ and ‘external stability condition’ as introduced originally by d’Aspremont *et al.* (1983).

Let π^c and π^o denote the profit of a cartel member and an outsider respectively. Formally, a firm has no incentive to leave a cartel formed by k firms when the following condition holds,

$$\pi^c(k) \geq \pi^o(k-1). \quad (2.2)$$

In a similar fashion, no fringe member has an incentive to join a cartel formed by k firms when,

$$\pi^o(k) \geq \pi^c(k+1). \quad (2.3)$$

If both conditions hold simultaneously, then a cartel is called *stable*. In what follows, we refer to (2.2) in conjunction with (2.3) as the *participation constraint*.²⁷

In the example presented above and ignoring the special case of a cartel formed by a single firm, the participation constraint only holds for a cartel comprising all three firms. Notice that the full cartel is trivially externally stable, because there are

²⁷Note that with these conditions, a firm only considers the changes after it has made its choice. A more sophisticated notion of this participation constraint, when firms are farsighted and take account of all participation choices of other firms, has been analyzed in Diamantoudi (2005). See also Thoron (1998) which studies coalition-proof stable cartels.

no outsiders. The all-inclusive cartel is internally stable if $\pi^c(3) \geq \pi^o(2) = 0$, which always holds. Firms therefore do have an incentive to form the full cartel in this setting. Moreover, the full cartel was shown to be viable when firms are sufficiently patient. Hence, in the Bertrand homogeneous goods example, the all-inclusive cartel is the only viable cartel and firms do have an incentive to engage in this cartel agreement.

2.3.3 *Why are Cartels Bad?*

Nowadays, the vast majority of capitalist societies has declared cartels illegal. A legitimate question to ask is, why? As Figure 2.1 reveals, an immediate consequence of a hard-core cartel is higher prices and lower outputs. An effective cartel arrangement therefore leads producers to earn more profits. Yet, this comes with a price. Customers with a sufficiently high willingness to pay have to spend a larger part of their budget to obtain the product. Additionally, some customers that would have bought the product in absence of collusion do not buy the product anymore. Cartels therefore cause a re-allocation of welfare from consumers to producers and a welfare loss for society at large. The latter is referred to as ‘allocative inefficiency’. The term used for inefficiencies associated with a misallocation of resources is ‘dead-weight welfare loss’. In Figure 2.1, the dead-weight welfare loss is indicated with the triangle *abc*.²⁸

In addition to allocative inefficiencies, cartels are likely to generate so-called X-inefficiencies.²⁹ Firms that shelter from competitive pressure too much are widely believed to lose the incentives to produce their products or services as efficient as possible. Not operating efficiently might also not be needed to survive, because supra-normal profits are more than sufficient to make up for higher costs. Moreover, depending on the cost structure, a cartel may lead to technical inefficiencies in the short run when the reduction of outputs yields an increase in unit costs. Note, however, that when unit costs are increasing in output a cartel may actually yield technical efficiencies.³⁰

Cartels are sometimes believed to cause dynamic inefficiencies. Firms that take part in a cartel are thought to have less incentives to innovate, because there is less need to improve the position in the market. However, there are two counterarguments to the idea that collusion is likely to result in fewer product and process innovations. First, a cartel member has typically more means available to innovate, because more profits are earned per period.³¹ The additional collusive gains may in principle allow for more R&D expenditures. Second, when innovative activities are successful, a firm may benefit more from its inventions when competition is not too fierce. For example, additional gains due to R&D efforts can be limited when rivals can relatively easy implement the product or process innovation. Thus, collusion might enhance dynamic efficiency. It is unclear which of these opposing forces dominates and the economic literature is inconclusive regarding this issue.

The economic damage caused by cartels is substantial. In real-world examples, it has been estimated that modern international cartels were responsible for an average

²⁸Note that such ‘allocative inefficiency’ is absent when demand is perfectly inelastic.

²⁹See Leibenstein (1966).

³⁰I thank Jan Tuinstra for pointing this out.

³¹Schumpeter (1912) argues that a monopolist has stronger incentives to invest in R&D.

price increase of around 28%, but it might well be substantially higher.³² The graphite electrode cartel, for instance, raised prices by more than 60% in the U.S. and affected 1.7 billion dollars in U.S. commerce alone. More generally, empirical studies suggest it is not uncommon for cartels to increase prices by more than 20%.³³ Yet, with a lower price increase, economic harm can still be significant. For example, in the well-known international lysine cartel, U.S. customers paid on average approximately 17% more than what they would have paid in absence of the cartel. This corresponds to an overcharge of more than 75 million dollars in the United States and 200 million dollars worldwide.³⁴ Another example is the vitamin cartel which made customers in the United States pay an additional amount of at least 1.2 billion dollars.³⁵ Baker (2003) provides plenty of other examples. Moreover, he loosely estimates total costs to the U.S. economy of exercising market power to exceed 100 billion dollars annually. It is unclear, however, what share can be attributed to cartel practices.

In addition, cartels may cause other types of harm. In the year 2001, a detailed study of the global price-fixing cartel *lysine* by John M. Connor was published. It is titled: *Global Price-Fixing: "Our Customers are the Enemy"*. The phrase "our customers are the enemy" is a quote from one of the participating managers who were secretly taped by an FBI undercover agent during one of their cartel meetings. This statement reflects the attitude a cartel might have towards direct purchasers of their products. More generally, the title suggests that cartels can have consequences that are not merely economic. Indeed, it could be argued that firms ought not to steal from customers and society. Competition in the market place is distinct from ultimate rivalry, because it takes place within legal boundaries. These 'rules of competition' are established in a more or less democratic fashion and as such can be viewed a fundamental part of the social contract. To put it differently, firms that operate in breach of competition laws behave contrary to what should be done, at least according to the majority of the people.

Cartels, like many other concentrations of power, might lead to abuse and other spin-offs that overall can be considered unwanted. The infamous international *DeBeers Diamonds* cartel is illustrative in this respect. This cartel formed in South Africa at the end of the nineteenth century and is arguably one of the most successful cartels in history.³⁶ The diamond producer DeBeers takes a leading role in this cartel and, in fact, imposes collusive behavior on the other members by use of coercive tactics. In the early 1980s, Zaire, which is the worlds largest supplier of industrial diamonds defected. Only two months later the market was flooded by industrial diamonds, which caused the price of Zairian diamonds to drop significantly. It is a public secret that DeBeers was responsible by supplying an enormous amount of stockpiled industrial diamonds. Of course, this is a very costly strategy to discipline fellow members, but also very effective. Ultimately, Zaire had to stop the fight, because its pockets were not deep enough to hold out against DeBeers. By 1983, Zaire was allowed to join the cartel

³²See Connor (2003).

³³See Bolotova (2006).

³⁴See Connor (2001).

³⁵See Connor (2001).

³⁶See Spar (1994) for a detailed discussion of the international Diamond Cartel.

again, but, not surprisingly, on less favorable terms. DeBeers clearly made its point and one executive posed the following question in a published interview: ‘Anyone want to follow Zaire?’³⁷

Part of the success of the cartel is, no doubt, due to the fact that the power of DeBeers reaches far beyond the diamond production. It is also involved in the distribution and marketing of diamonds. One of the diamond dealers is said to have said: “DeBeers is like the Mafia...but it is good for the trade”³⁸ Also, a broker vividly stated:³⁹

“The Syndicate [DeBeers] knows when you sneeze or take a leak. They have a spy system that would put the CIA to shame. They know everything there is to know about anyone of any significance in the diamond world.”

Clearly, cartels which are the result of a coercive sort of cooperation could easily lead to practices that are harmful in non-economic terms.

It must be noted, however, that there might be an exception to the rule that cartels are indeed bad. These are the so-called structural crisis cartels. During a business cycle downturn, competition potentially threatens the existence of some firms. One might argue that the duty of saving a company from bankruptcy and consequently protect jobs weighs more heavily than strict compliance with antitrust laws. The implication of a severe economic crisis forcing firms out of business is that fewer firms operate in the market once the economy recovers. The reduced competitive pressure may yield inefficiencies that potentially outweigh the cost of collusion during the crisis.⁴⁰

2.4 On the Profitability of Incomplete Cartels

In this section, we discuss the profitability of incomplete cartels. The key difference between a full cartel and an incomplete cartel is that the latter faces outside competition, which potentially limits the ability to charge a price above competitive levels. In the extreme, suppliers that do not take part in the conspiracy undermine the effectiveness of a cartel. To see this, consider the example depicted in Figure 2.1 above. Suppose, however, that two out of three sellers attempt to collude by fixing a price $p^c \in (c, p^m)$. For such a cartel to be profitable, the non-conspirator must charge a price $p^o \geq p^c$, because cartel demand would be zero with any other price. However, given p^c , the best response strategy of the outsider is typically to slightly undercut p^c . As a result, all customers will choose to buy from the fringe supplier. It follows

³⁷See Spar (1994)

³⁸See Spar (1994).

³⁹See Koskoff (1981).

⁴⁰The European Union and the U.S. have a different attitude towards structural crisis cartels. In the U.S. such cartels are prohibited and considered a *per se* violation of the Sherman Act. By contrast, in Europe structural crisis cartels are sometimes permitted. See, for example, Neumann (2001).

that there exists no profitable incomplete cartel in this simple setting.⁴¹ Therefore, the main question is under what conditions incomplete cartels are profitable.

To begin, we address the question: What are factors that affect the ability of a cartel to raise price above competitive levels? To that end, consider a given industry in which n identical firms produce a homogeneous product and simultaneously decide on their output level q . The profit function of firm i is given by,

$$\pi_i = (p(Q) - c)q_i, \quad (2.4)$$

with c being marginal production cost and $Q = \sum_{i=1}^n q_i$. In order to determine the profit-maximizing output level for firm i , we take the first-order derivative of (2.4) with respect to q_i and level to zero,

$$\frac{d\pi_i}{dq_i} = p - c + q_i p' \left(1 + \sum_{\substack{j=1 \\ j \neq i}}^n \frac{dq_j}{dq_i} \right) = 0. \quad (2.5)$$

Let $s_i = \frac{q_i}{Q}$ denote the market share of firm i and, for the sake of simplicity, we assume firms have equal market shares, i.e., $s_i = s = \frac{1}{n}$. Furthermore, let $\varepsilon = \frac{-p}{Qp'}$ be the price elasticity of demand and define $r_{ij} = \frac{dq_j}{dq_i}$ as the conjectural variation. The conjectural variation measures the expected change in output of firm j to a change in output of firm i . Using some basic calculations, we can rearrange (2.5) to,

$$L_i = \frac{p - c}{p} = \frac{s \left(1 + \sum_{\substack{j=1 \\ j \neq i}}^n r_{ij} \right)}{\varepsilon}, \quad (2.6)$$

which is the Lerner index of firm i . The Lerner index has an upper bound equal to 1 and higher values of L_i indicate more market power. Due to symmetry, all firms have a Lerner index with a similar structure, i.e., $L_i = L$. Observe that the price level (and mark-up) is positively correlated with the market share of a firm and negatively with the price elasticity of demand. The profitability of a firm further depends on beliefs about the strategy of competitors. For example, in Cournot competition, firm i does not expect the others to change outputs in response to a change in its own output level, i.e., $\sum_{\substack{j=1 \\ j \neq i}}^n r_{ij} = 0$. Hence, $L_i = \frac{s}{\varepsilon} = \frac{1}{n\varepsilon}$ when firms compete in quantity. By contrast, in Bertrand competition firm i expects a change in its production level to be exactly offset by its rivals, i.e., $\sum_{\substack{j=1 \\ j \neq i}}^n r_{ij} = -1$. As a result, $L_i = 0$ when firms compete in price.

Now consider the case of collusion. As in competition, the ability of cartel members to raise price above unit production cost depends on the elasticity of demand, the market share and beliefs about the strategies of rivals. Perfect collusion requires every firm to perfectly match each others output reduction, i.e., $\sum_{\substack{j=1 \\ j \neq i}}^n r_{ij} = n - 1$.

⁴¹This conclusion closely resembles the result that became known as the ‘Bertrand paradox’. See Tirole (1988) for a more extensive explanation and a formal proof.

Consequently, $L_i = \frac{ns}{\varepsilon} = \frac{1}{\varepsilon}$, which corresponds to the market power of a monopolist and is the maximum market power that can be obtained by a cartel. To see what happens when the cartel is not all-inclusive suppose that k firms form a hard-core cartel agreement and together reduce outputs. The market power of this incomplete cartel then depends on how the $n - k$ outsiders respond. If the outsiders do not perfectly follow the cartel policy the Lerner index of cartel members will be in the following interval $[0; \frac{1}{\varepsilon})$. For example, when the cartel believes fringe members will not adapt their output levels when participants lower their production then the Lerner index of a cartel member equals $L_i^c = \frac{k}{n\varepsilon}$. Clearly, the k cartel members have more market power than in competition, i.e., $\frac{k}{n\varepsilon} > \frac{1}{n\varepsilon}$ for $k > 1$. Yet, the market power is less than with a full cartel, i.e., $\frac{k}{n\varepsilon} < \frac{1}{\varepsilon}$ for $k < n$. It can be concluded that the ability to raise price above marginal cost positively depends on the number of firms that take part in the cartel, all else equal. Note, however, that when firms compete in price, outsiders perfectly compensate the output reduction of the cartel. As a result, no incomplete cartel is viable. If and to what extent an incomplete cartel is profitable therefore depends heavily on the assumptions one is working with.

Some empirical studies attempt to compute the Lerner index for cartels. Eckbo (1976) shows that in 19 out of 51 cartel agreements the cartel was able to raise price more than 200 percent above unit cost. Griffin (1989) analyzes 54 international cartels, 53 of which were less than all-inclusive. The average market share equals approximately 60%. For all 54 cartels, he computes the Lerner index. Among other things, he obtains that the highest Lerner index is equal to 0.80 for an incomplete cartel in the rubber industry that lasted from 1923-1928. The incomplete wheat cartel from 1933-1934 was computed to have the lowest Lerner index (-0.12). It must be noted, however, that this cartel was active in a period following the Great Depression of 1929 and therefore still may have been reasonable successful in relative terms.

The economic literature on collusion offers a variety of models that allow us to further explore the profitability of incomplete cartels. In the following, we briefly discuss four of the most prominent ones. All these theories are closely related to the basic setting laid out in the previous section. The difference lies in the fact that either one or two of the core assumptions are modified.

2.4.1 Collusive price leadership

Incomplete cartels are traditionally studied in a model of collusive price leadership.⁴² In this setting, it is assumed that a limited number of large firms engages in a price-fixing cartel. This cartel is supposed to operate as a price leader followed by a large number of small firms that take the cartel price as given.⁴³ Hence, the key difference with the basic setting described above is that in this model outsiders are considered

⁴²One of the first detailed discussions of collusive price leadership is Markham (1951). For a modern and advanced theoretical analysis see Deneckere and Kovenock (1992).

⁴³The asymmetry in firm size assumed in this model is commonly observed in oligopolistic markets. We can think of various reasons for why some firms might grow big. Examples include, more efficient production technologies, inventions that are temporarily protected by patents and advantageous geographic locations.

price takers. The competitive fringe is indeed competitive in the sense that profit maximization requires fringe firms to produce the quantity for which price equals marginal cost.

In order to explain the workings of this model consider the following diagram:

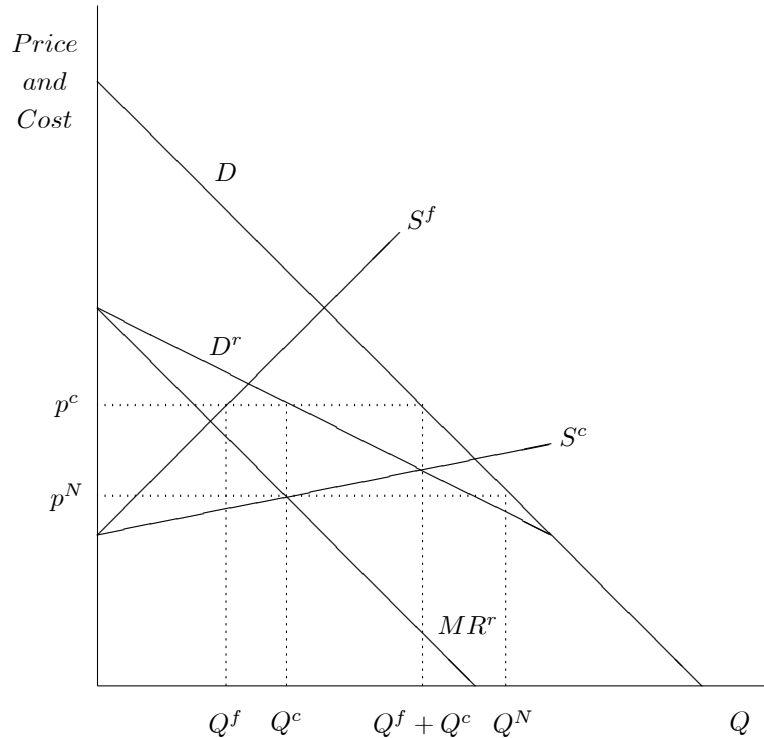


Figure 2.2 Collusive price leadership equilibrium.

The situation depicted in Figure 2.2 is very much similar to the situation sketched in Figure 2.1. An important difference with the previous model is that in the current setting it is assumed that marginal costs are not constant, but increasing in production. Prior to the cartel, all firms set price equal to marginal cost and determine the optimal output accordingly. Like before, the competitive output and price are respectively indicated by Q^N and p^N .⁴⁴

Suppose now that a subset of firms forms a price-fixing cartel, which attempts to maximize joint cartel profits. The cartel supply function is the horizontal sum of the marginal cost curves of members and indicated by S^c . In a similar fashion, one can construct the supply function of the competitive fringe, which is denoted S^f . The demand for cartel output depends both on industry demand and total fringe

⁴⁴For simplicity, the competitive supply curve is omitted. The market supply curve is the horizontal sum of firm's marginal cost curves and intersects the market demand curve D at point (p^N, Q^N) .

production. Residual cartel demand D^r is constructed by subtracting fringe supply from industry demand at any given price. The corresponding residual marginal revenue curve is denoted by MR^r .

The cartel maximizes joint profits by restricting outputs so that joint marginal costs of members equals MR^r . In Figure 2.2, the cartel optimally sets a price p^c and produces Q^c . As mentioned, the competitive fringe takes p^c as given and optimally produces total fringe output Q^f for which price equals marginal cost. Notice that in equilibrium total fringe supply plus cartel output equals market demand at the cartel price. Note further that, although price is the choice variable and products are homogeneous, an incomplete cartel is profitable when the cartel operates as a price leader. In particular, the assumption that fringe producers sell their products at the cartel price allows the cartel to increase prices without losing all its customers to producers not included in the price-fixing conspiracy. Finally, an interesting observation is that, due to the presence of outside competition, the residual demand (D^r) is more elastic than market demand (D). As a result, an incomplete cartel typically sets a lower price than an all-inclusive cartel would.

2.4.2 Differentiated goods

In the previous setting, the standard framework was modified by changing the cost structure and the order of moves. Alternatively, it may be assumed that products are less than perfect substitutes. With product heterogeneity, low price fringe members may not attract all customers and potentially leave part of the market to the cartel. As a result, incomplete cartels might be profitable when products are sufficiently different. In this section, we briefly discuss the setting discussed in Section 2.3 under the assumption that the price elasticity of demand is less than infinite.

Such a study has been conducted by Deneckere and Davidson (1985).⁴⁵ They consider a setting in which firms face the following Shubik-Levitan demand function,

$$q_i = a - p_i - \gamma \left(p_i - \frac{1}{n} \sum_{j=1}^n p_j \right). \quad (2.7)$$

Here, a is a demand scale parameter and n indicates the number of (symmetric) firms. The direct price effect and the cross-price effect are respectively given by $1 + \gamma(1 - \frac{1}{n})$ and by $\frac{\gamma}{n} \sum_{j \neq i} p_j$. The common parameter $\gamma \geq 0$ measures the degree of product differentiation. If $\gamma = 0$, goods are unrelated and collusion would make no sense. By contrast, if γ is large, goods are close substitutes and price competition becomes fierce.⁴⁶ We therefore naturally suppose γ to be neither very large nor very small.

For simplicity, we normalize marginal cost to zero, i.e., $c = 0$. Suppose that k firms engage in a price-fixing cartel with the objective to maximize joint profits. The profit-maximizing cartel price is equal to,

⁴⁵This paper is primarily concerned with mergers, but as the authors remark in a footnote the analysis could be equally well applied to cartels.

⁴⁶Perfect substitutability corresponds to $\gamma \rightarrow \infty$.

$$p^c = \frac{a[2n + \gamma(2n - 1)]}{4n + 2\gamma(3n - k - 1) + \gamma^2(\frac{n-k}{n})(2n + k - 2)}, \quad (2.8)$$

and independent outsiders optimally charge,

$$p^o = \frac{a[2n + \gamma(2n - k)]}{4n + 2\gamma(3n - k - 1) + \gamma^2(\frac{n-k}{n})(2n + k - 2)}. \quad (2.9)$$

Observe that $p^c > p^o$ for $k > 1$, which always holds. Hence, independent of the size of the coalition, a cartel member always sets a higher price than a fringe firm.⁴⁷ As a consequence, cartel members lose market share, i.e., outsiders become the larger firms in the industry.

Concomitant profits are respectively,

$$\pi^c = a^2 \left[\frac{2n + \gamma(2n - 1)}{4n + 2\gamma(3n - k - 1) + \gamma^2(\frac{n-k}{n})(2n + k - 2)} \right]^2 \left(1 + \gamma \frac{(n - k)}{n} \right), \quad (2.10)$$

and,

$$\pi^o = a^2 \left[\frac{2n + \gamma(2n - k)}{4n + 2\gamma(3n - k - 1) + \gamma^2(\frac{n-k}{n})(2n + k - 2)} \right]^2 \left(1 + \gamma \frac{(n - 1)}{n} \right). \quad (2.11)$$

It is easy to show that profits of cartel members are positively correlated with the number of participants, i.e., larger coalitions are more profitable than smaller ones. In particular, any incomplete cartel is profitable.

2.4.3 Quantity competition

A third way in which we can adapt the standard framework is by assuming that firms have output rather than price as their strategic variable. The key difference with the previous models is that in quantity setting games reaction functions are typically downward sloping, while these are often upward sloping in price setting games.⁴⁸ As we have seen in the Bertrand differentiated goods model, a price increase by cartel members leads to a price increase of fringe firms. In this particular setting, a price increase leads to an output reduction by both the cartel and outsiders. By contrast, when firms compete in quantity, an output reduction by cartel members leads to an *increase* of fringe output. Thus, given that the quantities supplied by firms function

⁴⁷In addition, note that both the cartel price and the fringe price are higher than the competitive price, $p^N = \frac{a}{2 + \gamma \frac{n-1}{n}}$.

⁴⁸Note, however, that under some conditions reaction functions may have upward sloping portions in quantity setting games even when products are substitutes. Likewise, reaction functions may in some settings not be upward sloping over the relevant range in price setting games. Yet, the positive slope of reaction functions in price setting games is felt to be natural for a broad class of situations. See, for example, Martin (2002) for an extensive discussion.

as strategic substitutes, an incomplete cartel will be profitable only when the output reduction of members is more than compensated for by a price increase. As before, we consider a situation in which firms make their decisions simultaneously and a situation in which the cartel moves first.

2.4.3.1 Simultaneous play

The question when incomplete collusion is profitable in a static Cournot setting with linear demand has been addressed by Salant *et al.* (1983). To explain their main findings, let N denote a set of n identical profit-maximizing firms, which choose their production level simultaneously. Products are homogeneous and produced at constant marginal and average cost equal to $c > 0$. Suppose that the inverse demand function is linear and given by,

$$p(Q) = a - bQ, \quad (2.12)$$

where p is the market price and $Q = \sum_{i=1}^N q_i$ is total market output. In order to guarantee positive production levels, we naturally suppose that $a > c$. In this symmetric setting, every firm earns the following competitive Cournot-Nash equilibrium profits,

$$\pi^N = \frac{(a - c)^2}{b(n + 1)^2}. \quad (2.13)$$

Now suppose that a number of k firms collude. Every cartel member then makes a profit equal to,

$$\pi^c = \frac{(a - c)^2}{bk(n - k + 2)^2}, \quad (2.14)$$

whereas independent outsiders act according to their best response function and earn,

$$\pi^o = \frac{(a - c)^2}{b(n - k + 2)^2}. \quad (2.15)$$

As can be observed, outsiders make more profits than in competition for all cartel sizes. In addition, a fringe member benefits more from collusion than a cartel member for all $k > 1$. Cartel members reduce their production levels and thereby create incentives for outsiders to increase their output. This is due to the assumption about quantities being strategic substitutes. As a result, the cartel creates a positive externality for firms that remain independent outsiders, i.e., fringe members enjoy the benefits of a higher market price, but do not incur the cost associated with a reduction of output. Furthermore, it is noteworthy that total industry profits always increase in the presence of a cartel.

A second key result of the analysis is that collusion is only (weakly) profitable if,

$$\frac{k}{n} \geq 1 + \frac{3 - \sqrt{4n + 5}}{2n}.$$

This implies that at least 80% of the firms in the industry must participate in the cartel to ensure that members are better off than in competition.⁴⁹ At first glance, this result may appear surprising, because in principle the multiplant firm is always able to copy the competitive strategy. This, however, is not an equilibrium following the cartel, because the best response strategy *given* unchanged outputs of outsiders is to reduce total cartel production. In fact, the possibility of lower profits for cartel members arises precisely because outsiders increase their output levels. The price increase resulting from the cartel may therefore not be sufficient to offset the loss in market share.⁵⁰ In sum, an incomplete cartel might be profitable if it takes a sufficiently dominant position in the market.

2.4.3.2 Sequential play

Shaffer (1995) and Martin (2002) study a model in which a cartel operates as a Stackelberg leader. To formalize, consider the same setting as discussed in the previous subsection. A cartel typically reduces outputs, which in equilibrium induces a fringe firm to adapt its output level according to the following best response function,

$$q^o = \frac{a - c - bQ^c}{b(n - k + 1)}, \quad (2.16)$$

where Q^c denotes total cartel output. Given that fringe firms act accordingly, the optimal output level of a cartel member equals,

$$q^c = \frac{a - c}{2bk}. \quad (2.17)$$

This yields the following per firm equilibrium profits,

$$\pi^c = \frac{(a - c)^2}{4bk(n - k + 1)}, \quad (2.18)$$

and,

$$\pi^o = \frac{(a - c)^2}{4b(n - k + 1)^2}. \quad (2.19)$$

To examine whether or not an incomplete cartel is profitable we compare (2.18) with the competitive Cournot profits as given by (2.13). Cartel members are (weakly) better off than in competition when the following condition holds,

$$(n + 1)^2 \geq 4k(n - k + 1). \quad (2.20)$$

⁴⁹This minimum is reached for $n = 5$ and $k = 4$, i.e., $\frac{k}{n} = \frac{4}{5} = 0.8$.

⁵⁰The original analysis of Salant *et al.* (1983) focuses on the profitability of mergers instead of cartels. However, as Perry and Porter (1985) rightly remarks, mergers are conceptually not well defined in this setting. The reason is that the new collusive equilibrium implicitly assumes that the merger does not differ from fringe firms, while in fact it is larger. Moreover, the number of firms will be lower post-merger.

The right-hand side of (2.20) reaches its maximum at $k = \frac{n+1}{2}$ and for this value (2.20) holds with equality. Hence, any cartel that operates as a Stackelberg leader in the presence of a Cournot fringe is profitable.

2.4.4 Comparison

At this point, it is useful to summarize and compare the findings discussed above. In this section, we considered five different oligopoly models and examined which incomplete cartels are profitable. The main conclusions are summarized in the following table.

Table 2.1: Profitable incomplete cartels.

Strategic variable	Order of moves	Homogeneous goods	Heterogeneous goods
Price	Simultaneous	none	all
	Sequential	all	-
Quantity	Simultaneous	dominant cartels	-
	Sequential	all	-

Clearly, whether or not an incomplete cartel is profitable very much depends on the assumptions one is working with. As Table 2.1 reveals, incomplete cartels are often profitable. In particular, in three out of five models, incomplete cartels of any size are profitable. This is not the case in simultaneous move games in which firms produce a homogeneous commodity. Under these assumptions, an incomplete cartel must take a dominant position in a market in which firms compete in quantity, while no incomplete coalition is profitable when firms compete in price.

2.5 On the Sustainability of Incomplete Cartels

In the previous section, it has been shown that incomplete cartels are profitable in a variety of settings. In this section, we explore the sustainability of these profitable coalitions. Recall that every cartel member has an incentive to defect from the agreement given that its fellow members abide by the cartel contract. Consequently, none of the profitable incomplete coalitions is sustainable in the one-shot version of the games discussed so far. As we have explained in Section 2.3, a cartel is potentially sustainable in an infinitely repeated version of the static games. A cartel can be sustained when the incentive compatibility constraint is satisfied for all participants, i.e., when cartel members are sufficiently patient. It is important to realize that the incentive constraint is not necessarily violated for unprofitable coalitions. For example, firms might not have an incentive to defect from an unprofitable cartel agreement in the simultaneous Bertrand game with homogeneous products, because deviating yields no additional gains. However, we are naturally interested in cartels that are profitable. As mentioned, all profitable cartels are sustainable in the context of an infinitely repeated game under the assumption that cartel members are sufficiently patient.

In the following, we discuss the infinitely repeated version of the remaining four models. In particular, we explore if the critical discount factor is increasing or decreasing in the number of cartel participants. This is of special interest, because if the critical discount factor is increasing in the number of participants, then the full cartel might not be sustainable. Alternatively, a full cartel might be sustainable only when firms set a lower price than the unconstrained profit-maximizing price. This could yield lower per-firm collusive gains compared to an incomplete cartel that is sustainable even when it charges the unconstrained profit-maximizing price and/or output. In other words, when δ^* is increasing in k , there might exist a range of values for δ for which incomplete cartels are the preferred subgame perfect equilibrium.

2.5.1 *Collusive price leadership*

A cartel that operates as a price leader chooses its price such that the joint marginal costs of its members equals marginal revenue corresponding to the residual demand curve. Given that all cartel participants are identical, this implies that every participant produces a quantity for which marginal cost equal the marginal revenue derived from residual demand. In principle, therefore, none of the members has an incentive to deviate *given the cartel price*. That is, increasing production is not a profitable strategy when the cheating firm does not alter its price. Deviating from the agreement, however, is potentially very profitable when a cartel member secretly undercuts the cartel price slightly and produces a quantity such that the deviation price equals marginal costs.

However, determining what is the optimal deviating strategy in the basic standard collusive price leadership model is difficult. The reason for this is that the demand structure is such that firm demand does not depend directly on the prices set by rivals. Hence, to derive an explicit equation for the critical discount factor one needs to take a slightly different approach. For instance, one could assume that products are less than perfect substitutes. Posada (2001) makes an attempt in this direction. He finds that for a market with Bertrand competition, the critical discount factor is an increasing function of the number of cartel members. Consequently, when the actual discount factor is not too high, the only sustainable cartels are less than all-inclusive. Overall, however, this contribution illustrates that an analysis of incomplete cartels in repeated Bertrand settings very easily becomes intractable.

2.5.2 *Differentiated goods*

Above, it has been shown that when firms compete in price and produce differentiated goods all cartel sizes are profitable. As a result, there always exists a δ for which any incomplete cartel is sustainable. However, for a certain cartel to be sustainable it is required that in the event of cheating the non-cheating members can credibly commit themselves to dissolve the cartel. Eaton and Eswaran (1998) considers the situation in which non-cheating members may find it profitable to continue to operate as a cartel and refrain from implementing the punishment strategy. In other words, they assume the rest of the cartel to remain after a firm defected from the agreement. The authors refer to this assumption as ‘stacked reversion’ in contrast to ‘Nash reversion’.

Eaton and Eswaran (1998) shows with a numerical example that the critical discount factor is increasing in the number of cartel participants. However, under the assumption of ‘stacked reversion’, it increases much more rapidly. The reason is that the punishment for deviating from the agreement is less severe when the remaining members adhere to the agreement. Consequently, some cartels that are sustainable with ‘Nash reversion’ are not sustainable with ‘stacked reversion’. The main finding of the study is that under the assumption of ‘stacked reversion’ only cartels comprising a small fraction of the industry might be viable. This result is driven by the fact that cartel participants may have an incentive to leave the cartel as long as non-cheating members find it profitable to adhere to the agreement.

2.5.3 Quantity competition

As before, we also consider the possibility that firms compete in quantity. In the following, a distinction is made between a simultaneous move game and a setting in which the cartel has a first mover advantage.

2.5.3.1 Simultaneous play

Friedman (1971) shows that when firms play a simultaneous Cournot game with an infinite horizon there always exists a δ for which the monopoly outcome can be sustained. In a similar setting, Escrihuela-Villar (2004) analyzes the possibility that not every firm participates in the cartel. Consider the simultaneous Cournot game discussed above. The maximum amount to be earned by a chiseling firm is,

$$\pi^d = \frac{(k+1)^2}{4k} \pi^c. \quad (2.21)$$

Observe that any cartel is inherently unstable, because $\pi^d > \pi^c$ for $k > 1$. Combining with (2.13) and (2.14) and normalizing $\frac{(a-c)^2}{b} = 1$ we get,

$$\delta \geq \delta^* = \frac{(k-1)^2(n+1)^2}{(k+1)^2(n+1)^2 - 4k^2(n-k+2)^2}. \quad (2.22)$$

This critical discount factor δ^* is increasing in the number of firms. Hence, the more firms are operating in a given industry the higher will be the critical discount factor and the less likely it is that the actual discount factor is high enough to allow for effective collusion, all else equal. At the same time, the critical discount factor is strictly decreasing in the number of cartel participants. That is, whenever a cartel of k firms is sustainable, i.e., spans at least 80% of the market and $\delta \geq \delta^*$, then larger coalitions are sustainable too.

The latter result is not obvious and, in fact, the result of a trade-off. On the one hand, the larger the number of participants, the higher is the profit level. On the other hand, the incentive to cheat is positively correlated with the number of cartel members. Yet, the first effect dominates. It must be noted that this result in part depends on the type of punishment strategy. For example, when firms adopt a more optimal punishment scheme as in Abreu (1986, 1988), the result is partly reversed.

That is, the critical discount factor is decreasing in the number of cartel participants, but only for certain parameter values.

2.5.3.2 Sequential play

Martin (1990, 2002) examines a setting in which the cartel operates as a Stackelberg leader. In every period, cartel profits per member are therefore given by (2.18). He further assumes that in the event of cheating all firms compete à la Cournot in all periods following the period of defection. The static Nash profits π^N are given by (2.13). It remains to determine what a cartel member can maximally earn when it decides to defect on the cartel arrangement. Given that $k-1$ firms stick to the collusive output and $n-k$ firms play their fringe outputs, a chiseling member optimally sets the following output level,

$$q^d = \frac{(a-c)(n+1)}{4bk(n-k+1)}, \quad (2.23)$$

with concomitant profits,

$$\pi^d = \frac{(a-c)^2(n+1)^2}{16b(k(n-k+1))^2}. \quad (2.24)$$

Collusion can therefore be sustained if and only if,

$$\delta \geq \delta^* = \frac{(n+1)^2}{4k(n-k+1)}. \quad (2.25)$$

It is easy to verify that δ^* depends non-monotonically on k . In fact, δ^* is decreasing in k as long as $k < \frac{n-1}{2}$ and the reverse holds for larger values of k . It can be concluded that no cartel is sustainable whenever an incomplete cartel of size $k = \frac{n-1}{2}$ is not sustainable.

2.5.4 Comparison

A sustainable cartel is profitable, but the converse is not necessarily true. In the previous section, we have shown that there exists a variety of profitable coalitions in the five scenarios that were considered. All these cartels, however, are not self-enforcing in the static setting. One way of explaining how firms can sustain some level of collusion is by assuming that the static games are repeated for an infinite number of periods. In this section, we have shown that when the static games become part of a supergame with infinite horizon many coalitions are sustainable as subgame perfect equilibrium of the repeated game. In fact, when firms are sufficiently patient, any profitable coalition is sustainable. Clearly, punishment strategies are a very effective means to sustain collusion, but in some sense they are too successful. Indeed, the number of subgame perfect equilibria in these games is, as noted by Tirole (1988), ‘an embarrassment of riches’.

It should be emphasized that an all-inclusive cartel could be less effective than an incomplete cartel. Particularly, it might occur that when δ is sufficiently low the ICC

for members of the full cartel is binding, while it is not binding for some incomplete cartel. Thus, for certain values of δ an incomplete cartel in principle could yield more profits per member than a cartel that encompasses all firms in the industry. Whether or not this is the case depends in part on the setting under consideration. For the five oligopoly models, we might ask for what cartel size(s) the critical discount factor reaches its minimum. The following table shows for what cartel size δ^* is lowest.

Table 2.2: Cartel size (k/n) for which δ^* is lowest.

Strategic variable	Order of moves	Homogeneous goods	Heterogeneous goods
Price	Simultaneous	1	$2/n$
	Sequential	$2/n$	-
Quantity	Simultaneous	1	-
	Sequential	$\pm 1/2$	-

As Table 2.2 reveals, for what cartel size the critical discount factor reaches its minimum varies per setting. In the simultaneous models with homogeneous products, δ^* is lowest when all firms take part in the cartel. In other words, full collusion is sustainable whenever some collusion can be sustained. An all-inclusive cartel is therefore always preferred by cartel members. By contrast, when δ is relatively low, an incomplete cartel could be more effective when the cartel has a first mover advantage or when products are differentiated. In these scenarios, the critical discount factor does not reach its minimum for a cartel that comprises the entire industry. As a result, there might exist ranges of values for δ for which an incomplete cartel yields higher profits per member than the full cartel.

2.6 The Participation Puzzle

We have established that a variety of incomplete cartels is viable provided that cartel members are sufficiently patient. So far, however, we did not address the question how such a cartel can emerge. That is to say, we did explore under what conditions none of the participants finds it profitable to deviate from the agreement, but did not consider how these firms became part of the cartel. This issue is of particular interest if one is concerned with cartels that are not all-inclusive, because then only a subset of firms should have an incentive to engage in an anticompetitive arrangement. In this section, we discuss the incentives of firms to collude or to become an outsider to a cartel formed by rivals instead.

The discussion in the previous sections reveals that a cartel functions as a public good. A group of firms that collectively reduces outputs and increases prices creates a positive externality for firms that do not take part in the conspiracy. The best response of fringe members is often to increase both their production levels and their prices, which implies windfall profits. However, the main issue is not so much that outsiders benefit from the cartel, but that outsiders typically benefit *more* than insiders. As a consequence, there exists a strong incentive to free-ride on a cartel formed by com-

petitors.⁵¹ The ensuing “participation-puzzle” has been clearly formulated by George Stigler in a discussion on horizontal mergers. He stated that,

“...the major difficulty in forming a *merger* is that it is more profitable to be outside a *merger* than to be a participant. The outsider sells at the same price but at a much larger output at which marginal cost equals price. Hence, the promoter of a *merger* is likely to receive much encouragement from each firm - almost every encouragement, in fact, except participation...”

This reasoning applies equally well to (incomplete) cartels. It is therefore especially difficult to understand how less than all-inclusive cartels can emerge in equilibrium. Although all firms in the industry would benefit from the cartel, the very fact that the choice for fringe membership often strictly dominates the choice of becoming a cartel member may prevent the formation of (incomplete) cartels.

Firms may have an incentive to form a particular cartel when the participation constraint is satisfied, i.e., when the cartel is stable. Recall that for a cartel to be stable it is required that none of the firms wants to change membership. To begin, consider the Bertrand homogeneous goods model. For this model, it was shown that the only sustainable coalition is one comprising the entire industry. Notice that the full cartel is trivially externally stable, because there are no outsiders. The all-inclusive cartel is internally stable if $\pi^c(n) \geq \pi^o(n-1) = 0$, which holds for all n . Firms therefore do have an incentive to form the full cartel in this setting. In the following, we apply the participation constraint to the remaining four oligopoly models.

2.6.1 Collusive price leadership

D’Aspremont *et al.* (1983) apply their notion of stable cartels to a basic model of collusive price leadership. To explain its workings, we assume that the inverse demand function is given by,

$$p = \frac{1}{bn} (an - Q), \quad (2.26)$$

and further suppose that firms have identical cost functions of the form,

$$C(q) = \frac{1}{2}cq^2. \quad (2.27)$$

Recall that in a setting of collusive price leadership fringe firms take the cartel price as given and produce quantities so that price equals marginal cost. We therefore have that $cq^o = p^c$, which determines the output for an individual fringe member $q^o = \frac{p^c}{c}$.

⁵¹A separate strand of literature focuses on how a firm or group of firms might raise rivals’ costs. An incomplete cartel could attempt to increase the costs of outsiders and thereby reduce free-rider incentives. Potential cost raising scenarios will not be discussed in detail here and the reader is referred to Salop and Scheffman (1983, 1987) for analyses and discussion.

With a number of $n - k$ outsiders this yields a residual demand for the cartel that amounts to,

$$kq^c = n(a - bp^c) - (n - k) \left(\frac{p^c}{c} \right). \quad (2.28)$$

The joint-profit maximizing price of the cartel is then equal to,

$$p^c = \frac{an(1 - c)}{\left(bn + \frac{n-k}{c} \right) (2 - c)}, \quad (2.29)$$

which is strictly higher than the competitive price and which yields per firm cartel profits equal to,

$$\pi^c(k) = \frac{ca^2}{2 \left((n + bc)^2 - k^2 \right)}. \quad (2.30)$$

Outsiders charge the same price for their products, but produce more. Their individual profits amount to,

$$\pi^o(k) = \frac{ca^2(n + bc)^2}{2 \left((n + bc)^2 - k^2 \right)^2}. \quad (2.31)$$

This analysis yields several interesting results. First, cartel profits per member are strictly increasing in the number of cartel participants. Second, outsiders earn more profits than insiders for all $k > 1$ so that the ‘participation puzzle’ is fully applicable.⁵²

The main result of d’Aspremont *et al.* (1983) is that, under the reasonable assumption that n is finite, there always exists a coalition for which the participation constraint is satisfied. It must be noted, however, that the proof does not exclude the coalition consisting of only one firm and that the stable cartel may as well comprise the entire industry.⁵³ Also, the fraction of firms in a stable cartel is decreasing in the size of the industry. That is, if n increases the number of firms in the stable cartel tends to zero.

2.6.2 Differentiated goods

Applying the participation constraint to the Bertrand model with differentiated goods yields intractable expressions. Observe, however, that outsiders always earn more profits than insiders, i.e., (2.11) is larger than (2.10) for all $k > 1$. Clearly, for all cartel sizes, firms prefer to be outside the cartel. However, whether or not firms have an

⁵²Note that $\pi^o(k) > \pi^c(k)$ holds for all k does not imply that there exists no internally stable coalition. That is, $\pi^o(k) > \pi^c(k)$ is not inconsistent with $\pi^c(k) > \pi^o(k - 1)$.

⁵³Donsimoni *et al.* (1986) shows that, in this setting, the stable cartel may not be unique when firms are cost-efficient relative to market demand. In particular, they show that there exist industry sizes for which two stable cartels exist, a cartel that is incomplete and a cartel that is all-inclusive.

incentive to change membership cannot be determined *per se*. In all likelihood, this will depend heavily on the parameter specifications.⁵⁴

2.6.3 Quantity competition

Clearly, the participation constraint can also be applied to games in which firms have output rather than price as their choice variable. Like before, the simultaneous and sequential move games are discussed respectively.

2.6.3.1 Simultaneous play

Consider the Cournot market with homogeneous goods and linear demand analyzed above. A minimum requirement for firms to have an incentive to form a cartel of k firms in this setting is that none of the k members finds it in their interest to be an outsider to a cartel consisting of $k - 1$ firms. To begin, recall that any cartel for which $k < n + \frac{3}{2} - \frac{1}{2}(4n + 5)^{\frac{1}{2}}$ is not profitable in a sense that members are worse off than in competition. Clearly, such a cartel is never internally stable, because fringe profits are always (weakly) larger than competitive profits. In other words, firms never have an incentive to form a cartel with a market share less than 80%.

We now focus on the situation in which $k \geq n + \frac{3}{2} - \frac{1}{2}(4n + 5)^{\frac{1}{2}}$ so that any cartel member is (weakly) better off than in competition. Formally, the cartel is internally stable if,

$$\pi^c(k) = \frac{(a - c)^2}{bk(n - k + 2)^2} \geq \pi^0(k - 1) = \frac{(a - c)^2}{b(n - k + 3)^2}.$$

It is easy to verify that this condition is violated for all $k > 1$ as long as $n \geq 3$. The condition does hold for $n = 2$ and $k = 2$. It can be concluded that cartel members have an incentive to leave the cartel as long as the remaining coalition is (weakly) profitable. Accordingly, firms have an incentive to form a cartel of size $k \simeq n + \frac{3}{2} - \frac{1}{2}(4n + 5)^{\frac{1}{2}}$. Such a cartel is externally stable, because outsiders always earn higher profits than insiders. The reason that no member has an incentive to leave the cartel is that this would lead to Cournot-Nash profits that are (weakly) lower.⁵⁵

2.6.3.2 Sequential play

Shaffer (1995) applies the participation constraint to the sequential Cournot model.⁵⁶ Using the expressions,

$$\pi^c = \frac{(a - c)^2}{4bk(n - k + 1)}, \tag{2.32}$$

and,

$$\pi^o = \frac{(a - c)^2}{4b(n - k + 1)^2}, \tag{2.33}$$

⁵⁴See, for instance, Bloch (2002).

⁵⁵See also Eseriuhela-villar (2004, 2008a).

⁵⁶See also Martin (2002).

it can be shown that the cartel is (weakly) internally stable if,

$$k \leq \frac{(n - k + 2)^2}{(n - k + 1)}. \quad (2.34)$$

Hence, none of the firms is willing to leave the cartel only if the number of conspirators is not too big. The intuition behind this result is that a coalition between a small number of firms gives a relative big share of cartel profits to participants. At the same time, outside competition is fierce, which makes joining the fringe not too attractive. In contrast, if the number of cartel participants is large, cartel profits are shared among many, while the number of outsiders is small. Leaving the cartel is attractive now, because the fringe profits are relatively high compared to cartel profits. Moreover, note that for some combinations (n, k) ,

$$\frac{(a - c)^2}{4b(n - k + 1)^2} \geq \frac{(a - c)^2}{4bk(n - k + 1)} \geq \frac{(a - c)^2}{4b(n - k + 2)^2}, \quad (2.35)$$

implying that the cartel is (weakly) internally stable even though outsiders earn higher profits. This boundary result emerges when joining the fringe lowers the profits of outsiders just enough so that staying in the cartel is preferred.

The cartel is (weakly) externally stable if,

$$k \geq n - k + 1 + \frac{1}{n - k}. \quad (2.36)$$

Hence, the number of conspirators must be sufficiently large to make joining the cartel not attractive. In other words, although joining the cartel always increases total cartel profits, it lowers the per firm cartel profits due to an increasing number of participants.

The (subgame) perfect equilibrium of this Stackelberg game yields the following stable cartels for $n \geq 3$,

- $n > 4 \implies k = \frac{1}{2}n + 1$, for even n ,
- $n > 4 \implies k = \frac{1}{2}(n + 1) + 1$, for odd n ,
- $n = 3 \implies k = 3$
- $n = 4 \implies k = 3 \vee k = 4$

Note that, apart from $n = 4$, the equilibrium number of cartel participants is uniquely determined. In particular, the equilibrium number of cartel participants facing a Cournot fringe is just over half the firms in the industry regardless of the total number of firms. Finally, note that outsiders in equilibrium earn more profits than insiders.⁵⁷ Nevertheless, under certain conditions, insiders may be better off than outsiders when the cartel operates as a quantity leader. Salant (1976), for example, analyzes the world oil market with a quantity setting game in which the cartel is a

⁵⁷When $k < \frac{n+1}{2}$, insiders earn more profits than outsiders. Note, however, that this will never occur in equilibrium, i.e., such a small cartel is not stable.

dominant firm. Participation is beneficial because the cartel builds up more reserves by restricting outputs, while the stock of fringe firms will be exhausted sooner leaving the cartel a monopoly position in the future.

2.6.4 Comparison

The incentive to form a particular cartel in the various models are summarized in Table 2.3.

Table 2.3: Profitable stable incomplete cartels.

Strategic variable	Order of moves	Homogeneous goods	Heterogeneous goods
Price	Simultaneous	none	unknown
	Sequential	unknown	-
Quantity	Simultaneous	$k \simeq n + \frac{3}{2} - \frac{1}{2} (4n + 5)^{\frac{1}{2}}$	-
	Sequential	$k \simeq \frac{1}{2}n$	-

Based on the studies discussed in this section, we can draw two main conclusions. First, when firms compete in price, it might be that no profitable stable incomplete cartel exists. This certainly holds when firms decide on their choice variable simultaneously and produce homogeneous commodities. Whether or not this conclusion carries over to a Bertrand setting with differentiated goods and a model of collusive price leadership is uncertain. Bloch (2002) shows numerically that, in a Bertrand setting with differentiated goods, profits per cartel member are increasing in cartel size. Yet, it remains unclear whether or not firms have an incentive to form a particular cartel. Given that these type of numerical analysis are very sensitive to parameter specifications, results, if any, are likely to be only of limited value. Second, a stable cartel always exists when firms compete in quantity. When the cartel is assumed to take a leading role in the industry, firms have an incentive to engage in a cartel that comprises about half the industry. Instead, when firms take their strategic decisions simultaneously, firms have an incentive to form a cartel that controls approximately 80-90% of industry supply.

It is important to emphasize that the existence of stable incomplete cartels does not completely solve the participation puzzle. We have shown that only a subset of firms might find it in their interest to collude, but we did not say anything about how firms arrive at a particular equilibrium cartel arrangement. That is, the existence of a stable incomplete cartel reveals that firms have an incentive to engage in an incomplete coalition, but nothing is said about how such a coalition actually forms. This issue is discussed in the next section.

2.7 Coalition Formation with Positive Externalities

In the previous section, we have discussed the incentives of firms to form an incomplete cartel, but did not consider the actual coalition formation process. A separate

strand of literature focuses on the manner in which a cartel is formed. This literature typically assumes firms to play a two-stage game. In the first stage, players form a coalition according to the specific rules of the game. Given the established coalition structure, the game is played noncooperatively in the second stage. The oligopoly models analyzed above are examples of games played in the second stage. Recall that in this type of models a coalition creates positive externalities, i.e., non-participants benefit from the cartel.⁵⁸ Hence, we are naturally interested in coalition formation games with positive externalities. In this section, we review some of the main studies that deal with cartel formation explicitly.⁵⁹

To begin, it must be noted that we limit ourselves to a discussion of the formation of a single coalition with independent outsiders. The possibility of forming multiple coalitions is only marginally considered. In particular, we do not survey the literature on “equilibrium binding agreements” in which coalitions are allowed to split up in smaller coalitions.⁶⁰ Moreover, this section is organized slightly differently from the previous sections. As before, we do categorize the games with respect to the timing of decision-making. However, we do not discuss the five oligopoly models separately. Alternatively, we make a distinction between, on the one hand, bargaining or “exclusive membership” games and, on the other hand, participation or “open membership” games. In settings of “exclusive membership” an undertaking is only allowed to join the cartel when all members agree. In “open membership” games, by contrast, players have the freedom to join whatever coalition they like.

2.7.1 *Simultaneous cartel formation*

In simultaneous coalition formation games, firms do know the strategy space of the game, but have no knowledge about the actual participation decision of competitors. Arguably, this approach is somewhat restrictive within the context of cartel formation. After all, a cartel is commonly initiated by one firm or a small group of firms which operates as the cartel ringmaster. The ringmaster typically contacts other competitors with the request to become part of the ring.⁶¹ This implies that firms often have some knowledge about the participation decision of rivals. Nevertheless, the assumption of firms taking formation decisions simultaneously allows for a deeper understanding of what type of cartels are likely to emerge. We can distinguish between simultaneous bargaining models and simultaneous participation models, which are subsequently discussed.

⁵⁸In contrast, a coalition could create negative externalities. A well-known example is a R&D joint-venture that creates efficiency gains for participating firms. These forms of cooperation may yield a competitive disadvantage for outsiders. See, for example, Bloch (1995).

⁵⁹For an extensive overview and comparative analyses of stable coalition structures with externalities see Yi (1997). For traditional coalition formation literature that does not consider externalities across coalitions see, for example, Aumann and Dreze (1974) and Shenoy (1979).

⁶⁰See Ray and Vohra (1999).

⁶¹Occasionally, firms are more or less forced to join the cartel. See the brief description of the international diamond cartel in this chapter.

2.7.1.1 Simultaneous bargaining

A simultaneous bargaining game can be considered a game of exclusive membership. In such a setting, all players simultaneously propose a coalition in which they are willing to engage. There are two well-known variations.⁶² In one version, a cartel is formed only when all participants proposed this particular coalition. For example, the coalition $C = \{i, j, z\}$ will form if players i, j and z all three propose C . The second version is less stringent and rules that a coalition will emerge if part of the players reach consensus. For instance, if C is proposed by i and j , but not by z , then i and j will form a cartel.

It is important to emphasize that players base their proposals on the expected gain from participating in a particular coalition. Players may be aware, however, that their *ex ante* expectations of participating may not coincide with *ex post* earnings. Hart and Kurz (1983), for instance, analyzes a non-cooperative setting in which players simultaneously announce coalitions. In addition, however, they examine the bargaining process within a coalition once it is formed and analyze what coalition structures are stable given the *ex ante* value of a player.

2.7.1.2 Simultaneous participation

Probably the first research that took this approach is the game-theoretic work *where four are few and six are many* by Selten (1973). The main aim of this contribution is to analyze the conjecture by Chamberlin (1933) and Stigler (1964) that joint profit maximization among the few is likely, while impossible in markets with many firms. Selten analyzes a three-stage simultaneous Cournot game. In the first stage, firms simultaneously decide if they want to participate in the cartel or not, the second stage determines the cartel output quota, which in turn are implemented in the third stage. The main finding is that in markets with four or less players all firms will participate in a cartel, while in industries with six or more firms no cartels will be formed leaving five as the intermediate case.

In the game considered by Selten (1973), cartel size is taken to be endogenous. He calculates the probability of a particular coalition to emerge in equilibrium. In an industry with four or fewer firms all-inclusive cartels form with probability one. Hence, in such concentrated industries incomplete cartels cannot be explained as a subgame perfect equilibrium. In an industry with five sellers, the probability of an incomplete cartel consisting of four firms is approximately 18%. For markets with more than five firms this probability is positive but very small. It is below 1.3% and tends to zero quickly for larger industries. Therefore, in markets with five or more undertakings the probability of incomplete collusion to occur is positive, but very small.

A similar approach has been conducted by Prokop (1999) who investigates the case of a dominant cartel with price leadership. Recall, that in this setting outsiders earn always higher profits than insiders independent of the number of participants. One may, therefore, *a priori* expect all firms to decide for the outsider position if such decisions are supposed to be taken independently and simultaneously. Indeed, the

⁶²See, for example, Bloch (2005).

situation is similar to that of the well-known prisoner's dilemma. Yet, at the same time we know all firms benefit if some cartel nevertheless emerges. The outcome of the game depends on parameter values.

In particular, if costs are sufficiently high, a complete cartel will form in industries with $n \leq 3$ with certainty, while in industries with $n \geq 4$ each firm joins the cartel with a positive probability. By contrast, if firms in the industry are sufficiently efficient a complete cartel will form with certainty for $n \leq 5$. For $n > 5$ the uniqueness of a pure strategy equilibrium, i.e., all firms are willing to participate, cannot be guaranteed and very much depends on the value of the cost parameters. It is further shown, that both the participation probability as well as the expected size of the cartel are declining in the industry size.⁶³ The reason is that within this setting the stable cartel size is independent of n , which makes the free-rider effect more severe when more firms are operating on the market. It is noteworthy that the expected result does not coincide with the stability of the cartel. So, even though such a stable cartel always exist, it may be difficult to reach this outcome when firms take their participation decision simultaneously.

The case of a cartel that operates as quantity leader has been examined by Hviid (1992). This contribution is particularly concerned with the question how private information (e.g., about a common parameter in the demand function) affects the incentives to form a cartel. Note that the Stackelberg assumption implies that part of the information is either directly disseminated to the fringe or indirectly, for example, because it is reflected in the output decision of the cartel that is known by fringe members with certainty. By assumption, however, the analysis solely focuses on equilibria that comprise all firms in the industry. Instead, changes in industry size are used as an indicator of how private information is likely to affect the incentive to form cartels. The main conclusion is that if outsiders can infer the information of the cartel, it will, *ceteris paribus*, lead to smaller cartels, i.e., the existence of private information yields a disincentive to cartel formation. What drives this result is that, apart from the existing free-rider effect, outsiders also benefit from correct inference of the cartel's output choice and the attempt of the cartel to manipulate this inference is shown to have an adverse effect on cartel profits.

2.7.2 Sequential cartel formation

In sequential cartel formation games, players take a decision or make a proposal in a particular order. This implies that some players have more information about the participation decisions of rivals than others. Arguably, a sequential approach more accurately reflects cartel formation in practice. After all, a manager that has to decide on whether or not to take part in the conspiracy typically has some information about the plans of (part of) the other firms. For instance, receiving an invitation to attend a meeting in a "smoke-filled room" at least partly reveals the intentions of the other firms that are invited.

⁶³Note the similarity with the symmetric Cournot game analyzed in Selten (1973).

However, assuming that participation decisions are taken sequentially comes with a price. The outcomes of sequential formation games tend to be very sensitive to the often arbitrary assumptions about the order in which firms take decisions. Clearly, firm a taking its decision before firm b could lead to an outcome that might differ substantially from a situation in which firm b decides first. Obviously, however, this drawback is far less severe under the assumption that firms are identical.

In sequential formation games, every decision maker has (perfect) knowledge about the history of the game, because otherwise it would be *de facto* simultaneous. At the same time, players anticipate the consequence of their decision for (the optimal) decisions made by subsequent players, i.e., players are farsighted. Moreover, for such a formation setting to work, it is important that once a decision is made it is binding for the rest of the formation stage. That is, a sequential formation game is only fundamentally different from a simultaneous formation game when firms commit to their choices. As before, we distinguish between bargaining and participation models, which are subsequently discussed.

2.7.2.1 Sequential bargaining

Bloch (1996) proposes a method that allows coalitions to form in a non-cooperative sequential process. Given a certain order of players, the first player proposes a coalition, which is formed when accepted by all prospective members. The first player to reject the proposal becomes the initiator in the next round. An important property of this method is that firms which successfully form a cartel are removed from the game so that the rest of the game is played only by the remaining players. Moreover, the focus is exclusively on the formation of coalitions given a certain sharing-rule of the value of a coalition. Hence, players *ex ante* know what their earnings will be in a given coalition.

The author applies the method to the simultaneous Cournot setting discussed above. Recall that a cartel must at least consist of 80% of the firms in order to be profitable and that outsiders benefit more from the cartel than insiders. He finds that the unique equilibrium of the formation game is such that the first $n - k$ firms remain independent, while the last k firms form a cartel. Given the number of firms in the industry the unique k is given by the following equation,

$$k \geq n + \frac{3 - \sqrt{4n + 5}}{2}. \quad (2.37)$$

That is, the first integer that “solves” (2.37) gives the equilibrium cartel size. The reason that the last $n - k$ firms form a cartel even though outsiders are better off is that their next best alternative is competition and competitive profits are lower than cartel profits given that the 80%-requirement is met.

In the previous section, it was shown that firms have an incentive to form this particular cartel. Due to the free-rider incentive, the unique cartel size is the minimum cartel size required to sustain some collusion. Yet, although this result is theoretically sound, it is difficult to give the formation mechanism a practical interpretation. That is, it is difficult to see how a couple of firms start the formation process by announcing “no”, leaving the remaining firms in a situation in which they are willing to say “yes”.

Part of this problem is due to the fact that commitment to the participation decision is not credible.

2.7.2.2 Sequential participation

In a sequential participation game firms, one after the other, take a binary decision, which is either to participate in a cartel or to remain an independent outsider instead. Prokop (1999) analyzes the process of cartel formation in relation to collusive price leadership as studied by d’Aspremont *et al.* (1983). In particular, he attempts to find an answer if such stable cartels will emerge if the participation decision is made endogenous. He finds that within the sequential-move game the first $n - k$ firms refuse to participate leaving the task for the last k firms to form the cartel. It is important to emphasize that, in contrast to the analysis of this model with simultaneous participation decisions, the cartel that is formed is also stable. The intuition underlying this result is similar to that in Bloch (1996). The first firms understand that saying “no” leaves a situation that is internally stable, but not externally stable for firms later on in the sequence of decision makers. Hence, after $n - k$ firms said “no” it is in the self-interest of the next firm to say “yes” and all the others that follow, given that a cartel of k firms is indeed stable.

2.8 Incomplete Cartels and Firm Heterogeneity

The theoretical studies of incomplete cartels that we have discussed up until now assume identical firms. This assumption has the major advantage that it often allows the researcher to obtain clean analytical results. In many instances, calculations are tedious and can easily become intractable without this assumption. However, assuming identical firms is potentially problematic when studying incomplete cartels, because it is *a priori* unclear how identical firms would take non-identical decisions, which is trivially needed for an incomplete cartel to emerge. That is, assuming identical firms implies that firms lack identity in the model and therefore we cannot analyze why a particular firm decided to join the cartel or became an independent outsider instead. In other words, nothing can be said about which or what type of firms are more inclined to join a cartel. Moreover, any collusive equilibrium is unique only up to permutations.

There are a few theoretical studies that analyze incomplete cartels when firms are heterogeneous, which are briefly surveyed in this section. We will focus exclusively on studies that assume firms to differ *ex ante*, i.e., prior to cartel formation. For example, the possibility that participants enjoy a cost advantage, as may be the case with a merger, is excluded. This is based on the belief that cartel organizations will create substantially less synergies than mergers.⁶⁴

⁶⁴Perry and Porter (1985) and Farrell and Shapiro (1990) analyze the ability of firms to create synergies in horizontal merger settings. The central idea is that cooperation may yield a significant reduction in production costs so that *ex post* insiders may be better off than outsiders. It seems difficult to convincingly defend the case of cost-saving cartels. On the contrary, restricting outputs in

Heterogeneity among firms can take many forms. For instance, a firm can have an information advantage over its rivals, access to more financial resources or control more production capacity. Also, some sellers may enjoy the benefits of having a more advantageous geographical location. Under a regime of basing-point pricing, for example, a group of firms located relatively close together potentially could use a distant firm as a natural focal point to determine a collusive base location. This not only allows for a stable agreement, but at the same time prevents outsiders from profitably entering local markets. Indeed, basing-point pricing is well-known to facilitate incomplete cartels.⁶⁵

Differences between sellers might explain why full collusion sometimes seems to be an unattainable ideal. For example, when firms have different cost levels it might be difficult to establish a cartel contract that is acceptable to all parties. Low-cost firms typically prefer a lower cartel price than high-cost firms.⁶⁶ We can imagine that it is easier to form a cartel agreement between firms that are more or less identical. That is, substantial differences in production costs could hurt the negotiation process and, as a result, full collusion could be difficult to arrange.⁶⁷ In addition, the stability of the cartel may depend on relative cost efficiencies of firms as is shown in Rothschild (1999).

Donsimoni (1985) studies incomplete cartels when firms differ in terms of unit production cost in a model of collusive price leadership as discussed above. Cost functions are identical among firms of the same type, but allowed to differ across types. The main question is what type of firms are more inclined to become a price leader, i.e., what type of firms have a stronger incentive to join the cartel. Among other things, she finds that the most efficient firms will be in the cartel, while less efficient firms remain independent outsiders. Collusion among firms that differ in terms of production costs is also studied by Cramton and Palfrey (1990). The authors analyze a static model with a finite number of firms and a continuum of cost types and find that a group of low-cost firms might find it in their interest to bribe high cost firms not to produce. However, the main purpose of this study is to analyze if heterogeneity in costs prevents firms from obtaining the monopoly outcome. The main result is that with sufficiently many firms in the industry, the monopoly outcome cannot be obtained. This is in line with the conventional wisdom that collusion becomes increasingly difficult the higher the number of firms.

A different form of asymmetry is introduced by McAfee (1994). He analyzes a particular advertising model to study cartel formation. Firms differ in their availability

the presence of potential scale economies cartel members are more likely to have higher production costs.

⁶⁵Chapter 5 provides an in-depth discussion of basing-point pricing.

⁶⁶This is not to say that collusion among firms with different efficiency levels yields relatively low cartel prices. For example, as is shown in Harrington (1991) the optimal collusive price might exceed the monopoly price of a low-cost firm given that the discount factor is sufficiently high.

⁶⁷Alexander (1997) analysis of the 1930s pasta industry suggests that cost heterogeneity is a problem to form a cartel. In particular, the large firms had low costs and therefore a different optimal collusive price than high costs small firms.

rate, which is defined as the probability that a consumer receives a price offer from a particular firm. He finds that there exists an equilibrium in which only the largest firms participate in the cartel. Moreover, cartels typically comprise at least two, but generally not every firm in the industry.

The studies discussed so far in this section use a static model and cartel agreements are therefore *de facto* binding. That is, firms face a participation constraint, but no incentive constraint. To the best of our knowledge, there is no contribution that considers incomplete cartels with firm heterogeneity in an infinitely repeated game setting. The effect of firm heterogeneity on collusion has been studied in a dynamic setting, but only under the assumption that the cartel is all-inclusive. Harrington (1989), for example, considers a setting in which firms differ in terms of their discount factors. Among other things, he finds that firms with a relatively low discount factor receive a relatively large share of collusive profits, which is required to prevent cheating.

There exist a few studies that explore how the size distribution of firms in the industry affects collusion. Kühn and Motta (1999) analyzes a model in which firms produce a variety of products. The authors find that small firms, i.e., firms with only one or a few type of products, have a stronger incentive to defect from the cartel agreement. In addition, implementing a punishment strategy is more costly for larger firms, i.e., firms with many type of products. Hence, a more equal distribution of product varieties facilitates collusion. Compte *et al.* (2002) and Vasconcelos (2005) analyze a supergame in which firms differ in terms of production capacity. Like Kühn and Motta (1999), both find that the highest critical discount factor is lowest when firms are of equal size. Under the assumption of an all-inclusive cartel, all three papers then find support for the idea that more symmetry among firms facilitates collusion.

2.9 Incomplete Bidding Rings

So far, we have restricted ourselves to a discussion of incomplete collusion among sellers. A substantial part of economic activity, however, is organized through auctions and procurements. These exchange mechanisms allow a group of bidders to compete for an object or contract offered by a seller. Like in almost every market there is a natural incentive for competitors to reduce competitive pressure among them with the aim to increase (short-run) gains from trade. There exists both formal and informal evidence that bidder collusion is a prevalent phenomenon and that the rings formed may or may not include all bidders.⁶⁸ Incomplete cartels in auctions have their own characteristics and are not necessarily comparable to incomplete cartels in more regular markets.⁶⁹

⁶⁸See, for example, Froeb (1988) who reports that between 1979-1988 there are 319 Sherman Act Section 1 criminal cases filed by the U.S. DOJ, 81% of which concerned auctions. Feinstein *et al.* (1985), for instance, analyzes a real-world example of an incomplete bidding ring in the highway construction industry

⁶⁹For example, as mentioned by Porter and Zona (1993,1999), some participants may purposely decide not to submit bids. It is difficult to see how some cartel members that operate in more regular markets could temporarily “disappear”, although participants may refuse to sell to certain groups of buyers.

In the following, a brief overview is provided of the theoretical economic literature on incomplete bidding rings.

Like in the discussion above, we may distinguish between the feasibility of a cartel and cartel formation. As regards to the sustainability of bidding rings there is a variety of methods available to guarantee cartel stability. For instance, the cartel may appoint an enforcer who is given the task to punish observed chiselers. Simultaneously, when bidders interact frequently, the ring may adopt punishment strategies to ensure compliance. A bidder who does not adhere to the cartel policy is confronted with lower competitive payoffs in future auctions. A bidding ring is then effective when discounting is sufficiently low. In short, bidding rings often can be sustained in much the same way as cartel arrangements in more regular markets.

In a recent contribution, Marshall and Marx (2007) develop a framework in which the ring not only faces competition from outside bidders, but is also confronted with potential stability problems. They conclude that collusion is less effective in a first-price rather than a second-price auction. The authors explain the difference in the incentives to chisel in a first- and second-price auction as follows,

“At a second-price auction, a ring must suppress the bids of all members except the bidder with highest value. The ring member with the highest value goes to the auction and bids as he would were he acting non-cooperatively. Any ring member who thinks of breaking ranks and competing at the auction faces the highest ring bidder and the highest non-ring bidder, each submitting bids that are the same as if all were acting non-cooperatively. Thus, there is no gain to deviant behavior. The first-price auction is quite different. In order to secure a collusive gain, the ring member with the highest value must lower his bid below what he would have bid acting non-cooperatively, and other ring members must suppress their bids. But when the highest-valuing ring member lowers his bid, the opportunity is created for a non-highest-valuing cartel member to enter a bid at the auction, either on his own or through a shill, and secure the item. This possibility jeopardizes the feasibility of a cartel at a first-price auction.”⁷⁰

Among other things, they show that when a ring adopts a so-called ‘bid coordination mechanism’ it cannot eliminate all ring competition at a first-price auction.⁷¹ Furthermore, the work illustrates that incomplete cartels in auctions quite generally are profitable.

Also, the ‘participation puzzle’ might well be different in auctions. More accurately, unlike for incomplete coalitions in regular markets, cartel formation may be less of a problem for bidding rings. The reason is that the incentive to free-ride on a cartel

⁷⁰See Marshall and Marx (2007), pp. 375-376.

⁷¹In a bid coordination mechanism (BCM), the cartel can install a side-payment scheme and recommend bids to all members. However, it lacks the power to control the bids of the ring members.

might be weaker or even absent. Indeed, whether or not an incomplete bidding ring creates externalities for non-conspirators typically depends on the type of auction. To illustrate, consider the following two settings.

McAfee and McMillan (1992) analyzes a first-price auction with n bidders of which k collude. Bidders have valuation v , which is assumed to follow a discrete distribution so that $v = 0$ or $v = 1$. Let $p = \Pr\{v = 1\}$ and assume the reserve price to be zero. The expected profit of a cartel member and an independent outsider are then respectively given by,⁷²

$$\pi^c = (1-p)^{n-k} \left(\frac{1}{k}\right) \left[1 - (1-p)^k\right],$$

and

$$\pi^o = (1-p)^{n-k} p.$$

Observe that outsiders always earn more profits than insiders independent of cartel size.⁷³ Moreover, in line with the notion of stable cartels discussed above, it can be shown that a stable cartel always exists. The cartel size can be (almost) uniquely predicted, assuming that the cartel is formed through an open membership game. Following the participation constraint as given by (2.2) and (2.3), the optimal cartel size k^* is implied by the following inequality constraints,

$$\frac{1 - (1-p)^{k^*}}{k^*} \geq p(1-p) \geq \frac{1 - (1-p)^{k^*+1}}{k^* + 1}.$$

The number of bidders participating in a stable ring is increasing in p and is always larger than three.

In contrast, it can be shown that the free-riding effect is absent in second-price auctions. To see this, assume that valuations are drawn independently from a common distribution function F with density f . The gains for respectively cartel members and outsiders are given by,

$$\pi^c = \frac{1}{k} \int_0^\infty \int_0^x (x-y)(n-k) F(y)^{n-k-1} f(y) k F(x)^{k-1} f(x) dy dx,$$

and

$$\pi^o = \int_0^\infty \int_0^x (x-y)(n-1) F(y)^{n-2} f(y) f(x) dy dx.$$

As can be observed, the expected gain of outside bidders is independent of k . Moreover, the expected gain of outsiders is always lower than the expected profits of participants for all values of k . Clearly, it is always beneficial to join the cartel and the unique stable cartel is all-inclusive.

⁷²It is common in the auction literature to refer to “utility” of bidders instead of profits. For convenience, we have chosen to keep notation in line with the previous sections.

⁷³Here, we only present some of the main results. For a detailed description of the game as well as derivations the reader is referred to the original article.

The reason that free-riding incentives are absent in second-price auctions is that it is typically optimal for the ring to submit a bid equal to the bid that its highest-valuing member would submit in competition. It is well-known that in second-price auctions submitting a bid equal to the personal valuation is a dominant noncooperative strategy for each bidder. Mailath and Zemsky (1991) show that in second-price private value auctions any subset of bidders can efficiently collude, but their analysis also indicates that an all-inclusive ring is most likely even when bidders are heterogeneous. Incomplete bidding rings in second-price auctions are therefore not easily explained as the result of strong free-rider incentives. Yet, the authors remark that,

“... one would expect that in many environments it will not be feasible to extend membership to all potential bidders. For example, legal considerations might lead the ring to limit membership to avoid detection. Or there might be a large number of potential bidders each with only a small probability of having a value above the reservation price so that the costs of coordinating the large ring exceeded the benefits.”⁷⁴

Bidder heterogeneity in first-price auctions has been studied by Marshall *et al.* (1994). This type of analysis becomes easily analytically intractable and, as a result, the authors have limited themselves to a numerical analysis. They propose an algorithm that proves to be particularly useful in the analysis of incomplete cartels that face competition from independent outsiders. The study conducted suggests that the expected revenue is higher in first-price auctions, but also that cartel formation is more complicated than in English auctions. As mentioned above, this is due to the fact that in first-price auctions a less than all-inclusive cartel creates positive externalities for bidders not part of the ring. Furthermore, a noteworthy result is that small all-inclusive rings are often feasible, but quite unlikely when the number of (potential) bidders is large.

The absence of free-rider and deviation incentives in second-price auctions forms a sound explanation for the prevalence of bidder collusion. It must be noted, however, that often some additional measures are required to make the ring effective. For instance, quite frequently only one or a few cartel members attend the main auction and consequently have a positive chance of winning the object or contract. Clearly, the remaining participants want their share of the pie and this may give rise to several organizational problems. Indeed, the problem of how to divide the (extra) rents among members can be quite severe.⁷⁵ This in particular implies that bidding rings, more often than not, form examples of overt collusion. A notable exception are multi-object auctions. Brusco and Lopomo (2002) consider a multi-object English auction and establish that in this setting side contracts may not be needed to sustain some level of collusion. The reason is that in such open ascending bid auctions bidders can signal what object they value most, e.g., by abstaining from competing over certain objects. However, tacit collusion in these type of auctions becomes difficult when the number of (potential) bidders grows large.

⁷⁴See Mailath and Zemsky (1991), p. 485.

⁷⁵Eckbo (1976) shows empirical evidence that in a sample of international cartels almost half collapsed due to disagreement on how to share the profits.

There exists some real-world evidence of bidding rings that have developed sophisticated methods to cope with these type of problems. Graham and Marshall (1987), for example, reports that bidding rings address the issue of how to allocate collusive gains through “nesting”, i.e., forming rings within rings.⁷⁶ Interestingly, it appears that some rings have adopted a scheme in which every participant is awarded its “Shapley-value”, i.e., every member receives its own average marginal contribution. Graham *et al.* (1990) builds on this work and theoretically explores these so-called secondary auctions or knockouts. Among other things, they establish that in equilibrium side-payments to ring members are indeed equal to their Shapley value in a variety of settings.

2.10 Discussion

In this chapter, we have provided an overview of available theories of incomplete cartels. Three issues have been discussed in reference to five basic oligopoly models: The profitability and sustainability of incomplete cartels and the incentives of firms to form a particular cartel. It has been shown that incomplete cartels are often profitable and sustainable provided that coalition members are sufficiently patient. When firms compete in quantity, there is an incentive to form an incomplete cartel and, given industry size, the size of the incomplete cartel is uniquely determined. By contrast, in Bertrand competition conclusions are less clear cut and whether or not firms have an incentive to form a less than all-inclusive cartel is unknown.

In the remainder of the chapter, we have briefly explored cartel formation with positive externalities, incomplete cartels with heterogeneous firms and incomplete bidding rings. In simultaneous-move formation games, no incomplete cartels will form in concentrated industries, i.e., when the number of firms is weakly smaller than four. In industries with five or more sellers, firms might form an incomplete cartel, but the probability of an incomplete cartel to emerge in equilibrium is typically low. By contrast, incomplete coalitions are quite likely to emerge in sequential formation games. There is only a modest number of studies that assumes non-identical firms. These analyses suggest that larger firms are more inclined to join a cartel. Finally, we have discussed incomplete bidding rings. One of the key observations is that both the incentive problem and the participation problem are typically absent in second-price auctions. As a result, incomplete cartels in second-price auctions cannot be explained by free-rider incentives. Still, the optimal collusive arrangement might not include all bidders when the number of competitors is large, because with many (potential) bidders coordination costs can be formidable.

At this point, it is interesting to note that the Cournot models seem to better explain the stylized facts that were distilled from real-world examples of incomplete cartels. Indeed, taking account of both the incentive constraint and the participation con-

⁷⁶The authors list several ‘stylized facts’ that illustrate the organization of rings. For instance, ring members tend not to bid meaningfully against each other and the ring attempts to conceal its existence from the auctioneer.

straint, cartels are typically not all-inclusive and tend to take a dominant position in the market when firms compete in quantity. So far, the economic literature on incomplete cartels remains largely silent about what happens when firms compete in price. Also, independent of the strategic variable, the literature that was discussed cannot fully explain how incomplete cartels lose market share over time and why incomplete cartels often comprise the larger firms in the industry. The latter is suggested by some studies, but these contributions do not take account of the incentive problem of the cartel. To provide a rational for these ‘stylized facts’ one necessarily needs to take a dynamic approach in which firms differ in at least one respect.

To address these issues, one could, for example, take a structural dynamic approach as in Fershtman and Pakes (2000). In particular, altering the structure of the game potentially could shed light on how and why firms take different decisions. An example of a structural analysis applied to an incomplete cartel is deRoos (2001) which examines the vitamin c market that was part of the global *vitamin* cartel. The vitamin cartel was confronted by a group of Chinese producers. At first, the cartel decided to accommodate the emergence of Chinese vitamin c production, but eventually, after a considerable growth of fringe production it caused the demise of the cartel. The model as developed endogenizes decisions of firms on investment, entry, exit and collusion and allows for a deeper understanding of why the cartel initially accepted fringe production, but ultimately could not, or did not wish to, maintain the cartel organization anymore. Among other things, it is found that the cartel tends to accept fringe competition as long as competitive pressure remains small. In addition, less effort is taken to deter entry when it is likely that the cartel could be maintained in light of entry. Interestingly, the entrant tends to invest heavily in the presence of a cartel, which broadly speaking can be understood as the result of the free-rider effect emphasized above. It must be noted, however, that a major drawback of a structural dynamic analysis is that it typically yields no analytical results. Consequently, one has to settle for numerical solutions, which are quite sensitive to parameter specifications.

The strategic dynamic approach as discussed in this chapter is more suited for obtaining analytical solutions. In order to provide a rational basis for the existence of incomplete cartels one should take into account both the incentive problem and the participation problem of firms. As noted, this at a minimum requires a dynamic setting in which firms differ in at least one respect. However, introducing firm heterogeneity in a repeated-game setting significantly complicates the analysis. The study of incomplete cartels forms no exception. As remarked by Motta (2004, p. 181), “Unfortunately, the analysis of partial collusion raises several difficulties, as one should model a situation where a group of firms collude whereas others simply best respond. Solving such a model analytically is not easy, and further work is needed on this issue.” In the next chapter, we make an attempt in this direction.

3

A Theory of Incomplete Cartels with Heterogeneous Firms

“There are two kinds of people: those who do the work, and those who take the credit. Try to be in the first group; there is less competition there.”
– Indira Gandhi¹

3.1 Introduction

Most economic theories on collusion predict that the optimal cartel arrangement is one in which all sellers in the industry participate.² To explain how firms could fail to eliminate all market competition it is typically argued that an all-inclusive cartel might not be stable and therefore may be too high of a goal. The potential instability is due to the incentive of firms to cheat on a cartel. Such cheating may occur both *ex ante* and *ex post*. Cheating *ex ante* refers to the incentive of firms to free-ride on a cartel formed by competitors. The output reduction and price increase of cartel members allows nonconspirators to raise their prices and expand their production, which yields windfall profits. The incentive to cheat *ex post* results from the fact that a cartel member commits to a production level for which marginal revenue exceeds marginal cost. As a consequence, all members feel a temptation to increase outputs. Yet, if every firm expects its rivals to collude, no cartel will emerge. Cartels are therefore the result of a complex trade-off between, on the one hand, incentives to collude and, on the other hand, incentives to cheat.

¹See <http://quotationsbook.com>.

²This chapter is in part based on joint work with J.E. Harrington. See Bos and Harrington (2008).

This chapter provides a rational basis for the existence of incomplete cartels. We contribute to the existing literature in at least two important ways. First, we focus on explicit cartel agreements as opposed to tacit collusion. Communication between firms makes that cartelizing is costly and taking into account these cost of colluding could provide an additional explanation for why many discovered cartels were less than all-inclusive. The intuition is that excluding one or more firms from the cartel may significantly reduce the cost of colluding so that cartel profits per member are higher for an incomplete cartel. Second, we take into account both the incentive to cheat *ex ante* and the incentive to cheat *ex post*. Indeed, the main challenge in building a theory of incomplete cartels is to model a situation in which a group of sellers finds it in their interest to collude, whereas others prefer to remain outside the cartel. That is to say, in order to understand how incomplete cartels emerge, a subset of firms must have no incentive to cheat, neither *ex ante* nor *ex post*, and the remaining firms should lack the incentive to join the cartel. At a minimum, this requires a model in which firms face both an incentive problem and a participation problem. In addition, we want firms to differ in at least one respect in order to address the question: What (type of) firms take part in a cartel and what (type of) firms remain independent outsiders?

To that end, we analyze a price setting supergame in which firms differ in terms of capacity stock, which is taken as a proxy for firm size.³ There are two reasons for choosing this setting. First, it is a simple and convenient way to deal with firm heterogeneity. In particular, it allows us to derive analytical solutions that are difficult to obtain in other type of models. Second, the descriptive cartel studies surveyed in Chapter 2 suggest that firm size is an important determinant in a firm's decision on whether or not to take part in a cartel. There exist quite a few studies that analyze collusion in price setting supergames with capacity constraints. Examples include Brock and Scheinkman (1985), Benoit and Krishna (1987), Staiger and Wolak (1992) and more recently Compte *et al.* (2002).⁴ All these studies assume the cartel to be all-inclusive. In the supergame literature on collusion, this assumption can sometimes be defended on the ground that a full cartel is sustainable whenever some collusion is sustainable. In more technical terms, the critical discount factor might be decreasing in the number of cartel participants, all else equal.⁵ However, whether or not firms have an incentive to form an industry-wide cartel is a different matter. The main difference with our analysis is that these studies focus exclusively on the incentive problem of the cartel and that only opportunities to collude tacitly are considered.⁶

³In this chapter, the terms capacity stock, capacity and production capacity are used interchangeably. All refer to the maximum number of products that a firm can supply in a given period.

⁴Brock and Scheinkman (1985) consider the role of production capacity in enforcing collusion among a fixed number of identical firms. They establish that there exists no monotone relationship between the number of firms and the maximum sustainable profits. Benoit and Krishna (1987) allows firms to choose both capacity and price. Among other things, they find that colluding firms carry excess capacity. Staiger and Wolak (1992) introduce stochastic demand in this type of setting and find support that excess capacity may lead collusion to breakdown. See Chapter 2 of this dissertation for a discussion of Compte *et al.* (2002).

⁵See Chapter 2 of this thesis.

⁶The standard supergame literature on collusion often fails to discriminate between tacit and overt collusion. One convenient and popular way to tailor the analysis to explicit cartel agreements is by

The profound questions to be addressed are (*i*) under what conditions is the optimal cartel size less than all-inclusive?, (*ii*) given that the cartel is not all-inclusive, what (type of) firms have the strongest incentive to join a cartel?, (*iii*) do firms have an incentive to form the cartel for which total cartel profits are highest?, and (*iv*) how does a change in the size distribution of firms due to a merger affect the impact of a cartel?

The price setting supergame with asymmetric capacity constraints that we consider is closest to the model analyzed in Compte *et al.* (2002).⁷ Our analysis differs in at least three aspects. First, they exclusively focus on how firm heterogeneity affects the sustainability of an *all-inclusive* cartel and free-riding incentives are therefore not considered. As documented in the previous chapter, there exists a separate strand of literature that focuses on the participation problem of firms. In order to take account of the incentives of firms to free-ride on a cartel, we extend the price setting supergame with a participation stage. Second, although we do analyze the possibility that cartels emerge tacitly, our main focus is on explicit cartel agreements.⁸ Third, Compte *et al.* (2002) analyze how a redistribution of capacity, for example due to a merger, affects the sustainability of an all-inclusive cartel. Instead, we examine the relationship between mergers and the composition of the (incomplete) cartel. For example, two firms not part of a cartel might find it in their interest to join the cartel post-merger. The main question therefore is what type of mergers have the strongest coordinated effects.

The main results of our analysis are as follows. We find that the optimal cartel size might not be all-inclusive when colluding is costly and the smallest firms are sufficiently small. We establish that the incentive to take part in a cartel is positively related to firm size and that very small firms have no incentive to take part in any cartel. Also, the most profitable cartel is formed by the largest firms in the industry and firms do have an incentive to form this cartel whenever its smallest member is sufficiently large. Furthermore, we show that sellers lack the incentive to merge in absence of collusion. In particular, firms only have an incentive to engage in a merger when they are part of the cartel post-merger. Our analysis further suggests that the strongest coordinated effects may come from a merger between moderate-sized firms.

The chapter proceeds as follows. Section 2 presents the model and derives some benchmark results. Section 3 explores what is the optimal cartel size both under the assumption that colluding is costless and when colluding is costly. Here, we do not yet consider the incentives of firms to take part in a cartel. Section 4 extends the analysis of the previous section by introducing a participation stage. In this section, we examine what type of firms have a stronger incentive to collude. In Section 5, we

introducing extra costs in the model (e.g., cartels that require members to communicate are subject to antitrust enforcement and participants therefore face an additional cost equal to the probability of being discovered times the level of punishment.). See, for instance, McCutcheon (1997), in which it is assumed that collusion requires communication, but the communication process is not modelled explicitly.

⁷The main difference with their model is that we impose an upper bound on the production capacity of the largest firm(s). We further assume demand to be downward-sloping instead of unit demand.

⁸It must be noted, however, that the model presented in this chapter in principle applies equally well to tacit collusion.

examine if and when firms have an incentive to form the cartel for which total cartel profits are highest. In Section 6, we explore how the change in the size distribution of firms due to a merger affects collusion. We analyze merger incentives and potential coordinated effects of a merger. Section 7 concludes.

3.2 A Model of Collusion with Asymmetric Capacity Constraints

In a given industry, let N denote the set of firms containing a fixed number of n profit-maximizing sellers that simultaneously set prices, which is assumed to be their single strategic variable. Commodities are homogeneous and produced at common unit cost $c > 0$. For simplicity, fixed costs are normalized to zero. The market demand function is denoted $D(p)$, which we assume to be twice continuously differentiable with $D'(p) < 0$ and $D''(p) \leq 0$. We naturally suppose that society values the production of the first unit, i.e., $D'(c) > 0$. Firms differ in terms of capacity stock. The production capacity of firm i is denoted k_i and is taken to be fixed for all $i \in N$. Without loss of generality, we index the firms so that,

$$k_1 \geq k_2 \geq \dots \geq k_n.$$

Let firm i 's demand be $D_i(p_i, \mathbf{p}_{-i})$, $\forall i \in N$, which is a function of the price charged by firm i (p_i) and the prices set by all its rivals (\mathbf{p}_{-i}). In what follows, let $\Delta(p) \equiv \{j \in N : p_j < p\}$ denote the set of sellers that price strictly below p and let $\Omega(p) \equiv \{j \in N : p_j = p\}$ be the set of firms that price at p . Assuming the efficient rationing rule, we impose the following condition on individual demand,

$$D_i(p_i, \mathbf{p}_{-i}) \leq \max \left\{ D(p_i) - \sum_{j \in \Delta(p_i)} k_j, 0 \right\}, \forall i \in N.$$

Basically, this condition ensures that customers prefer cheaper products. In particular, it implies that firm i has positive demand only when firms that price below p_i have excess demand.

We make several additional assumptions on firm demand.

- Assumption 1:

- (i) If $0 < D(p) - \sum_{j \in \Delta(p)} k_j < \sum_{j \in \Omega(p)} k_j$, then $0 < D_i(p_i, \mathbf{p}_{-i}) < k_i$, $\forall i \in \Omega(p)$.
- (ii) If $\sum_{j \in \Omega(p)} k_j < D(p) - \sum_{j \in \Delta(p)} k_j$, then $k_i < D_i(p_i, \mathbf{p}_{-i})$, $\forall i \in \Omega(p)$.

Assumption 1(i) means that if firms, which price at p , have combined capacity that exceeds residual demand, then demand is allocated such that all firms that price at p have positive demand and excess capacity. Assumption 1(i) therefore imposes a lower and upper bound on the production level of firms that charge the same price. Note that,

within these boundaries, market shares of firms can vary substantially. Assumption 1(ii) mirrors the first part. It states that, when residual demand exceeds combined capacity of firms that price at p , demand is allocated such that all firms that price at p produce up to capacity. Assumption 1 therefore imposes some symmetry across firms.

The next two assumptions put some limits on the market power of the largest firm(s).

- Assumption 2: $D(p^m) > k_i, \forall i \in N$, i.e., none of the firms has sufficient capacity to meet monopoly demand $D(p^m)$.
- Assumption 3: $\sum_{j \in N \setminus \{i\}} k_j \geq D(c), \forall i \in N$, i.e., any combination of $n - 1$ firms has combined capacity that is sufficient to satisfy maximum demand.

Arguably, the last two assumptions are somewhat restrictive, because it rules out the possibility of firms being very large in *absolute* terms. Yet, it still allows firms to be very large in *relative* terms. Also, an immediate consequence of Assumption 2 in conjunction with Assumption 3 is that duopolistic market structures are excluded. However, the condition $n \geq 3$ is only mildly restrictive, because our focus is on cartels that are not all-inclusive.

For technical reasons, firms will choose a price from the set,

$$\{0, \varepsilon, \dots, c - \varepsilon, c, c + \varepsilon, \dots\},$$

with ε denoting the smallest monetary unit, which we typically assume to be sufficiently small.⁹

3.2.1 Static Nash Equilibrium

In competition, firm i 's profit function is given by,

$$\pi_i = (p_i - c) D_i(p_i, \mathbf{p}_{-i}). \quad (3.1)$$

The following result establishes the static Nash equilibrium of the game.

Proposition 3.1 *As $\varepsilon \rightarrow 0$, all firms price at marginal cost in competition.*

Proof. Suppose that all firms price at marginal cost, which implies $\pi_i = 0, \forall i \in N$. Given that $p_j = c, \forall j \neq i$, firm i may attempt to increase its profits by charging $p_i \neq c$. Setting $p_i < c$ would yield $\pi_i < 0$, which is clearly not an improvement. Alternatively, firm i may set a price $p_i > c$, but this means that firm i will lose all its customers due to Assumption 3, which implies $\pi_i = 0$. Hence, $p_i = c, \forall i \in N$, constitutes a (symmetric) Nash equilibrium of the game.

Next, suppose without loss of generality that $p_i \geq p_j, \forall j \in N$ and $i \neq j$ and further assume that all firms charge a price that is weakly higher than $c + \varepsilon$. If $p_i > c + \varepsilon$, then

⁹The reason for choosing a discrete price distribution instead of a continuous price distribution is that with the latter the best response functions of firms is not well-defined.

$D_i = 0$ or $D_i > 0$. If $D_i = 0$ firm i earns zero profits, which is clearly not optimal. For instance, $p_i = c + \varepsilon$ would always yield $\pi_i > 0$ due to Assumption 1. We therefore focus on the situation in which $p_i > c + \varepsilon$ and $D_i > 0$. Then, in light of Assumption 3 there must exist at least one firm j that charges $p_j \geq p_i$, which yields a contradiction unless $p_i = p_j$. However, given that firm i does not produce up to capacity at p_i , it can increase its demand by charging $p_i - \varepsilon < p_j$, which yields higher profits for ε sufficiently small. To show that firm i does not experience excess demand at p_i we will assume the opposite and derive a contradiction. Suppose that firm i does produce up to capacity at p_i . Then, in light of Assumption 1 (ii), all firms that price at p_i face excess demand. That is, $D(p_i) - \sum_{j \in \Delta(p_i)} k_j \geq \sum_{j \in \Omega(p_i)} k_j \implies D(p_i) \geq \sum_{j \in N} k_j$, which violates Assumption 3.

We have established that all firms charge a price strictly higher than $c + \varepsilon$ with zero probability. Note that a further reduction in price to c is not optimal, because this would yield zero profits, while at $c + \varepsilon$ profits are positive. All firms charging a price equal to $c + \varepsilon$ therefore constitutes a (symmetric) Nash equilibrium. It was already shown that $p = c$ constitutes an equilibrium. Yet, as $\varepsilon \rightarrow 0$, this difference disappears, which leaves $p_i = c, \forall i \in N$, as the competitive Nash equilibrium. ■

The above result implies that in competition none of the sellers is making economic profits. Clearly, it is in the interest of all undertakings to reduce competitive pressure and raise industry prices. Denote $\Gamma \subseteq N$ a (sub)set of firms that form a price-fixing cartel. The cartel is assumed to operate as a multiplant unit, which sets a single cartel price $p^c > c + \varepsilon$. Obviously, such a coalition is profitable only if cartel demand is strictly positive. This is naturally the case for an all-inclusive cartel. If the cartel is incomplete, however, its demand will depend in part on the characteristics and behavior of independent outsiders. The next result establishes the optimal behavior of firms that do not take part in the price-fixing conspiracy.

Lemma 3.1 *If a profitable cartel sets a price $p^c > c + \varepsilon$, then the best response of an individual profit-maximizing outsider j is to sell k_j units at a price $p_j^o = p^c - \varepsilon$, for ε sufficiently small.*

Proof. First, suppose that the cartel price is the (weakly) lowest price in the industry and consider the case in which $p_i^o > p^c, \forall i \in N \setminus \Gamma$. Without loss of generality, let p_j^o denote the highest price in the industry. Then, following Assumption 3, firm j has zero demand if it charges the strictly highest price, which is clearly not optimal, e.g., charging $p_j^o = p^c$ would yield $\pi_j^o > 0$. Suppose therefore that $D_j > 0$. In light of Assumption 3, this implies that there exists at least one other firm s that sets $p_s^o \geq p_j^o$ and therefore $p_s^o = p_j^o$. Note that, due to Assumption 1, all firms that price at p_j^o face positive demand. Yet, given that firm j does not produce up to capacity at p_j^o , reducing the price slightly to $p_j^o - \varepsilon$ yields a discrete increase in demand, which is profitable for ε sufficiently small. To show that firm j does not produce up to capacity at p_j^o we will assume the opposite and derive a contradiction. Suppose firm j does fully utilize its capacity at p_j^o . Then, by Assumption 1 (ii) we know that all firms that price at p_j^o produce up to capacity. This implies $D(p_j^o) - \sum_{i \in \Delta(p_j^o)} k_i \geq \sum_{i \in \Omega(p_j^o)} k_i \implies D(p_j^o) \geq \sum_{i \in N} k_i$, which violates Assumption 3. Hence, we have shown that all outsiders charge a price that is strictly higher than p^c with zero probability.

Now suppose that all outsiders price at p^c . Then, Assumption 3 in conjunction with Assumption 1 (i) implies that all firms face a positive demand and have excess capacity. Firm j therefore earns profits equal to $\pi_j^o = (p^c - c) D_j(p^c)$, with $D_j(p^c) < k_j$. However, by reducing its price to $p^c - \varepsilon$ it can earn $\pi_j^o = (p^c - \varepsilon - c) k_j$, which is larger for ε sufficiently small. Charging $p_i^o = p^c - \varepsilon, \forall i \in N \setminus \Gamma$, is therefore optimal if all outsiders produce up to capacity at $p^c - \varepsilon$, i.e., if $D(p^c - \varepsilon) \geq \sum_{i \in \Omega(p^c - \varepsilon)} k_i$. This must necessarily hold, because if $D(p^c - \varepsilon) < \sum_{i \in \Omega(p^c - \varepsilon)} k_i$ cartel demand would be zero, which means that the cartel is not profitable. Note that a further reduction in price by fringe members is not optimal, because all outsiders produce up to capacity at $p^c - \varepsilon$. ■

Lemma 3.1 implies that a cartel Γ is profitable only if $D(p^c) > \sum_{j \notin \Gamma} k_j$, which is the case when outsiders do not have sufficient capacity to meet total market demand. In other words, the only viable cartels are those that have sufficient control over industry capacity. Also, as $\varepsilon \rightarrow 0$, fringe firms optimally charge the cartel price. Note the close similarity with the approach taken in collusive price leadership models in which it is typically assumed that outsiders take the cartel price as given.

As a consequence, total cartel profits are given by,

$$\pi^c = (p^c - c) D^c = (p^c - c) \left[D(p^c) - \sum_{j \notin \Gamma} k_j \right]. \quad (3.2)$$

The cartel has to divide π^c among its members according to some profit sharing rule. Suppose therefore that the cartel establishes a profit allocation rule α . Given a capacity vector (k_1, k_2, \dots, k_n) , a cartel Γ and cartel price $p^c(\Gamma)$, α prescribes an allocation of cartel demand. Hence, every firm $i \in \Gamma$ receives a share $\alpha_i \in (0, 1)$ of total cartel profits, i.e.,

$$\pi_i^c = (p^c - c) \left[D(p^c) - \sum_{j \notin \Gamma} k_j \right] \alpha_i, \forall i \in \Gamma. \quad (3.3)$$

We require α to be efficient in the sense that cartel members allocate all cartel profits among themselves, i.e., $\sum_{i \in \Gamma} \alpha_i = 1$.

3.2.2 Infinitely Repeated Game

In the following, we will analyze the infinitely repeated version of this game. The collusive value of firm $i \in \Gamma$ is given by,

$$V_i^c(p^c, \Gamma) = \sum_{t=1}^{\infty} \delta^{t-1} \pi_i^c,$$

with t indicating the date of a single period and $\delta \in (0, 1)$ denoting the common discount factor. We assume that collusive arrangements are sustained through grim-trigger strategies. That is, every member of Γ adheres to the collusive strategy until one firm deviates. In the event of cheating, the coalition collapses with a one-period time

lag, i.e., all firms compete in all periods following a period of defection.¹⁰ Consequently, for collusion to be sustainable as a subgame perfect Nash equilibrium (SPNE) of the infinitely repeated game, the following incentive compatibility constraint must be satisfied for all participants,

$$V_i^c(p^c, \Gamma) \geq \pi_i^d, \quad (3.4)$$

with π_i^d denoting the maximum one-period gain from defection. The following result determines the maximum profit to be obtained by a chiseling firm during the period of defection.

Lemma 3.2 *As $\varepsilon \rightarrow 0$, $\pi_i^d = (p^c - c) k_i$.*

Proof. Consider a cartel that sets a price $p^c > c + \varepsilon$. By Lemma 3.1 we know that if the cartel is not all-inclusive outsiders set a price $p_j^o = p^c - \varepsilon, \forall j \in N \setminus \Gamma$. A chiseling firm i aims to maximize π_i^d . If it deviates by charging $p_i^d > p^c$, then its demand would drop to zero due to Assumption 3, which is clearly not optimal. Instead, suppose firm i reduces its price to $p_i^d = p^c - \varepsilon$. This defection strategy is optimal if firm i sells up to capacity at $p^c - \varepsilon$. In this case it would earn $\pi_i^d = (p^c - \varepsilon - c) k_i$. Note that $\varepsilon \rightarrow 0$ yields $\pi_i^d = (p^c - c) k_i$. Alternatively, firm i may have excess capacity at $p^c - \varepsilon$. Therefore, given that ε is sufficiently small, it could be profitable to reduce p_i^d further to $p_i^d = p^c - 2\varepsilon$, which would be the lowest price in the industry. Assumption 2 guarantees that firm i fully utilizes its capacity at this price. That is, $D(p^m) \geq k_i$ implies $D(p^c - 2\varepsilon) \geq k_i$, because $p^m \geq p^c$. Hence, a further reduction of p_i^d is never optimal. However, as $\varepsilon \rightarrow 0$, we have that $p^c - 2\varepsilon = p^c$. ■

Note that as $\varepsilon \rightarrow 0$, $\pi_i^d = \pi_i^o$. We therefore might say that fringe firms permanently and optimally cheat on the cartel, without being punished. In other words, from the start of the game, outsiders behave *as if* they optimally defect from the cartel agreement in every period. Furthermore, $\pi_i^d > \pi_i^c$ for all $i \in \Gamma$ whenever $\sum_{i \in N} k_i > D(p^c)$, which due to Assumption 3 always holds. Nevertheless, it is well-known that, despite the potential instability, collusion may be an equilibrium strategy if the interest rate that is used to discount future profit streams is sufficiently low.¹¹ Using Lemma 3.2, (3.4) can be rearranged to,

$$\delta \geq \delta_i^* = 1 - \alpha_i \frac{\pi_i^c}{\pi_i^d} = 1 - \frac{\alpha_i}{k_i} \left[D(p^c) - \sum_{i \in N} k_i + \sum_{j \in \Gamma} k_j \right]. \quad (3.5)$$

Observe that δ_i^* is decreasing in α_i and increasing in k_i , all else unchanged. Clearly, a member has a lower incentive to deviate the larger the share of total cartel profits it receives. Also, given some profit allocation rule, an increase in production capacity

¹⁰Note that the grim-trigger strategy is the most severe *credible* threat. Hence, whenever some level of collusion cannot be sustained by the threat of eternal competition it cannot be sustained by any other credible punishment strategy. For a detailed analysis of optimal penal codes in price-setting supergames the reader is referred to Lambson (1987, 1994).

¹¹The seminal work is due to Friedman (1971).

yields a stronger incentive to deviate, because larger capacity allows a firm to earn more during a single period of defection. In addition, (3.5) reveals that a cartel is only viable if it has control over at least $\sum_{i \in N} k_i - D(p^c)$ of total industry capacity. Finally, for sufficiently high δ , none of the firms can prevent its rivals from forming a cartel that is effective. This is due to Assumption 2, which ensures that none of the suppliers has sufficient capacity to meet $D(p^c)$.

Notice that the individual profits of conspirators increase with total cartel value. The objective of a coalition is therefore to choose a cartel price that maximizes total cartel profits, while satisfying (3.5). Hence, the optimal cartel price of a cartel Γ is defined by the following constraint optimization problem,

$$\max_{p^c} V^c(p^c, \Gamma) = \frac{1}{1 - \delta} (p^c - c) \left[D(p^c) - \sum_{i \in N} k_i + \sum_{j \in \Gamma} k_j \right], \quad (3.6)$$

subject to,

$$\frac{\alpha_i}{k_i} \left[D(p^c) - \sum_{i \in N} k_i + \sum_{j \in \Gamma} k_j \right] - (1 - \delta) \geq 0, \forall i \in \Gamma.$$

The next result shows that the solution to (3.6) is increasing in the combined capacity of colluders.

Lemma 3.3 Fix $\frac{\alpha_i}{k_i}$, $\forall i \in \Gamma$. The optimal cartel price is strictly increasing in total cartel capacity.

Proof. Consider a sustainable cartel Γ and for notational convenience, let $K^c = \sum_{j \in \Gamma} k_j$. The incentive compatibility constraint (3.5) may or may not be binding for one or more members. First, suppose it is not binding for any member. The optimal cartel price is then defined by,

$$D(p^c) - \sum_{i \in N} k_i + K^c + (p^c - c) D'(p^c) = 0.$$

Rearranging yields,

$$K^c = \sum_{i \in N} k_i - D(p^c) - (p^c - c) D'(p^c). \quad (3.7)$$

The first-order derivative of (3.7) with respect to p^c is,

$$\frac{\partial K^c}{\partial p^c} = -2D'(p^c) - (p^c - c) D''(p^c).$$

In order to evaluate how a change in total cartel capacity affects the optimal cartel price we take the inverse,

$$\frac{\partial p^c}{\partial K^c} = -\frac{1}{2D'(p^c) + (p^c - c) D''(p^c)} > 0.$$

Now suppose (3.5) is binding for at least one member, which means that the cartel price is chosen such that $\delta = \max \delta_i^*$, $i \in \Gamma$, i.e., the firm(s) for which the ratio $\frac{\alpha_i}{k_i}$ is smallest. Hence,

$$\delta = 1 - \frac{\alpha_i}{k_i} \left[D(p^c) - \sum_{i \in N} k_i + K^c \right].$$

Rearranging yields,

$$K^c = \frac{k_i}{\alpha_i} (1 - \delta) + \sum_{i \in N} k_i - D(p^c). \quad (3.8)$$

Taking the first-order derivative of (3.8) with respect to p^c yields,

$$\frac{\partial K^c}{\partial p^c} = -D'(p^c).$$

Taking the inverse gives,

$$\frac{\partial p^c}{\partial K^c} = -\frac{1}{D'(p^c)} > 0.$$

Notice that the function $p^c(K^c)$ is continuous, but might be non-differentiable. It can be concluded that the larger the share of industry capacity that is under the control of a cartel, the higher will be the optimal cartel price. ■

To what extent collusion yields higher industry prices therefore essentially depends on *total* cartel capacity relative to *total* fringe capacity, which, given industry size, are ‘communicating vessels’. Clearly, industry prices are highest when all firms take part in the cartel arrangement.

3.3 Optimal Cartel Size

In the previous section, it has been shown that any cartel that controls a large enough share of industry capacity is feasible provided that its members are sufficiently patient. In this section, we explore what is the optimal cartel size. We define optimal cartel size as follows.

Definition 3.1 *The size of a cartel is optimal when: (i) the cartel does not find it profitable to include additional firms, and (ii) no subset of participants finds it in their interest to exclude one or more members.*

At this point, it is important to emphasize that any result depends quite heavily on the profit allocation rule α . In particular, total cartel profits must be allocated in such a way that the cartel under consideration is sustainable. One way to think about this problem is by assuming that firms attempt to form a cartel with the support of a cartel manager (e.g., a ringleader). The prime task of the cartel manager is then to design an optimal cartel contract by deciding on a profit allocation rule.

In the following, we distinguish between a situation in which cartelizing is costless and a situation in which cartelizing is costly. By ‘costless’ we mean that, apart

from opportunity costs, membership of a price-fixing coalition is free. Hence, cartel members do not incur costs associated with the formation and management of the cartel. Furthermore, expected costs due to antitrust enforcement are assumed absent. By contrast, these type of costs are present when cartelizing is costly.

3.3.1 Costless Collusion

In this subsection, we investigate what cartel size is optimal under the assumption that cartelizing is costless. A cartel considered here is therefore *de facto* a tacit agreement among firms.

In a wide variety of settings, cartel profits are positively correlated with the size of the cartel.¹² Therefore, we might *a priori* expect the cartel manager to maximize the number of participants. However, fringe profits also tend to increase with cartel size. As a result, we may suspect that, *ceteris paribus*, it is increasingly difficult to expand the size of the cartel without rendering the cartel unstable. Yet, as the following result shows, larger cartels are feasible given that the total cartel value is allocated properly.

Theorem 3.1 *A cartel can always allocate its profits in such a way that it Pareto dominates every smaller cartel.*

Proof. Consider a stable cartel Γ consisting of x members that were willing to join given some profit sharing rule. Consequently, $n - x$ outsiders did not want to join given this profit sharing rule. For a firm $i \in N \setminus \Gamma$ to join the cartel it must be offered an amount that exceeds its current earnings. Following lemma 3.1, an outsider i earns,

$$\frac{1}{1 - \delta} (p^c(\Gamma) - \varepsilon - c) k_i.$$

Suppose therefore that it is offered the following amount,

$$\frac{1}{1 - \delta} (p^c(\Gamma) - c) k_i.$$

In what follows, let $\Psi \subseteq N \setminus \Gamma$ denote a (sub)set of outsiders and let $\Gamma_{+\Psi}$ indicate the cartel consisting of all members of Γ and Ψ . In order for all $j \in \Psi$ to join, they must be offered a total amount that is equal to,

$$\frac{1}{1 - \delta} (p^c(\Gamma) - c) \sum_{j \in \Psi} k_j.$$

After paying these fringe members, there must be a sufficient amount of money left to cover, at a minimum, the total earnings of Γ , which amount to,

$$\frac{1}{1 - \delta} (p^c(\Gamma) - c) \left[D(p^c(\Gamma)) - \sum_{i \in N} k_i + \sum_{j \in \Gamma} k_j \right].$$

¹²Note that there does not always exist a monotonic relationship between cartel size and cartel profitability. Consider, for example, a standard symmetric Cournot setting with linear demand and constant unit costs. In such a setting, the ‘empty cartel’ is more profitable than any coalition that spans less than 80% of the market. See Salant *et al.* (1983).

Hence, including outsiders is (weakly) profitable only if the following condition holds,

$$\frac{1}{1-\delta} (p^c(\Gamma_{+\Psi}) - c) \left[D(p^c(\Gamma_{+\Psi})) - \sum_{i \in N} k_i + \sum_{j \in \Gamma} k_j + \sum_{j \in \Psi} k_j \right] \geq$$

$$\frac{1}{1-\delta} (p^c(\Gamma) - c) \left[D(p^c(\Gamma)) - \sum_{i \in N} k_i + \sum_{j \in \Gamma} k_j \right] + \frac{1}{1-\delta} (p^c(\Gamma) - c) \sum_{j \in \Psi} k_j,$$

which is equivalent to,

$$(p^c(\Gamma_{+\Psi}) - c) \left[D(p^c(\Gamma_{+\Psi})) - \sum_{i \in N} k_i + \sum_{j \in \Gamma} k_j + \sum_{j \in \Psi} k_j \right] \geq \quad (3.9)$$

$$(p^c(\Gamma) - c) \left[D(p^c(\Gamma)) - \sum_{i \in N} k_i + \sum_{j \in \Gamma} k_j + \sum_{j \in \Psi} k_j \right].$$

It is *a priori* unclear if this condition holds generally. Suppose however that the new cartel $\Gamma_{+\Psi}$ is not changing its price, i.e., $p^c(\Gamma_{+\Psi}) = p^c(\Gamma)$. Then, (3.9) holds with equality. It remains to be shown that such a cartel is indeed sustainable. Using (3.4), the incentive compatibility constraint of a newcomer is given by,

$$\frac{1}{1-\delta} (p^c(\Gamma) - c) k_i \geq (p^c(\Gamma) - c) k_i,$$

which is satisfied for all values of δ . After paying all new participants, the old cartel members have the following pie to share,

$$\frac{1}{1-\delta} (p^c(\Gamma) - c) \left[D(p^c(\Gamma)) - \sum_{i \in N} k_i + \sum_{j \in \Gamma} k_j \right].$$

This is equal to the value of the old cartel Γ , which was given some profit sharing rule sustainable.

To complete the proof, it is important to note that with $p^c(\Gamma_{+\Psi}) = p^c(\Gamma)$, (3.9) only holds with equality and that none of the members of the old cartel strictly benefits from the cartel expansion. However, unlike $p^c(\Gamma)$, $p^c(\Gamma_{+\Psi}) = p^c(\Gamma)$ is not the optimal cartel price for the new cartel $\Gamma_{+\Psi}$. In fact, following Lemma 3.3, a cartel $\Gamma_{+\Psi}$ would optimally set $p^c(\Gamma_{+\Psi}) > p^c(\Gamma)$, which implies that the left-hand side of (3.9) will be strictly larger than the right-hand side. ■

The result that a cartel Pareto dominates any smaller cartel is non-trivial and depends in part on the structure of the model. In particular, the result does not necessarily hold when firms compete in quantity (strategic substitutes) instead of price (strategic complements). To see this, consider the following example.

Example 3.1 Consider a standard Cournot market with n identical firms that produce a homogeneous product. For simplicity, production costs are normalized to zero and there are no fixed costs. Let the linear inverse demand function be given by,

$$P = 1 - Q,$$

where $Q = \sum_{i=1}^n q_i$ is total market output.

Now suppose there exists a cartel of x members that attempts to include t outsiders. In this setting, Pareto-improving cartel expansion requires,

$$\frac{1}{(n - x - t + 2)^2} \geq \frac{t + 1}{(n - x + 2)^2},$$

which does not always hold (e.g., this condition is violated for $n = 10$, $x = 8$ and $t = 1$).

Hence, when firms compete in quantity it does not always pay to expand cartel size. In Cournot competition without capacity constraints, outsiders benefit a lot from the collusive behavior of rivals. Consequently, firms have a strong incentive to free-ride on a cartel formed by rivals. The incentive to free-ride is prevalent within our model, but it is mainly caused by the ‘umbrella effect’ of cartel pricing. In the absence of capacity constraints, we may expect external suppliers not only to increase prices, but also to expand fringe production. In principle, therefore, fringe profits might be too high in that, even when colluding is costless, it would be too costly to include a subset of outsiders in the cartel.

The implication of the above result is less surprising. The unique Pareto efficient collusive equilibrium is the one in which all firms take part in the price-fixing agreement.

Corollary 3.1 *The optimal cartel size is all-inclusive.*

This result holds quite generally, because the profits of an all-inclusive cartel are typically higher than the combined profits of an incomplete cartel and fringe members. The above analysis suggests that the way in which total cartel profits are allocated plays a significant role in designing this optimal anticompetitive arrangement. Establishing a profit sharing rule acceptable to all parties is unlikely to evolve naturally and typically requires direct communication between firms. However, when collusion is explicit, it is more natural to assume that cartelizing is costly.

3.3.2 Costly Collusion

A coalition comprising the entire industry yields the highest possible gain for sellers under the assumption that cartelizing is costless. Yet, the mere fact that the total pie must be divided in a particular way to achieve this outcome reveals that this assumption is restrictive. Establishing a profit sharing rule that is acceptable to all

participants requires at least one, but typically multiple negotiation rounds.¹³ Also, once a cartel becomes effective, members often have to monitor each other in order to ensure compliance with the agreement.¹⁴ In addition, the content of these discussions makes participating firms subject to antitrust enforcement. All these factors make that cartelizing is a costly exercise. Arguably, firms will take these (expected) costs into account when deciding on whether or not to join a cartel.

Not much is known about the magnitude of the cost of cartelizing, which is in large part due to the secret nature of anticompetitive organizations. Moreover, these costs possibly vary widely among cartels. Be that as it may, based on the factors mentioned above it appears reasonable to assume that the costs of cartelizing are increasing in the number of parties involved. Discussions are, *ceteris paribus*, easier in small groups than in large groups, e.g., because lines of communication are shorter. Also, diversity among participants is likely to be higher in larger groups, which is generally believed to adversely affect the probability of reaching consensus. In addition, we may suspect that (expected) costs that are created by antitrust law enforcement will increase in the number of participants. In its simplest form, the costs created by antitrust enforcement equal the probability of conviction times the level of the fine. We may conjecture that the risk of detection increases with the size of the cartel, because larger cartels are presumably “more visible”.¹⁵ Furthermore, many antitrust agencies adopted a policy to grant amnesty to undertakings that confess and prove their involvement in a cartel. Obviously, such a leniency program is effective only if the first firm to self-report receives the highest benefits and the probability of being the first is, *ceteris paribus*, decreasing in the number of participants. In sum, there is a variety of reasons for why we may expect the cost of colluding to be increasing in cartel size.

We will take account of the cost of colluding in the simplest possible way. Let the per-period cost of cartelizing be given by some function $T(x)$, with x being the number of cartel participants. Note that we assume T to be independent of firm size. The reason for this is that most costs associated with collusion do not seem to depend on the market position of cartel members. For example, it is difficult to see how the risk of a cartel member applying for leniency and costs associated with the formation of a cartel would depend on the production capacity of a firm. In light of the above discussion, we naturally assume $T(\cdot)$ to be strictly increasing in the number of cartel members. No additional assumptions are made to further define the shape of $T(\cdot)$. For example, it may have a concave shape if we believe that the marginal cost of colluding is decreasing in x . Alternatively, if we believe that including more firms in

¹³In addition, firms often make agreements about other factors, e.g., what to do if market conditions change? The study of the U.S. sugar cartel conducted by Genesove and Mullin (2001) illustrates how frequent communication might help firms to collude.

¹⁴Levenstein and Suslow (2006) and Harrington (2006) report on cartels that have installed monitoring mechanisms (e.g., a joint sales agency). Although, these investments are potentially quite costly, it still is beneficial for the cartel if it prevents punishment phases caused by (perceived) cheating.

¹⁵One could argue, however, that a partial cartel may be more visible than an industry-wide cartel, because with the latter there is no outside competition that can be used as a benchmark. I thank Jeroen Hinloopen for pointing this out. In this case there indeed might be a discontinuity in $T(\cdot)$. See Chapter 4 of this thesis for a discussion on how outside competition can be used in detecting (incomplete) cartels.

the coalition adds more than proportionally to the cost of colluding, we may think of $T(\cdot)$ as a convex function. For instance, the latter may be the case when negotiations become increasingly complex and when the rate of detection significantly increases with the number of participants.

To see how the cost of cartelizing affects the internal stability of a cartel, consider a firm $i \in \Gamma$. This firm will adhere to the agreement if,

$$\delta \geq \delta_i^* = 1 - \alpha_i \frac{\pi^c}{\pi_i^d} = 1 - \alpha_i \frac{(p^c - c) \left[D(p^c) - \sum_{i \in N} k_i + \sum_{j \in \Gamma} k_j \right] - T(\cdot)}{\pi_i^d}. \quad (3.10)$$

As can be observed, taking into account the cost of colluding tightens the incentive compatibility constraint for all $i \in \Gamma$, because $T(\cdot) > 0$ for all cartel sizes, but the ‘‘empty cartel’’. The set of collusive equilibria is therefore smaller compared to a situation in which colluding is costless, which is in line with expectations. Clearly, no cartel is viable for T sufficiently large and we therefore will restrict our attention to situations in which T is ‘not too high’.

The impact of the cost of colluding on optimal cartel size depends on a trade-off. On the one hand, both residual cartel demand as well as the optimal cartel price are increasing in the size of the cartel. On the other hand, every additional cartel member increases total cartel costs. What cartel size is optimal therefore depends on the magnitude of these two antagonistic forces. Due to the general structure of the current setting, the value of these effects cannot be determined explicitly. Nevertheless, as the following result indicates, if the smallest firm in the industry is sufficiently small, then the optimal cartel size is not all-inclusive.

Theorem 3.2 *If the smallest firm in the industry is sufficiently small, then the optimal cartel size is less than all-inclusive.*

Proof. We will compare the all-inclusive cartel, denoted Γ , with a cartel Γ_{-n} , which is the cartel consisting of the $n - 1$ (weakly) largest firms in the industry. First, note that whenever $\pi^c(\Gamma) > 0$, $\exists \delta$ for which Γ is sustainable. Clearly, the optimal cartel size is not all-inclusive if the full cartel is not sustainable. Suppose therefore that the following condition holds,

$$\delta \geq \delta_i^* = 1 - \alpha_i \frac{(p^m - c) D(p^m) - T(n)}{\pi_i^d}, \forall i \in \Gamma.$$

The total value of Γ_{-n} is given by,

$$\frac{1}{1 - \delta} [(p^c(\Gamma_{-n}) - c) [D(p^c(\Gamma_{-n})) - k_n] - T(n - 1)]. \quad (3.11)$$

This value exceeds the value of the full cartel if,

$$(p^c(\Gamma_{-n}) - c) [D(p^c(\Gamma_{-n})) - k_n] - T(n - 1) > (p^m - c) D(p^m) - T(n).$$

Rearranging yields,

$$T(n) - T(n-1) > (p^m - c) D(p^m) - (p^c(\Gamma_{-n}) - c) [D(p^c(\Gamma_{-n})) - k_n]. \quad (3.12)$$

It is *a priori* unclear if this inequality is satisfied or not, because the terms on both sides of the inequality sign are positive. However, as $k_n \rightarrow 0$, we have that $p^m = p^c(\Gamma_{-n})$ due to Lemma 3.3. Therefore, as $k_n \rightarrow 0$, the right-hand side of (3.12) equals zero, while the left-hand side of (3.12) is strictly positive because $T(x) > T(x-1)$ by assumption.

We have established that, for k_n sufficiently small, Γ_{-n} is more profitable than Γ . It remains to be shown that $\exists \delta$ for which Γ_{-n} is sustainable. Hence, we must show that the following condition holds for all $i \in \Gamma_{-n}$,

$$\delta \geq 1 - \alpha_i \frac{(p^c(\Gamma_{-n}) - c) [D(p^c(\Gamma_{-n})) - k_n] - T(n-1)}{\pi_i^d}.$$

This condition holds, because for $k_n \rightarrow 0$ it was shown that $\pi^c(\Gamma_{-n}) - \pi^c(\Gamma) > 0$, while π_i^d remains the same in both cases. Therefore, the maximum critical discount factor for Γ_{-n} is lower than for Γ , which was sustainable. ■

This result is intuitive and holds quite generally. In particular, the result does not depend on the size distribution of firms in the industry. For example, the optimal cartel size is also less than all-inclusive when firms are of equal size and sufficiently small. The intuition behind this result is that it does not pay to include a firm for which the marginal benefit to the cartel falls short of the marginal cost of colluding. Very small firms hardly contribute to cartel revenues, but at the same time are responsible for a nonnegligible part of the cost of colluding. As a result, the value of an incomplete cartel might exceed that of a full cartel when colluding is costly and the smallest firms are sufficiently small.

3.4 Incentives to Collude

It has been established that the optimal cartel size is not all-inclusive when colluding is costly and the smallest firms are sufficiently small. However, what cartel will emerge is a different matter. This typically depends on the incentives of firms to collude or to remain an independent outsider to a cartel formed by others instead. In this section, we explore what (type of) undertakings are likely to participate and what (type of) firms are likely to become fringe members.

Clearly, in competition all firms have an incentive to coordinate actions, because none of them is earning economic profits. At the same time, however, there exists a strong incentive for all sellers to wait in the hope that a cartel is formed by rivals. This incentive to free-ride is caused by the fact that cartel members do not fully utilize their capacity, while non-participants produce up to capacity and sell their products at approximately the same price. It has long been recognized that the incentive to

cheat *ex ante* makes it difficult to understand how incomplete coalitions emerge.¹⁶ This problem is particularly severe in symmetric settings, because it is difficult to see why identical firms would take non-identical decisions. However, in the current analysis firms are characterized by their capacity stock and we may conjecture this to cause a diversity in free-riding incentives.

In order to address this issue, we extend the model laid out in the previous sections by assuming that prior to the cartel there exists a period in which a cartel might be formed. More specifically, we assume that at $t = 0$ firms play an open membership game in which they simultaneously announce whether or not to join the coalition. In making this participation decision, a firm naturally will take into account how all sellers will adjust their behavior *ex post*, i.e., after it has made its choice.

Let π^c and π^o denote the profit of a cartel member and an outsider respectively. Formally, outsiders have no incentive to join a cartel Γ if the following ‘external stability condition’ is satisfied,

$$\pi_i^o(p^c(\Gamma)) \geq \pi_i^c(p^c(\Gamma_{+i})), \forall i \in N \setminus \Gamma, \quad (3.13)$$

with $p^c(\Gamma_{+i})$ being the price set by a coalition consisting of Γ and firm i . In a similar fashion, cartel members prefer to remain in the cartel if the following condition holds,

$$\pi_i^c(p^c(\Gamma)) \geq \pi_i^o(p^c(\Gamma_{-i})), \forall i \in \Gamma, \quad (3.14)$$

with $p^c(\Gamma_{-i})$ being the price set by a coalition consisting of Γ excluding firm i . We will refer to (3.13) in combination with (3.14) as the ‘participation constraint’. In addition, it is convenient to refer to the game analyzed so far as \mathcal{K} and to denote \mathcal{K}^E the extended version of \mathcal{K} , i.e., \mathcal{K}^E is \mathcal{K} including the participation stage. Hence, a cartel can be explained as an equilibrium outcome of \mathcal{K}^E only when (3.13), (3.14) and (3.10) are satisfied. As such, the participation constraint can be viewed *a refinement criterion*. Thus, for profitable coalitions, the incentive to participate implies sustainability, but not necessarily *vice versa*. That is to say, extending the game with a participation stage at $t = 0$ (weakly) narrows the set of subgame perfect equilibria of \mathcal{K} .

To make the participation decision, firm i must be able to determine $\pi_i^c(p^c(\Gamma_{+i}))$ *ex ante*, which is difficult because so far α was not defined explicitly. Hence, in order to analyze the incentives to collude we must specify how firms value the various coalitions. To that end, we suppose in the following that participants receive a proportional share of total cartel profits, i.e.,

$$\alpha_i = \frac{k_i}{\sum_{j \in \Gamma} k_j}, \forall i \in \Gamma. \quad (3.15)$$

There are at least two reasons for this assumption.¹⁷ First, capacity may be taken as a proxy for market share and there exists evidence that some cartels based their profit sharing rule on the market shares of members in years prior to the cartel.¹⁸

¹⁶See Stigler (1950).

¹⁷In Bos and Harrington (2008) it is shown that the proportional sharing rule can also be derived endogenously based on a Rawlsian notion of justice.

¹⁸See, for example, Harrington (2006).

Second, this allocation rule facilitates collusion in the sense that it minimizes the highest critical discount factor. In other words, if a cartel is not sustainable with a proportional allocation of profits, it cannot be sustained with any other profit sharing rule.

Typically, there exist multiple coalitions that can be sustained as subgame perfect equilibrium of \mathcal{K}^E and it is *a priori* unclear which of these will actually emerge. To begin, we might ask how the incentive to join a cartel relates to the production capacity of a firm. The following result shows that the incentive to collude is positively correlated with firm size.¹⁹

Theorem 3.3 *Assume $\delta > \frac{K-D(c)}{K_\Gamma}$ and consider $i, j \notin \Gamma$. If $k_j > k_i$ then: i) if firm i finds it optimal to join cartel Γ then so does firm j ; and ii) if firm j does not find it optimal to join cartel Γ then neither does firm i .*

Proof. Consider a cartel Γ comprising $x - 1$ firms and a firm that is not a member of Γ . If its capacity is k , it prefers to join cartel Γ iff

$$([p^c(K_\Gamma + k) - c][D(p^c(K_\Gamma + k)) - (K - K_\Gamma - k)] - T(x)) \left(\frac{k}{K_\Gamma + k} \right) \geq [p^c(K_\Gamma) - c]k,$$

where the left hand side expression is the stationary profit from joining the cartel and the right hand side is the stationary profit from remaining outside the cartel. This condition can be re-arranged to

$$\frac{[p^c(K_\Gamma + k) - c][D(p^c(K_\Gamma + k)) - (K - K_\Gamma - k)] - T(x)}{K_\Gamma + k} - [p^c(K_\Gamma) - c] \geq 0. \quad (3.16)$$

Take the derivative of the expression in (3.16) with respect to k ,

$$\begin{aligned} & p^{c'}(K_\Gamma + k) \left(\frac{D(p^c(K_\Gamma + k)) - (K - K_\Gamma - k)}{K_\Gamma + k} \right) + [p^c(K_\Gamma + k) - c] \left(\frac{1}{K_\Gamma + k} \right)^2 \times \\ & [D'(p^c(K_\Gamma + k))p^{c'}(K_\Gamma + k)(K_\Gamma + k) + (K_\Gamma + k) - D(p^c(K_\Gamma + k)) + (K - K_\Gamma - k)] \\ & + \frac{T(x)}{(K_\Gamma + k)^2} \\ & = p^{c'}(K_\Gamma + k) \left(\frac{1}{K_\Gamma + k} \right) \times \quad (3.17) \\ & [D(p^c(K_\Gamma + k)) - (K - K_\Gamma - k) + (p^c(K_\Gamma + k) - c)D'(p^c(K_\Gamma + k))] \\ & + \frac{[p^c(K_\Gamma + k) - c](K - D(p^c(K_\Gamma + k))) + T(x)}{(K_\Gamma + k)^2}. \end{aligned}$$

The second term in (3.17) is positive because $p^c(K_\Gamma + k) - c > 0$, $K - D(p^c(K_\Gamma + k)) > 0$ and $T(x) > 0$. Since $p^{c'}(K_\Gamma + k) > 0$ was established in the proof of Lemma 3.3, the first term in (3.17) is non-negative iff:

$$D(p^c(K_\Gamma + k)) - (K - K_\Gamma - k) + (p^c(K_\Gamma + k) - c)D'(p^c(K_\Gamma + k)) \geq 0. \quad (3.18)$$

¹⁹A similar result without the cost of colluding can be found in Bos and Harrington (2008).

(3.18) holds with equality when the ICC is not binding and with inequality when the ICC is binding. Hence, (3.17) is positive.

Suppose $k_j > k_i$. Since the expression in (3.16) is increasing in k , if (3.16) holds for firm i then it holds for firm j ; and if (3.16) does not hold for firm j then it does not hold for firm i . ■

Thus, the larger the production capacity of a firm, the stronger its incentive to take part in a cartel, all else unchanged. We can make a stronger claim by showing that very small firms never have an incentive to join a cartel. The following result shows that the participation constraint is always violated for firms that are sufficiently small.

Lemma 3.4 *A sufficiently small firm has no incentive to join any cartel.*

Proof. Consider a firm $r \in N \setminus \Gamma$, which has to decide on whether or not to join a cartel Γ that is formed by x members. Following the external stability condition as defined by (3.13), firm r has no incentive to join coalition Γ if,

$$(p^c(\Gamma) - c)k_r \geq \left[(p^c(\Gamma_{+r}) - c) \left[D(p^c(\Gamma_{+r})) - \sum_{i \in N} k_i + \sum_{j \in \Gamma} k_j + k_r \right] - T(x+1) \right] \frac{k_r}{\sum_{j \in \Gamma} k_j + k_r},$$

with $p^c(\Gamma_{+r})$ being the solution of (3.6) for the new cartel consisting of members of Γ and firm r . Rearranging yields,

$$k_r \leq \frac{(p^c(\Gamma_{+r}) - c) (\sum_{i \in N} k_i - D(p^c(\Gamma_{+r}))) + T(x+1)}{p^c(\Gamma_{+r}) - p^c(\Gamma)} - \sum_{j \in \Gamma} k_j. \quad (3.19)$$

Following Lemma 3.3 we have that $p^c(\Gamma_{+r}) = p^c(\Gamma)$ as $k_r \rightarrow 0$. Therefore, the right-hand side of (3.19) tends to infinity for ever smaller values of k_r , while the left-hand side of (3.19) tends to zero. ■

The intuition behind these results is as follows. Both residual cartel demand and the extent to which a cartel can raise industry prices depend in part on the combined capacity of its members. As Lemma 3.3 indicates, the more industry capacity that is under the control of the cartel, *ceteris paribus*, the higher is the optimal cartel price. If a small outsider joins the coalition, then the change in both cartel and fringe capacity is only marginal. Consequently, the marginal contribution of small sellers to the cartel is limited. These small firms prefer to remain independent outsiders, because as a fringe member they do not have the commitment to reduce production. By contrast, large fringe members leave little residual demand for a cartel. As a result, a cartel facing a competitive fringe with significant production capacity has limited power to raise the cartel price above marginal costs. In turn, this limits the incentive to free-ride for a large firm, because even though it would produce up to capacity, it would also face a relatively small mark-up. Instead, if a large firm decides to join the cartel it will sell fewer products, but at the same time taking part in the cartel allows for a substantial

price increase. Therefore, a firm will find it in its interest to join a conspiracy when it is sufficiently large.

The previous results indicate that, *ceteris paribus*, there exists a negative correlation between the incentive of firms to free-ride on a cartel formed by others and firm size. The next proposition shows that a cartel comprising the largest firms in the industry is indeed a solution of \mathcal{K}^E .

Proposition 3.2 *If δ is sufficiently close to one, then there exists $m \in \{2, \dots, n\}$ such that $\{1, \dots, m\}$ is a SPNE of \mathcal{K}^E .*

Proof. Consider a cartel Γ in which the largest firms participate. Suppose that whenever the smallest participant leaves Γ this would render the cartel unstable. This implies that (3.14) is satisfied for all $i \in \Gamma$. It remains to be shown that (3.13) holds for all $i \in N \setminus \Gamma$. If so, the proof is complete. If not, then at least one outsider has an incentive to join Γ . By Theorem 3.3, whenever (3.13) is violated for some $i \in N \setminus \Gamma$, it must be violated for the largest fringe member. Let $r \in N \setminus \Gamma$ denote the largest outsider. If r joins Γ , it becomes the smallest member of the new cartel Γ_{+r} . Clearly, (3.14) is satisfied for firm r , because otherwise firm r would have had no incentive to join Γ . Note that Theorem 3.3 implies that all firms larger than firm r prefer to stay in Γ . Hence, the participation constraint is satisfied for all members of Γ_{+r} . Repeating this exercise leaves two possible outcomes. Either, at some point (3.13) is satisfied for the largest outsider. Then, by Lemma 3.4, all smaller fringe members have no incentive to join either. Alternatively, no outsiders are left, i.e., the cartel is all-inclusive. ■

Notice, however, that there typically exists a whole range of coalitions for which (3.13), (3.14) and (3.10) hold simultaneously. Cartels in which the smallest members are smaller than the largest outsiders might well be a solution of the game. More generally, whether or not moderate-sized firms have an incentive to take part in a cartel often depends on what type of ‘formation game’ is played.

To illustrate, suppose that firms do not take their participation decision simultaneously, but instead play a sequential formation game as in Bloch (1996) and Prokop (1999). Recall that both studies conclude that when players are identical and farsighted, the last x firms in a sequence of proposers will join the cartel, leaving the first $n - x$ firms as outsiders.²⁰ This result is driven by the fact that in equilibrium outsiders earn more profits than insiders and, as a result, all firms have a strong incentive to free-ride on the cartel. The first firms to propose anticipate that announcing “no” yields a situation in which the remaining firms have an incentive to collude, because the alternative for the latter is competition.²¹

The predicted equilibrium cartel size is unique in these games, which is due to the assumption of symmetry. The symmetry assumption, however, also implies that the outcome is unique up to permutation. Applying such sequential formation rules to

²⁰Bloch (1996) considers a situation in which firms sequentially propose a coalition and applies this approach to a simultaneous symmetric Cournot game. Prokop (1999) allows firms only to say “yes” or “no” and applies this approach to the symmetric model of collusive price leadership as analyzed in d’Aspremont *et al.* (1983). See Chapter 2 of this thesis for a more detailed description.

²¹Note that this outcome requires all firms to commit to their choices.

the current heterogeneous setting typically yields a variety of equilibria. In particular, what coalition will form depends heavily on the position firms take in the sequence of proposers. For instance, a relatively large firm at the beginning of the sequence may well announce “no” realizing that a relatively small firm will have an incentive to say “yes” later on.

It has been shown in Lemma 3.4. and Theorem 3.3 that sufficiently small firms will always say “no” and that large firms are more inclined to say “yes” to a coalition. This result is *independent of the order of moves*. To what extent moderate-sized firms have an incentive to collude is sensitive to the order of moves and therefore strongly depends on the specific rules of the formation game that is played. The order in which firms take their decision is quite arbitrary and what cartel will actually form remains therefore difficult to predict.

3.5 The Most Profitable Cartel

In a wide array of settings, total cartel value is highest for a cartel that is all-inclusive. We have already established that when cartelizing is costly this might not be true. In this section, we explore if and under what conditions firms have an incentive to form the cartel agreement for which *total cartel profits* are highest. Note that the most profitable cartel is not necessarily all-inclusive as in the current setting colluding is costly. In the following, let Γ^* denote the cartel which generates the highest total cartel value. First, we examine what firms take part in Γ^* and then we explore whether or not Γ^* is a solution of \mathcal{K}^E .

The following result shows that, for a fixed number of cartel participants, the most profitable cartel consists of the largest firms in the industry.

Proposition 3.3 *For any given cartel size, the most profitable cartel comprises the largest firms in the industry.*

Proof. Consider a sustainable cartel Γ in which x members participate. Total cartel value amounts to,

$$V^c(p^c, \Gamma) = \frac{1}{1 - \delta} \left[(p^c - c) \left[D(p^c) - \sum_{i \in N} k_i + \sum_{j \in \Gamma} k_j \right] - T(x) \right].$$

Observe that, given x , $V^c(p^c, \Gamma)$ is increasing in total cartel capacity. With the proportional profit allocation rule as given by (3.15) all cartel members face the following critical discount factor,

$$\delta^* = \frac{(p^c - c) (\sum_{i \in N} k_i - D(p^c)) + T(x)}{(p^c - c) \sum_{j \in \Gamma} k_j},$$

which is decreasing in total cartel capacity. It can be concluded that for a given number of participants the most profitable cartel consists of the largest firms in the industry. ■

The above result implies that, given the number of cartel members, whenever some collusion is sustainable, than a cartel comprising the largest firms is sustainable too. Hence, we know that Γ^* is formed by the largest sellers in the market. In addition, we have already established that the most profitable cartel is less than all-inclusive when colluding is costly and the smallest firms are sufficiently small. As a result, the size of Γ^* is uniquely determined.²²

To see whether or not firms have an incentive to form the most profitable cartel Γ^* notice first that (3.10) trivially holds, because otherwise Γ^* would not be viable. Furthermore, it has been established that the smallest member of Γ^* is (weakly) larger than the largest outsider. By Proposition 3.2 we know that a cartel comprising the largest firms in the industry is indeed an equilibrium outcome of \mathcal{K}^E . However, this result is in itself not sufficient. For Γ^* to be a solution of \mathcal{K}^E it is also required that the right number of sellers has an incentive to participate.

The next result reveals that (3.13) is always satisfied for firms that are not included in Γ^* .

Lemma 3.5 $\pi_i^o(\Gamma^*) \geq \pi_i^c(\Gamma_{+i}^*)$ for all $i \in N \setminus \Gamma^*$.

Proof. Without loss of generality, assume Γ^* consists of x members and let firm $r \in N \setminus \Gamma^*$ denote some outsider. Firm r has no incentive to join Γ^* whenever the following condition holds,

$$\pi_r^o(\Gamma^*) \geq \pi_r^c(\Gamma_{+r}^*),$$

which is equivalent to,

$$(p^c(\Gamma^*) - c) \left(\sum_{j \in \Gamma^*} k_j + k_r \right) \geq \pi^c(\Gamma_{+r}^*). \quad (3.20)$$

Note that by definition $\pi^c(\Gamma_{+r}^*) \leq \pi^c(\Gamma^*)$, because otherwise Γ^* would not be the most profitable cartel arrangement. Hence, (3.20) holds whenever the following condition is satisfied,

$$(p^c(\Gamma^*) - c) \left(\sum_{j \in \Gamma^*} k_j + k_r \right) \geq \pi^c(\Gamma^*),$$

or,

$$(p^c(\Gamma^*) - c) \left(\sum_{j \in \Gamma^*} k_j + k_r \right) \geq (p^c(\Gamma^*) - c) \left[D(p^c(\Gamma^*)) - \sum_{j \in N} k_j + \sum_{j \in \Gamma^*} k_j \right] - T(x).$$

Rearranging yields,

$$(p^c(\Gamma^*) - c) k_r \geq (p^c(\Gamma^*) - c) \left[D(p^c(\Gamma^*)) - \sum_{j \in N} k_j \right] - T(x),$$

²²Note that when some firms have equally large capacity stocks the composition of the most profitable cartel arrangement might not be unique.

which always holds. It can be concluded that (3.13) is satisfied for all outsiders to Γ^* . ■

Hence, whenever Γ^* is less than all-inclusive, none of the fringe members has an incentive to take part in the most profitable cartel arrangement. For Γ^* to be a SPNE of \mathcal{K}^E it remains to be shown that none of the participants prefers to be a fringe member. We will show that this is not generally true. In fact, whether or not firms have an incentive to form Γ^* depends on the size of the smallest member.

Theorem 3.4 *If the smallest member of Γ^* is sufficiently large, then Γ^* is a SPNE of \mathcal{K}^E .*

Proof. Recall that for Γ^* (3.10) is naturally satisfied. Moreover, by Lemma 3.5 we know that (3.13) always holds. Hence, for Γ^* to be a SPNE of \mathcal{K}^E it remains to be shown that $\pi_i^c(\Gamma^*) \geq \pi_i^o(\Gamma_{-i}^*)$ for all $i \in \Gamma^*$. To that end, let r be the smallest member of Γ^* and suppose without loss of generality that the most profitable cartel consists of x firms. Firm r prefers to stay in Γ^* if,

$$\pi_r^c(\Gamma^*) \geq \pi_r^o(\Gamma_{-r}^*).$$

This is equivalent to,

$$\left((p^c(\Gamma^*) - c) \left[D(p^c(\Gamma^*)) - \sum_{i \in N} k_i + \sum_{j \in \Gamma^*} k_j \right] - T(x) \right) \frac{k_r}{\sum_{j \in \Gamma^*} k_j} \geq (p^c(\Gamma_{-r}^*) - c) k_r.$$

Rearranging yields,

$$k_r \geq \frac{(p^c(\Gamma^*) - c) (\sum_{i \in N} k_i - D(p^c(\Gamma^*))) + T(x)}{p^c(\Gamma^*) - p^c(\Gamma_{-r}^*)} - \sum_{\substack{j \in \Gamma^* \\ j \neq r}} k_j. \quad (3.21)$$

Observe that this inequality holds for k_r sufficiently large. Note that whenever (3.21) is satisfied for firm r it will be satisfied for all larger members too (by Theorem 3.3.), which completes the proof. ■

The above result shows that for the most profitable cartel to be a solution of \mathcal{K}^E it must be the case that the smallest member owns a sufficient part of total industry capacity. To illustrate, suppose that the cost of colluding is so low that Γ^* is all-inclusive. Whether or not firms have an incentive to form Γ^* then depends on the share of industry capacity that is under the control of the smallest market player. If this firm is very small it prefers to be a fringe member and the most profitable cartel arrangement will not emerge.

3.6 Incomplete Cartels and Mergers

The incentive to collude is positively related to the capacity stock available to sellers. If and what cartel is likely to emerge therefore partly depends on the size distribution of

firms in the industry. For instance, a cartel is unlikely to be all-inclusive in markets with one or more very small undertakings. A particular way in which the size distribution of firms can change is through a merger. In this section, mergers are discussed in relation to incomplete cartels. We address two issues: Merger incentives and the potential coordinated effects of a merger. In what follows, we assume that mergers do not create synergies, i.e., the level of production costs remains unchanged post-merger.²³ The effect of a merger is then a reduction in the number of firms and two or more firms become one larger firm with a capacity stock equal to the sum of capacities of the merging parties.

3.6.1 Merger Incentives

In this subsection, the incentives of firms to engage in a merger are analyzed. A distinction can be made between merger incentives in the presence and absence of collusion. Note that, under the assumptions made, the competitive equilibrium is unaffected by a merger. The reason is that the static Nash equilibrium is neutral to the number of firms as long as $n \geq 3$, which holds by assumption. As a result, none of the firms has an incentive to engage in a merger in absence of collusion. It must be noticed, however, that competing firms might find it in their interest to merge when collusion is anticipated at a later stage.

Alternatively, consider a situation in which a cartel is already present. To begin, notice that a cartel Γ which adopts a proportional profit allocation rule as given by (3.15) faces the following constrained optimization problem,

$$\max_{p^c} V^c(p^c, \Gamma) = \frac{1}{1-\delta} \left[(p^c - c) \left[D(p^c) - \sum_{i \in N} k_i + \sum_{j \in \Gamma} k_j \right] - T(\cdot) \right], \quad (3.22)$$

subject to,

$$\delta \geq \delta^* = \frac{\sum_{i \in N} k_i - D(p^c)}{\sum_{j \in \Gamma} k_j} + \frac{T(\cdot)}{(p^c - c) \sum_{j \in \Gamma} k_j}. \quad (3.23)$$

Observe that δ^* is the same for all members and decreasing in total cartel capacity, all else equal. Furthermore, (3.23) reveals that the sustainability of a cartel does not depend on the size distribution of participants, but on the total share of industry capacity that is under the control of the cartel. This does not mean that a merger between two or more cartel members has no effect on the incentive constraint. In fact, a merger between two or more cartel participants loosens the ICC. The reason is that there are fewer firms in the cartel post-merger, which leads to a reduction in $T(\cdot)$.

Hence, when the cartel is all-inclusive, all firms have an incentive to merge. If the cartel is not all-inclusive, there are two more cases to consider: (i) a merger between fringe members, and (iii) a merger between one or more cartel members with one or more fringe firms. In analyzing the latter, we assume that whenever the new larger firm does not take part in the cartel this will be perceived as cheating by cartel members

²³Clearly, it is beneficial for firms to merge whenever this leads to lower unit costs.

not involved in the merger. The following result shows that firms have an incentive to merge only when they are part of the cartel or become part of the cartel post-merger.

Proposition 3.4 *In the presence of a cartel,*

- (i) *Fringe firms have no incentive to merge,*
- (ii) *Cartel members do have an incentive to merge,*
- (iii) *Fringe firms and cartel members have an incentive to merge only if they are part of the cartel post-merger.*

Proof. Fix some collusive equilibrium in which there exists a cartel Γ consisting of x members. Three type of mergers may occur (i) a merger among two or more fringe members, (ii) a merger among two or more cartel members and (iii) a merger between one or more cartel members with one or more outsiders. We will discuss each case in turn.

(i) Note that a merger among two or more outsiders does not alter total cartel capacity. Following Lemma 3.3, we know that the optimal cartel price and therefore the mark-up of merging parties remains unaltered. The merger yields no additional profits, because total capacity of the merger equals the sum of production capacities of the merging firms.

(ii) The total cartel value of Γ equals,

$$\frac{1}{1-\delta} \left((p^c(\Gamma) - c) \left[D(p^c(\Gamma)) - \sum_{i \in N} k_i + \sum_{j \in \Gamma} k_j \right] - T(x) \right).$$

Observe that a merger between two or more cartel participants only leads to a decline in the number of cartel participants, i.e., the merger only negatively affects $T(x)$. Consequently, any merger between cartel members increases total cartel value. Notice further that the cartel is naturally sustainable post-merger, because $\frac{\partial \delta^*}{\partial T} > 0$. It can be concluded that cartel members always have an incentive to merge.

(iii) In the following, we show that a merger between cartel members and fringe members yields a larger firm that has no incentive to defect from the cartel agreement. To that end, let Ψ denote the coalition of merging parties. The total value that this coalition could earn by deviating optimally amounts to,

$$(p^c(\Gamma) - c) \sum_{j \in \Psi} k_j.$$

Such deviation is beneficial only if total deviating profits exceed the sum of profits that the merging parties earn separately, i.e.,

$$(p^c(\Gamma) - c) \sum_{j \in N \setminus \Gamma, \Psi} k_j + \sum_{j \in \Gamma, \Psi} \pi_j^d \geq \frac{1}{1-\delta} (p^c - c) \sum_{j \in N \setminus \Gamma, \Psi} k_j + \sum_{j \in \Gamma, \Psi} V_j^c.$$

This inequality is violated, because Γ is sustainable, i.e., $V_j^c \geq \pi_j^d$ for all $j \in \Gamma, \Psi$. Finally, note that (3.23) is decreasing in total cartel capacity and that the number of cartel firms will not increase. Hence, the cartel is always sustainable post-merger.

Fringe firms and cartel members therefore might have an incentive to merge and when they do they will be part of the cartel post-merger. ■

The above result is in contrast with Escrihuela-Villar (2008b) which analyzes incomplete cartels in a symmetric infinitely repeated Cournot model. He shows that, in the presence of a cartel, mergers among outsiders and mergers among insiders lead to a price increase. In the current setting, the price level depends on total cartel capacity and therefore remains unaffected by a merger.²⁴ He also finds that a cartel reduces merger incentives, while in our case merger incentives are higher under a collusive regime. Furthermore, he shows that a cartel provides additional incentives for fringe firms to merge. In our model, an incomplete cartel provides no incentives for outsiders to organize their production capacity in a single company.²⁵ It is important to emphasize that the above result is largely driven by the fact that competitive profits do not depend on the number of sellers in the market. This is typical for homogeneous goods markets in which firms compete in price. If, instead, firms compete in quantity or when goods are differentiated, firms might find it beneficial to merge in absence of collusion.

3.6.2 Coordinated Effects of Mergers

In this subsection, we study potential coordinated effects that result from a merger. To provide some preliminary intuition, consider a merger between two very small sellers so that, following Lemma 3.4, both sellers have no incentive to join a cartel. Clearly, if the new firm is sufficiently small, merging parties still have no incentive to join a cartel post-merger. Hence, a merger between sufficiently small firms has no impact. In a related vein, a merger between sufficiently large firms, in that both would have joined a cartel pre-merger, has no impact either. That is, a merger between sufficiently large firms leaves total cartel capacity unaltered and therefore the merger has no impact on the equilibrium cartel price. The analysis presented below suggest that the composition of a stable cartel is most likely affected by a merger among moderate-sized firms.

Compte *et al.* (2002) and Vasconcelos (2005) are two important contributions that explore how the distribution of capacity stocks affects collusion. Assuming an all-inclusive cartel, both studies examine an infinitely repeated game in which firms differ in terms of production capacity. Compte *et al.* (2002) assumes that firms compete in price and find that the critical discount factor is increasing in the size of the largest firm(s). Vasconcelos (2005) considers a Cournot setting and finds that not only the size of the largest firms matter, but also that of the smallest. The smallest firm has the highest incentive to deviate from the cartel agreement, while the largest firm has the strongest incentive not to punish in the event of cheating. Consequently, redistributing capacity from the largest to the smallest seller facilitates collusion. Both papers show that the minimum critical discount factor is obtained with a symmetric capacity

²⁴This is not true for the case in which cartel members merge with fringe firms. However, in Escrihuela-Villar (2008b) this possibility is not considered.

²⁵Likewise, fringe firms have no incentive to form a cartel. As a consequence, in the current setting it is never optimal to have multiple cartels.

distribution and it is in this sense that more symmetry facilitates collusion. Hence, both studies suggest that the strongest coordinated effects result from a redistribution of capacity involving either the largest firms or the smallest firms, or both.

By contrast, when cartel membership is endogenous and incomplete the strongest coordinated effects might well come from a redistribution of capacity among moderate-sized firms. For example, a merger between two firms of medium size results in a larger firm that might have an incentive to join a cartel post-merger. It is not immediately clear, however, if cartel capacity increases due to a merger, because a merger that joins the cartel may induce other members to leave the coalition. It is difficult to derive analytical results, because any outcome quite generally depends on the entire capacity distribution. In Bos and Harrington (2008), we perform simulations to explore this issue in more detail. These results are discussed below. First, however, we believe it is instructive to examine some numerical examples.

3.6.2.1 Numerical Examples

For the sake of simplicity, suppose that colluding is costless and that market demand $D(p)$ is linear,

$$D(p) = 1 - p.$$

Thus, monopoly demand is equal to $\frac{1}{2}$. By Assumption 2, we then must have that,

$$\frac{1}{2} > k_1 \geq k_2 \geq \dots \geq k_n > 0,$$

Furthermore, assume that $c = 0$ so that by Assumption 3,

$$\sum_{i=2}^n k_i \geq 1.$$

We further suppose that $\delta \simeq 1$, which implies that the ICC is not binding. Hence, in equilibrium the cartel price of a cartel Γ is either:

$$p^c(K_\Gamma) = \frac{1 - K + K_\Gamma}{2} \text{ if } K_\Gamma > K - 1,$$

or

$$p^c(K_\Gamma) = c = 0 \text{ when } K_\Gamma \leq K - 1.$$

For Γ to be stable, we must have that none of its members has an incentive to leave the coalition and all outsiders should prefer to stay part of the competitive fringe. A firm $i \in \Gamma$ has no incentive to leave a profitable cartel Γ if this renders the cartel unprofitable, i.e., when $p^c(K_\Gamma - k_i) = c = 0$. When $p^c(K_\Gamma - k_i) > c = 0$, it has no incentive to join the fringe if,

$$[p^c(K_\Gamma) - c] \left(\frac{D(p^c(K_\Gamma)) - K + K_\Gamma}{K_\Gamma} \right) k_i > [p^c(K_\Gamma - k_i) - c] k_i,$$

which is equivalent to,

$$k_i > \frac{K_\Gamma^2 - (K - 1)^2}{2K_\Gamma}, \forall i \in \Gamma \quad (3.24)$$

Likewise, a firm $i \notin \Gamma$ has no incentive to join Γ if,

$$[p^c(K_\Gamma) - c]k_i \geq [p^c(K_\Gamma + k_i) - c] \left(\frac{D(p^c(K_\Gamma + k_i)) - K + K_\Gamma + k_i}{K_\Gamma + k_i} \right) k_i,$$

which can be rearranged to,

$$k_i \leq \sqrt{K_\Gamma^2 - (K - 1)^2}, \forall i \notin \Gamma. \quad (3.25)$$

Hence, a cartel Γ is stable when both (3.24) and (3.25) hold.

We consider an industry of eight firms ($n = 8$) pre-merger. Pre-merger, firms can take three sizes:

- Large firms with capacity $\frac{4}{10}$
- Medium firms with capacity $\frac{2}{10}$
- Small firms with capacity $\frac{1}{10}$

In the following, we analyze three cases.

Case 3.5 *A merger among non-cartel members could result in the merged firm joining the cartel and expand cartel membership.*

Suppose that the capacities of the eight firms are distributed as follows,

$$\begin{aligned} k_1 &= k_2 = \frac{4}{10} \\ k_3 &= k_4 = k_5 = \frac{2}{10} \\ k_6 &= k_7 = k_8 = \frac{1}{10}. \end{aligned}$$

- Pre-merger: Suppose that the pre-merger cartel is formed by firm 1 and 2, i.e., $\Gamma = \{1, 2\}$. This cartel is profitable because $K_\Gamma = \frac{8}{10} > K - 1 = \frac{7}{10}$. Furthermore, if any of the two members leaves the cartel it renders unprofitable, which implies it is internally stable. Γ is also externally stable, because

$$k_i^2 \leq \frac{15}{100}, \forall i \notin \Gamma.$$

Hence, the pre-merger cartel $\Gamma = \{1, 2\}$ is stable.

- Post-merger: Now suppose firm 3 and 4 merge, so that $k_3 + k_4 = \frac{2}{10} + \frac{2}{10} = \frac{4}{10} = k_m$. Firm 3 and 4 have an incentive to join Γ post-merger if,

$$k_m > \sqrt{K_\Gamma^2 - (K - 1)^2} \Rightarrow \frac{16}{100} > \frac{15}{100}.$$

Clearly, the firms want to join post-merger. Also, the new cartel $\Gamma + \{3/4\}$ is stable. It is internally stable, because we have shown that firms of size $\frac{4}{10}$ have an incentive to join a cartel of size $\frac{8}{10}$. It is also externally stable because:

$$k_5^2 = \frac{4}{100} \leq \frac{95}{100},$$

Thus, a merger among non-cartel members could result in the merged firm joining the cartel and expand cartel membership.

Case 3.6 *A merger between a cartel and a non-cartel member may leave the composition of the cartel unaffected and result in more cartel capacity.*

Now suppose the capacity distribution is as follows,

$$\begin{aligned} k_1 &= \frac{4}{10}, \\ k_2 &= k_3 = k_4 = \frac{2}{10} \\ k_5 &= k_6 = k_7 = k_8 = \frac{1}{10}, \end{aligned}$$

so that $K - 1 = \frac{4}{10}$. Also, $\sum_{i=2}^n k_i = \frac{10}{10} \geq 1$.

- Pre-merger: again, assume that there is a cartel $\Gamma = \{1, 2\}$, which is profitable because $K_\Gamma = \frac{6}{10} > K - 1 = \frac{4}{10}$. This cartel is internally stable, because when any of the two members leaves we have $K_\Gamma \leq K - 1$, which implies zero profits for all firms. Moreover, Γ is externally stable because,

$$k_i^2 \leq K_\Gamma^2 - (K - 1)^2 \Rightarrow k_i^2 \leq \frac{20}{100}, \forall i \notin \Gamma.$$

Therefore, $\Gamma = \{1, 2\}$ is a stable cartel.

- Post-merger: Now suppose firm 2 and firm 3 merge, so that $k_2 + k_3 = \frac{2}{10} + \frac{2}{10} = \frac{4}{10} = k_m$, which will be part of the cartel. The new cartel $\Gamma^m = \Gamma - \{2\} + \{2/3\}$ is internally stable, because whenever one of the two firms leave we have that $K_\Gamma \leq K - 1$, which implies zero profits, while profits are positive pre-merger. To see that it is also externally stable note that,

$$k_i^2 \leq K_\Gamma^2 - (K - 1)^2 \Rightarrow k_i^2 \leq \frac{48}{100}, \forall i \notin \Gamma^m.$$

Hence, Γ^m is stable. We have shown that a merger between a cartel and a non-cartel member may leave the composition of the cartel unaffected and result in more cartel capacity.

Case 3.7 *Total cartel capacity can be lower post-merger.*

Consider the following capacity distribution:

$$\begin{aligned} k_1 &= \frac{4}{10} \\ k_2 &= k_3 = k_4 = k_5 = k_6 = \frac{2}{10} \\ k_7 &= k_8 = \frac{1}{10}, \end{aligned}$$

so that $K - 1 = \frac{6}{10}$ and $\sum_{i=2}^n k_i = \frac{12}{10} \geq 1$.

- Pre-merger: Suppose firm 1, 2 and 3 are in the cartel, i.e., $\Gamma = \{1, 2, 3\}$. Total cartel capacity is then given by $K_\Gamma = \frac{8}{10}$. Thus, Γ is profitable because $K_\Gamma = \frac{8}{10} > K - 1 = \frac{6}{10}$ and internally stable, because whenever one of the participants leaves we have that $K_\Gamma \leq K - 1$. It is also externally stable because,

$$k_i^2 \leq K_\Gamma^2 - (K - 1)^2 \Rightarrow k_i^2 \leq \frac{28}{100}, \forall i \notin \Gamma.$$

As a result, $\Gamma = \{1, 2, 3\}$ is stable.

- Post-merger: Now suppose firm 2 merges with firm 7 so that $k_2 + k_7 = \frac{2}{10} + \frac{1}{10} = \frac{3}{10} = k_m$. Denote Γ^m the cartel post-merger. Note that both firm 1 and firm m must be part of Γ^m because total cartel capacity must be at least $\frac{6}{10}$ in order to be profitable. For instance, firm 1 and 3 have total capacity of $\frac{6}{10}$, which is not enough to make the cartel profitable. Yet, firm 3 potentially can leave the cartel post-merger without making it unprofitable because combined capacity of the merger and firm 1 equals $\frac{7}{10} > \frac{6}{10}$. Firm 3 has an incentive to leave the cartel $\Gamma^m = \Gamma - \{2\} + \{2/7\}$ when,

$$k_3 \leq \frac{K_{\Gamma^m}^2 - (K - 1)^2}{2K_{\Gamma^m}} = \frac{\frac{81}{100} - \frac{36}{100}}{\frac{180}{100}} \Rightarrow k_3 \leq \frac{1}{4},$$

and this holds because $k_3 = \frac{1}{5}$. Hence, $\Gamma^m = \Gamma - \{2\} + \{2/7\}$ is not stable and firm 3 will leave the cartel. Firm 1 and 2/7 will not leave because that would imply zero profits. Consequently, $\Gamma - \{2\} + \{2/7\} - \{3\}$ is internally stable. It is also externally stable, otherwise firm 3 would have had no incentive to leave (all outsiders are weakly smaller than firm 3). Finally, note that total cartel capacity post-merger is equal to $\frac{4}{10} + \frac{3}{10} = \frac{7}{10}$, while pre-merger it was equal to $\frac{4}{10} + \frac{2}{10} + \frac{2}{10} = \frac{8}{10}$. As a result, total cartel capacity is lower post-merger.

The first two examples illustrate how a merger between moderate-sized firms can increase the impact of a cartel. The third example shows how a merger involving a medium firm can lower cartel capacity. These examples suggest that the strongest

coordinated effects might come from mergers involving moderate-sized firms. It is clear, however, that a more advanced analysis is required to say anything more substantial.

In Bos and Harrington (2008), we make an attempt to address this issue more rigorously by means of simulations. We assess the impact of a two-firm merger on a stable cartel comprising the largest firms. By Proposition 3.2 we know that a cartel agreement between the largest firms is an equilibrium outcome of the game. For simplicity, assume that colluding is costless. The next result shows that when demand is linear there is a unique stable cartel consisting of the largest undertakings.

Theorem 3.8 *If demand is linear, then there is a unique stable cartel comprising the largest firms.*

Proof. Denote the smallest stable cartel $\Gamma = \{1, \dots, m\}$ so that all cartels smaller than Γ are unstable. We will show that any larger cartel is unstable too. To that end, we need to consider three cases. First, suppose the ICC is not binding for Γ , then we know the ICC is not binding for any larger cartel, because

$$\delta^* = \frac{\sum_{i \in N} k_i - D(p^c)}{\sum_{j \in \Gamma} k_j},$$

is the same for all cartel members and decreasing in total cartel capacity. Second, for δ relatively low, the ICC may be binding for Γ and for all larger cartels. Thirdly, the ICC might be binding for Γ , but not for some larger cartel(s). We discuss the three cases in the following.

(a) If δ is sufficiently high, then the ICC is not binding for Γ and for any larger cartel. By assumption, Γ is stable, which implies firm $m + 1$ has no incentive to join, or,

$$k_{m+1} \leq \sqrt{(K_\Gamma)^2 - (K - D(c))^2}. \quad (*)$$

A similar condition holds for outsiders to larger cartels. Note that,

$$k_{h+1} \leq k_{m+1} \leq \sqrt{\left(\sum_{j=1}^m k_j\right)^2 - (K - D(c))^2} < \sqrt{\left(\sum_{j=1}^h k_j\right)^2 - (K - D(c))^2}, \text{ for all } h > m. \quad (**)$$

Hence, if $\{1, \dots, m^*\}$ is a stable cartel then (*) holds. By (**), $\{1, \dots, f\}$ is not a stable cartel, for $f > m^*$, since firm f does not want to join.

(b) Suppose the ICC is binding for Γ and for all larger cartels. Γ is stable, which means that firm $m + 1$ has no incentive to join, or:

$$k_{m+1} \leq \frac{D(c) - K + \delta \left(\sum_{j=1}^m k_j\right)}{1 - \delta}.$$

Next note that:

$$k_{h+1} \leq k_{m+1} \leq \frac{D(c) - K + \delta \left(\sum_{j=1}^m k_j\right)}{1 - \delta} < \frac{D(c) - K + \delta \left(\sum_{j=1}^h k_j\right)}{1 - \delta}, \text{ for all } h > m. \quad (**)$$

By the same argument as above, there is a unique value for m whereby $\{1, \dots, m\}$ is a stable cartel.

(c) Suppose the ICC is binding for Γ , but non-binding for some larger cartel(s). There are two cases to consider:

(i) The ICC is binding for Γ , but non-binding for any larger cartel. Thus, we have that,

$$p^*(K_\Gamma) = \frac{a - K + \delta K_\Gamma}{b}, p^*(K_\Gamma + k_{m+1}) = \frac{a + bc - K + K_\Gamma + k_{m+1}}{2b},$$

and

$$k_{m+1} \leq \frac{D(c) - K + \delta K_\Gamma}{1 - \delta}.$$

Note that because Γ is assumed stable, the cartel $\Gamma + \{m + 1\}$ is unstable, because firm $m + 1$ wants to be an outsider to Γ . Therefore, consider cartel $\Gamma + \{m + 1\} + \{m + 2\}$, which faces an ICC that is non-binding. Firm $m + 2$ wants to leave this cartel if,

$$k_{m+2} \leq \sqrt{(K_\Gamma + k_{m+1})^2 - (K - D(c))^2}.$$

Rearranging gives,

$$k_{m+2}^2 - k_{m+1}^2 \leq K_\Gamma^2 + 2K_\Gamma k_{m+1} - (K - D(c))^2.$$

Note that $k_{m+1} \geq k_{m+2}$ so that the LHS is weakly negative. It therefore suffices to show that the RHS is larger or equal to zero.

$$K_\Gamma^2 + 2K_\Gamma k_{m+1} - (K - D(c))^2 \geq 0,$$

which is equivalent to,

$$\sqrt{K_\Gamma^2 + 2K_\Gamma k_{m+1}} \geq K - D(c).$$

This condition holds because $K_\Gamma > K - D(c)$ otherwise Γ would not be stable. Hence, the cartel $\Gamma + \{m + 1\} + \{m + 2\}$ is unstable. By repeating the exercise it can be shown that all larger cartels are unstable too.

(ii) Suppose the ICC is binding for Γ and non-binding for one or more larger cartels, but not all. Note that this implies that the ICC is binding for $\Gamma + \{m + 1\}$. Let $\Gamma^* = \{1, \dots, t\}$ be the largest cartel for which the ICC is binding, which implies $t \geq m + 1$. Note that by a similar analysis as in (b) any larger cartel than Γ that faces a binding ICC is unstable. For example, for the smallest member of Γ^* it holds that,

$$k_t \leq k_{m+1} \leq \frac{D(c) - K + \delta \left(\sum_{j=1}^m k_j \right)}{1 - \delta} \leq \frac{D(c) - K + \delta \left(\sum_{j=1}^{t-1} k_j \right)}{1 - \delta}.$$

Hence, if there exists another stable cartel it will face a non-binding ICC. In the following, Let $\Gamma^{**} = \{1, \dots, t + 1\}$ be the smallest cartel for which the ICC is non-binding. We have to show that Γ^{**} is unstable. To that end, consider firm $t + 1$ which

compares being in Γ^{**} with being outside Γ^* . Γ^{**} is unstable if firm $t + 1$ wants to leave the cartel, which is the case if,

$$(p^*(K_{\Gamma^{**}}) - c) \left[\frac{a - bp^*(K_{\Gamma^{**}}) - K + K_{\Gamma^{**}}}{K_{\Gamma^{**}}} \right] - (p^*(K_{\Gamma^*}) - c) \leq 0,$$

which is equivalent to,

$$\delta \geq \frac{(D(c) - K + K_{\Gamma^{**}})^2 + 4K_{\Gamma^{**}}(K - D(c))}{4K_{\Gamma^{**}}K_{\Gamma^*}}. \quad (3.26)$$

In this particular situation it holds that,

$$\frac{K - D(c) + K_{\Gamma^*}}{2K_{\Gamma^*}} \geq \delta \geq \frac{K - D(c) + K_{\Gamma^*} + k_{t+1}}{2(K_{\Gamma^*} + k_{t+1})},$$

because the ICC is binding for Γ^* and non-binding for Γ^{**} . The condition (3.26) therefore holds if:

$$\frac{K - D(c) + K_{\Gamma^{**}}}{2K_{\Gamma^{**}}} \geq \frac{(D(c) - K + K_{\Gamma^{**}})^2 + 4K_{\Gamma^{**}}(K - D(c))}{4K_{\Gamma^{**}}K_{\Gamma^*}}.$$

Rearranging yields,

$$\begin{aligned} K_{\Gamma^*}^2 &\geq (K - D(c))^2 - 2D(c)k_{t+1} + 2Kk_{t+1} + k_{t+1}^2, \\ K_{\Gamma^*}^2 &\geq (K - D(c) + k_{t+1})^2, \\ k_{t+1} &\leq D(c) - K + K_{\Gamma^*}. \end{aligned}$$

This holds because $K_{\Gamma^*} - k_{t+1} \geq K_{\Gamma} > K - D(c)$ otherwise Γ would not have been stable. Hence, Γ^{**} is unstable. Following the analysis in (i) above it can be shown that firm $t + 2$ does not want to join Γ^{**} so that a cartel $\{1, \dots, t + 2\}$ is unstable. In a similar fashion, it can be shown that all larger cartels are unstable too.

Consequently, there is a unique stable cartel comprising the largest firms when demand is linear. ■

We perform simulations for $D(p) = 1 - p$ and $c = 0$. A single simulation involves the following five steps.

1. Fix the number of firms, n .
2. Randomly select a vector of capacities (k_1, \dots, k_n) according to a uniform distribution over $(0, \frac{1}{2})^n$. Requiring that each firm's capacity is less than $\frac{1}{2}$ ensures that Assumption 2 is satisfied. Next check that $\sum_{h \neq i, j} k_h \geq 1, \forall i, j$. If that condition holds then Assumption 3 is satisfied both in the pre-merger and any post-merger situation, in which case go to step 3. If instead $\sum_{h \neq i, j} k_h < 1$ for some i, j then redraw the vector of capacities until the condition is satisfied.
3. Given (k_1, \dots, k_n) , randomly select the discount factor δ according to a uniform distribution over $(\frac{K - (a - bc)}{K}, 1)$ where $K \equiv \sum_{i=1}^n k_i$. By drawing δ from this interval, some collusion is sustainable. (Otherwise, a merger has no effect.)

4. Given (k_1, \dots, k_n) and δ , derive the unique stable cartel involving the largest firms. Record the pre-merger price.
5. Consider every possible two-firm merger and, for each of them, derive the new stable cartel and post-merger price. Record the change in price due to the merger as well the rank of the firms (in terms of capacity) involved in the merger.

This procedure is repeated 100,000 times and we compute the average price change that results from a merger for all two-firm mergers. With n firms, the number of possible mergers is equal to $1 + 2 + \dots + (n - 1)$. The simulations are performed for $n \in \{5, 6, 7, 8, 9, 10\}$. The results for $n = 5$ are reported in Table 3.1.

Table 3.1: Average price change due to a merger, $n = 5$
Capacity Rank of
Merger Partners

Firm A	Firm B	Average Price Change
4	5	0,0297
3	5	0,0185
1	5	0,0114
2	5	0,0114
3	4	0,0099
1	4	0,0047
2	4	0,0047
1	2	0
1	3	0
2	3	0

With $n = 5$, there are ten possible mergers. In Table 3.1, the mergers are ordered in terms of the size of their average price effects. The highest average price change is equal to 0,0297 and obtained with a merger between the two smallest firms. A merger between the median firm (firm 3) and the smallest firm has the next biggest price effect. A merger between any of the three largest firms has no impact. Observe that the price effects tend to be larger the smaller are the firms involved in the merger. Yet, with only five firms, the number of moderate-sized firms is limited. We therefore performed the same exercise for $n = \{6, 7, 8, 9, 10\}$.

When $n > 5$, there are many types of merger and, as a result we obtain quite a few average price changes. To organize the data, we partition firms in three categories: large, medium, and small. When there are six or nine firms in the industry, there is a natural categorization; when $n = 6$ (9), the largest two (three) firms are labelled *large*, the firms with the two (three) smallest capacities are labelled *small*, and the remaining firms are labelled *medium*. When the number of firms is not divisible by three, we consider the three partitions that are closest to having $n/3$ in each category. For example, when $n = 7$, one partition has firms ranked 1st, 2nd, and 3rd (in terms of capacity) being large, those ranked 4th and 5th being medium, and those ranked 6th and 7th being small; a second partition has firms ranked 1st and 2nd being large, those ranked 3rd, 4th, and 5th being medium, and those ranked 6th and 7th being small; and a third partition has firms ranked 1st and 2nd being large, those ranked 3rd

and 4th being medium, and those ranked 5th, 6th, and 7th being small. The results are presented in Table 3.2 below.

Table 3.2: Average price change due to a merger, $n \in \{6, 7, 8, 9, 10\}$

Categorization of Firms*				Merger Type**					
n	Large (L)	Medium (M)	Small (S)	L/L	L/M	L/S	M/M	M/S	S/S
6	1,2	3,4	5,6	0	0,0036	0,0161	0,0153	0,0255	0,0240
7	1,2,3	4,5	6,7	0	0,0125	0,0139	0,0258	0,0202	0,0138
	1,2	3,4,5	6,7	0	0,0067	0,0131	0,0201	0,0186	0,0138
	1,2	3,4	5,6,7	0	0,0034	0,0132	0,0144	0,0193	0,0184
8	1,2,3	4,5,6	7,8	0	0,0110	0,0099	0,0212	0,0149	0,0087
	1,2,3	4,5	6,7,8	0	0,0096	0,0112	0,0200	0,0173	0,0125
	1,2	3,4,5	6,7,8	0	0,0054	0,0109	0,0151	0,0155	0,0125
9	1,2,3	4,5,6	7,8,9	0	0,0081	0,0090	0,0169	0,0136	0,0094
10	1,2,3,4	5,6,7	8,9,10	0,0011	0,0085	0,0081	0,0163	0,0118	0,0074
	1,2,3	4,5,6,7	8,9,10	0	0,0063	0,0077	0,0136	0,0111	0,0074
	1,2,3	4,5,6	7,8,9,10	0	0,0063	0,0083	0,0117	0,0120	0,0097

* “Categorization of Firms” allocates a firm to being large, medium or small based on its rank in terms of capacity.

** “Merger Type” refers to the size - large, medium or small - of the firms participating in the merger.

Table 3.2 reports the average price changes from various mergers. Observe that in all the cases considered the biggest impact comes from mergers involving either two medium sized firms or one medium and one small firm. Furthermore, a merger involving two large firms always has the lowest impact. Also, a merger between two small firms never yields the highest average price change. Contrary to previous research that assumed an all-inclusive cartel, we do not find that it is mergers between large firms and/or small firms that have the biggest impact. Intuitively, a merger involving a medium sized firm may lead to a significant increase of cartel output post-merger.

3.7 Concluding Remarks

This chapter provides a rationale for the existence of incomplete cartels. We show that the value of an incomplete cartel might exceed that of a full cartel if colluding is costly and the smallest firm is sufficiently small. Also, there is a positive correlation between firm size and the incentive to join a cartel. Moreover, we have established that the most profitable cartel comprises the largest firms in the industry and that firms have an incentive to form this cartel when its smallest member is sufficiently large. Finally, we have analyzed merger incentives and the potential coordinated effects that may result from a merger. We have shown that firms have an incentive to merge only when they are part of a cartel or become part of the cartel post-merger. Our analysis further suggests that the most severe coordinated effects might come from mergers involving firms of medium size.

Part of the theoretical predictions find considerable support in ‘real-world’ examples of incomplete cartels. In Chapter 2, some stylized facts were distilled from descriptive

cartel studies and antitrust cases. Many cartels were incomplete, but took a dominant position in the market. In the model developed in this chapter, it is shown that an incomplete cartel is viable only when it controls a sufficiently large part of industry capacity. Furthermore, incomplete cartels often consisted of the larger firms in the market. We have established theoretically that larger firms are indeed more inclined to join a cartel. Finally, incomplete cartels tend to lose market share over time. In the current model, fringe firms do have an incentive to increase production in response to an output reduction by cartel members. However, outsiders are bound by the available production capacity and as a result output expansion by fringe members is limited. A potentially interesting extension of the model would be to endogenize total industry capacity.

There exists various ways in which total industry productivity might change. Sellers investing in additional production capacity is all but one example.²⁶ Investments not only lead industry capacity to expand, but might also cause a change in the size distribution of firms. Note that, akin to the merger incentives discussed above, firms have no incentive to invest in absence of collusion. The reason is that, under the assumptions made, the competitive equilibrium is unaffected by a change in firm size. However, given that total cartel profits are allocated proportionally, cartel members might have an incentive to increase their production capacity. The reason is that it allows a participant to claim a bigger share of total cartel value, all else unchanged. Whether or not participants decide to invest in part depends on how this will affect the sustainability of the cartel. All members face an identical incentive compatibility constraint, which is strictly increasing in cartel investments. One can imagine that participants do not want to undermine the stability of the cartel. To that end, a cartel might attempt to control the investment plans of its members.²⁷ As to fringe members, investments lead to an increase of fringe profits as long as the optimal cartel price remains unchanged. However, additional fringe capacity directly reduces cartel demand and creates a downward pressure on the cartel price. In addition, an expansion of fringe capacity potentially undermines the stability of the cartel, because it leads to an increase of the critical discount factor. In sum, it is to be expected that both the cartel participants as well as the competitive fringe have an incentive to invest in additional production capacity, which potentially threatens the existence of the cartel arrangement. To make this claim, however, a formal analysis is warranted, which is left for future research.

Another potentially interesting extension is to model the actual cartel formation process. At this point, it is important to re-emphasize that the above analysis yields no (unique) prediction about what cartel will actually form. That is, we have focused exclusively on the incentive of firms to collude and did not model the cartel formation process. In the last decades, some progress has been made in the analysis of group formation with externalities.²⁸ Yet, the study of endogenous coalition

²⁶Market entry/exit is another prominent example.

²⁷Notice that with a different profit allocation rule investments in additional capacity do not threaten cartel stability as long as the members facing the highest critical discount factor abstain from investing.

²⁸For an overview of this strand of literature see Yi (1997).

formation with externalities is complex. In particular, the problem easily becomes analytically intractable if one attempts to introduce firm heterogeneity in coalition formation games.²⁹ Also, as discussed in Section 3.4, results tend to be highly sensitive to the precise structure of the game. Part of this problem can be overcome by assuming identical firms, but as we have seen, this assumption makes it difficult to give results a practical interpretation.

Firm heterogeneity is an important issue also because it is commonly believed to hamper collusion. Indeed, diversity among sellers complicates both the formation and management of a cartel. However, this should not lead one to conclude that no collusion will occur in industries in which firms differ substantially. Although, a cartel that controls the entire industry is unlikely to emerge, a subgroup of sufficiently identical sellers can potentially form a cartel that is effective. One of the peculiar observations in our model is that symmetry does not foster collusion. In fact, the asymmetry of firms has no effect at all on whether or not the cartel is sustainable. Clearly, this is due to the restrictions that we imposed on the absolute size of the largest firms. As the study of Compte *et al.* (2002) reveals, relaxing these assumptions yields a setting in which the size distribution of cartel participants affects the sustainability of a cartel.

Assuming an upper bound on firm size is convenient in that it greatly simplifies the analysis. It might be that allowing for firms that are very large, in a sense that their capacity stock is sufficient to serve the entire market, does not affect the main results of this study. When colluding is costly and the smallest firms are sufficiently small, then the optimal cartel size will still be less than all-inclusive. Clearly, any cartel would be viable only when very large firms participate. This is in line with the prediction that sufficiently large firms have an incentive to take part in any cartel. However, whether or not the analysis yields similar results without Assumption 2 and Assumption 3 cannot be determined without further inquiry. Furthermore, it may be interesting to conduct a similar analysis in a Cournot setting. It must be noted, however, that one is likely to encounter several problems. For example, optimal punishment schemes are typically more involved compared to the current model. A particular problem occurs if one wants to assess the coordinated effects of merger in a standard Cournot framework. The reason is that a two-firm merger is unprofitable, because their combined market share is typically less than 80%.

Finally, this chapter is a modest attempt to delineate tacit collusion and overt collusion. Traditionally, theoretical work on cartels focuses on conditions under which collusion can be sustained without communication between undertakings. Our analysis suggests that communication can be a powerful ‘tool’ in dealing with the incentive of firms to cheat on a cartel, either *ex ante* or *ex post*. A better understanding of what distinguishes tacit from explicit collusion is particularly important in view of antitrust enforcement, because only the latter is considered illegal. One can imagine that certain market structures are more conducive to overt collusion as opposed to tacit collusion. As such, a more narrow focus on explicit cartel arrangements in future theoretical research may prove to be particularly valuable in the detection of cartels.

²⁹Belleflamme (2000) is one of the few attempts to study group formation with externalities and firm heterogeneity.

4

Cartel Detection and Antitrust Law Enforcement

“The chase is better than the catch” – Scooter¹

4.1 Introduction

In many fields of law enforcement, pro-active measures are taken to ensure compliance. For example, tax authorities frequently carry out ‘fishing expeditions’ in order to verify if certain groups of people truthfully reported their earnings, financial markets are screened for security fraud and the police regularly monitors highways to check if speed limits are respected and whether or not drivers are under the influence of alcohol or other drugs. In contrast, antitrust agencies around the world tend to rely on more passive instruments of enforcement. In particular, hard-core cartel activity is discovered almost exclusively through in-depth investigations in response to other clear signals of collusion. In other words, antitrust authorities are mainly *reacting*, not *acting*.²

An antitrust agency does, however, have the right to search for collusive behavior upon its own initiative, which we denote ‘cartel detection’. More precisely, we define it as follows,

Definition 4.1 *Cartel detection is a **pro-active, conscious search** for cartels by an antitrust agency.*

¹From the song “How much is the Fish?” by Scooter (1998).

²See, for example, Friederiszick and Maier-Rigaud (2007, 2008).

In principle, cartels can be brought to light by anybody. Yet, discoveries by, for instance, employees, direct and indirect purchasers, the board of directors, competitors of the cartel, are rather accidental and are usually communicated to the antitrust authority via a complaint procedure and/or a leniency program.³ The, admittedly somewhat narrow, definition of cartel detection therefore primarily serves the purpose of isolating those cases in which the antitrust agency is the first non-participant to find out about potential anticompetitive conduct. Additionally, the possibility that the antitrust authority discovers a cartel accidentally, for example, as a spin-off of some other investigation, is excluded from the definition.

This chapter discusses the potential role of cartel detection in antitrust enforcement and explains how economic theory can be used in the search for cartels. In antitrust practice, examples of cartel detection methods include tracking of individuals and media, infiltration and economic methods of detection.⁴ One can imagine that it might be useful to closely follow the careers of managers that have been involved in cartels before. Also, one potentially can find hints of anticompetitive behavior in relevant trade press and business related websites. In addition, the authority may attend business events with the aim to become familiar with existing social relations between firms and managers within a particular industry. Finally, use can be made of economic theory and econometric techniques to search for cartel activity.⁵ The latter is doubtless the least popular and many antitrust authorities, including the Antitrust Division of the U.S. Department of Justice, do not use economics to detect cartels.⁶

Currently, authorities that do make use of economics often use it to identify industries that are prone to collusion. This procedure typically works as follows. As a first step, investigators loosely define a particular market. For instance, the market for bicycles consists of about x players and the products involved are only slightly differentiated. A second step is then to collect more detailed information, which is partly used to determine the validity of the initial assumptions, e.g., there may be more sellers operating on the market for bicycles than was initially believed. Hence, the procedure essentially works top-down in an iterative manner. Once investigators are confident to have a sufficiently precise picture of the markets under consideration, these markets are categorized according to “likelihood of collusion”. Markets that are believed to be very conducive to anticompetitive conduct are monitored more closely and are possibly subject to further in-depth investigations. This categorization of markets is in addition based on historical findings (e.g., was there any cartel activity in this market in the past?) and cartel cases in other countries (e.g., are there any cartels discovered in similar markets in other countries?).

³See, for example, McAnney (1991).

⁴For an overview of cartel detection methods in antitrust practice the reader is referred to the Anti-Cartel Enforcement Manual, Chapter 4, 2007. This manual has been released recently by the International Competition Network and can be found at www.internationalcompetitionnetwork.org.

⁵The Anti-Cartel Enforcement Manual further mentions settlement strategies, education and cooperation between authorities as “pro-active methods of detecting cartels”.

⁶See the Anti-Cartel Enforcement Manual, Chapter 4, 2007. However, the U.S. Department of Justice does attempt to raise awareness of cartel activity as witnessed by the flyer ‘Price Fixing, Bid Rigging, and Market Allocation Schemes: What They Are and What to Look For: An Antitrust Primer’. This flyer is available at: <http://www.usdoj.gov/atr/public/guidelines/211578.htm>.

The role of economics in cartel detection is therefore profoundly limited to the identification of potential “crime scenes”, i.e., industries that are prone to collusion. It is hardly used in “crime scene investigations”, i.e., in the evaluation of certain business practices.⁷ There are several reasons for this. First, using (economic) techniques to analyze and judge firm behavior in a given industry is typically involved and because antitrust agencies face both time and budget constraints, it is often not perceived as an attractive option. Second, suitable economic methods may not always be available. Third, a more detailed economic analysis usually requires good quality micro-data, which either may not exist or which cannot be used for detection purposes due to legal constraints (e.g., because of privacy protection). Finally, both the complaint procedures as well as the leniency program are thought to be sufficiently effective and there exists no convincing evidence that available economic detection methods are indeed successful.

In this chapter, we argue that economics is likely to play an increasingly important role in cartel detection and, in turn, that cartel detection itself is likely to play a more prominent role in antitrust enforcement. To that end, this chapter provides an overview of economic methods of cartel detection and discusses the merits and demerits of these techniques. In particular, we make the case that significant improvements are within reach when detection methods are designed exclusively for a certain (type of) industry.

This chapter proceeds as follows. The goal and scope of cartel detection are discussed in Section 2. Section 3 provides an overview of economic methods of detection that are currently available. Section 4 discusses potential pitfalls of economic detection tests. In Section 5, special attention is given to detection techniques that can be used to search for incomplete cartels. In Section 6, we make the case that part of the pitfalls can be dealt with through the design of detection methods that are tailored to a certain (type of) industry. Section 7 concludes.

4.2 Goal and Scope of Cartel Detection

In order to protect competition as an institution, free market societies have adopted sets of competition rules. To enforce these rules of competition, most countries have allocated corresponding powers to an antitrust authority to pursue two main goals, deterrence and desistance. The prime goal of antitrust enforcement is to prevent firms from restricting competition in an unlawful manner, i.e., antitrust policy should be a deterrent. If nevertheless firms infringe rules of competition, the task is to stop them from doing so. However, to safeguard competition in the market place it must be known if and where rules of competition are being violated. Unlike with murder, robbery, rape, arson, *etcetera*, it is *a priori* unclear whether companies have indeed acted in breach of the law, because cartels hardly leave obvious traces that indicate

⁷Some antitrust agencies do analyze bid-patterns in auctions to search for symptoms of bid-rigging. Occasionally use is made of econometric techniques to analyze bidding behavior. See the Anti-Cartel Enforcement Manual, Chapter 4.

anticompetitive conduct.⁸ To make sure antitrust enforcement is effective it is essential to make potential anticompetitive behavior visible.

In this section, we argue that cartel detection, in all likelihood, will play an increasingly important role in antitrust policy. In addition, we make the case that economics will become a key determinant of cartel detection.

4.2.1 *Why do we need Cartel Detection?*

Some scholars and antitrust practitioners are skeptical about the potential of cartel detection. The most frequently raised objections against cartel detection are that methods currently available are complex and difficult to implement and that results, if any, are often ambiguous.⁹ Moreover, the large majority of cartels known today were disclosed by means of a complaint procedure or a leniency program.¹⁰ Be that as it may, cartel detection will arguably become increasingly important in antitrust enforcement. Perhaps the best way to explain why is by asking a classic economic question, “what is the alternative?”. In other words, what does antitrust enforcement look like without cartel detection?

Not denying the quantitative success of both the complaint procedure and the leniency program, it is questionable if these are equally successful in qualitative terms, i.e., in terms of social welfare. One may, for instance, wonder why competitors would submit a complaint given the well-established result in the economic literature that non-participants often benefit from the presence of a cartel.¹¹ Indeed, as noted by Porter (2005, p. 149),

“...one should be suspicious of complaints by a rival firm not party to the conspiracy. Rivals typically gain from higher prices and they suffer from more intense competition.”

But even if the cartel is harmful for customers and outsiders, they will only ever consider filing a complaint if they know the cartel exists. There are, however, good reasons to believe that the more sophisticated cartels, which are arguably among the most harmful ones, have effective strategies at their disposal to hide their illegal practices.¹²

Without an active search for cartels, a similar argument holds for the leniency program.¹³ The main question would be why any of the participants would ever start

⁸See also Schinkel (2008).

⁹See, for example, Rey (2007).

¹⁰Harrington (2007) reports that cartel detection was to some extent successful in finding a cartel in the Dutch shrimp industry.

¹¹See Chapter 2 and 3 of this dissertation.

¹²For example, the cartel may gradually increase its price so as not to attract too much attention by its customers. See Harrington and Chen (2006).

¹³The effectiveness of leniency programs is theoretically discussed in, for example, Aubert *et al.* (2006) and Motta and Polo (2003). There further exists experimental evidence that a leniency policy may enhance the deterrent effect of antitrust policy and may lead to lower cartel prices. See Hinlopen and Soetevent (2008). However, both theoretical and experimental contributions on leniency programs typically assume positive detection probabilities.

such a “race to the courthouse”. That is, why would a criminal turn himself in and confess his guilt if the police is not even looking for him? Either, the authority is about to discover the cartel for other unclear reasons, or cartel members are trying to use the program to their own advantage, i.e., once the cartel breaks down the first firm to self-report avoids a fine, but at the same time allows the authority to impose a, possibly substantial, fine on fellow members.¹⁴ Cartel detection therefore has an important indirect effect as well. When effective, cartel detection increases the probability of discovery, which in turn encourages members to apply for leniency.¹⁵

Arguably, well-organized sophisticated cartels will take measures to avoid discovery via complaint procedures or leniency programs.¹⁶ In particular, one is likely to find out mostly those cartels that are already falling apart or the least professional ones.¹⁷ Hence, a more active search and destroy policy is warranted in order to guarantee effective antitrust enforcement.

4.2.2 Why do we need Economics to Detect Cartels?

It is often not immediately clear whether or not markets are “infected” by collusive behavior. This implies that before hunting the bad guys, we need to know if and where they are likely to operate. Economists are traditionally well-trained in identifying such potential “crime scenes”. For example, it has long been recognized that cartels are, *ceteris paribus*, more likely to be present in industries with a limited number of firms and with significant barriers to entry.¹⁸ However, once suspect industries have been identified, it is not yet clear if any cartel activity indeed takes place. The main, but at the same time most challenging, exercise is to determine to what extent observed behavior is led by ‘visible hands’.

A straightforward way to discover cartel activity is through interception of relevant communication between cartel members, e.g., finding minutes of a cartel meeting. Frankly however, to attempt to detect cartels by sifting through E-mail, phone calls, *etcetera*, between firms is not only legally objectionable, but is also likely to be a dead end road. The reason is that with modern technology it is becoming easier to encode pieces of information.¹⁹ Also, there is no need anymore for managers to be physically present in the same room or to use the phone. An increasing number of criminal organizations make use of environments on the internet that cannot be located by, for example, search engines. This so-called “freenet” (or related technologies) allows participants to exchange information almost anonymously.²⁰ Use is made of programs similar to MSN messenger, while it remains very difficult and sometimes even impossible to trace the source of the information exchanged.

¹⁴See Goppelsroeder *et al.* (2008) which argues that leniency programs potentially can be abused by sophisticated cartel organizations.

¹⁵See Friederiszick and Maier-Rigaud (2008) for an argument along similar lines.

¹⁶See Schinkel (2008).

¹⁷See Posner (1970) and Goppelsroeder *et al.* (2008)..

¹⁸See Section 3 of this chapter.

¹⁹See Halliday and Seabright (2001).

²⁰See Clarke *et al.* (2001).

Yet, in contrast to the internal organization of cartels, the *result* of the agreement is far less easy to cover.²¹ Implementing the cartel contract implies that members have to restrict production levels and increase their selling prices, which is often observable. The key question is therefore if we recognize cartel behavior on the basis of what we actually observe and what we typically observe are *economic* data. Consequently, the challenge is to use available economic information to determine if overt collusion provides the best explanation for observed firm conduct. In other words, the objective of economic detection methods is to identify market conduct that is the result of an explicit cartel agreement between firms.

A first major challenge is then to delineate tacit collusion and explicit collusion, because only the latter is considered illegal.²² Firms that collude tacitly coordinate actions without direct communication between them. Collusion is explicit when firms directly communicate to establish a cartel agreement. Given some collusive market result, tacit collusion is arguably the preferred strategy. On the one hand, such tacit coordination of strategies is, *ceteris paribus*, more efficient (e.g., you save cost associated with the formation and management of the cartel). On the other hand, firms substantially reduce the risk of being caught for acting in breach of antitrust laws, because they leave no traces of physical evidence. However, colluding tacitly can be quite complex, in particular when there exists no natural focal price. A firm that feels the current price is too low may aim to hike the price, but this may be costly when the price increase is not followed by other firms, e.g., it may lose market share. Similarly, reducing the price to signal that collusive prices should be lower is potentially costly, because it may trigger a price war if misunderstood by competitors. When received as the start of a price war it typically cannot be stopped by individual firms without communication.²³

Instead, communicating directly on the preferred price-quantity combinations avoids a lot of these risks. Also, shocks that change market conditions are less likely to undermine the stability of the cartel. Direct communication between firms allows them to coordinate on a new equilibrium. Such explicit collusion, however, contains drawbacks. It not only implies costly negotiations, but it is also more risky in the sense that a firm probably has to communicate more private information than in situations of conscious parallelism. In addition, gathering together in “smoke-filled rooms” potentially leaves extra evidence that can be used in trial. For example, there may exist minutes of the meeting or evidence that the suspect managers were together in the same hotel on the same day.

Despite its importance, economists do not have a clear understanding yet of how to delineate tacit from overt collusion. In particular, it is often not immediately clear if observed market conduct is the result of tacit or overt collusion. For example, in some markets firms may find out about prices of rivals faster than consumers. High priced firms may then decide to buy out the low priced firm(s), i.e., buy all the products of

²¹See also Bos (2007).

²²Tacit collusion is not prohibited by law. In Europe, for example, tacit collusion without further evidence is not sufficient to establish an infringement of article 81. See, for instance, *Ahlström v. Commission*, 1988, ECR 5193 (*Wood Pulp*).

²³See Motta (2004).

the low priced firms. The obtained products may then be sold to customers at higher prices. Van Cayseele and Furth (1996b) show that in a Stackelberg setting firms often can obtain the monopoly outcome by buying up rival output. Note that, in principle, firms can express buy-out behavior non-cooperatively.²⁴ That is, if one observes buy-out behavior, this may simply be the result of a tacit coordination of actions. First refusal contracts, i.e., agreements between firms to offer the output to each other first, may have a similar effect.²⁵ However, with first refusal contracts firms typically have to communicate “more directly”, which is why this practice should be considered more suspect.

To further illustrate the problem, consider the study by Christie and Schultz (1994) who aim to find an answer to the question why the vast majority of NASDAQ market makers seem to avoid so-called odd-eight quotes in 1991. They consider various potential explanations for this surprising phenomenon and provide indirect evidence indicating that this market outcome is the result of tacit collusion among market players. This conclusion finds additional support in a subsequent study by Christie *et al.* (1994), which shows a rapid and significant increase in the use of odd-eight quotes after the results of the previous study were published in leading newspapers. Yet, even though both studies provide convincing evidence of the existence of a collusive agreement the possibility that this agreement may have been explicit is not considered. In particular, the very fact that the NASDAQ has many market makers that know each other reasonably well in combination with the sudden breakdown after May 26, 1994 (i.e., the day that the results of Christie and Schultz (1994) became publicly known), suggests that overt collusion may well form a better explanation for observed practices.²⁶

The second major challenge is to delineate (imperfect) competition and collusion. Currently, economic methods of cartel detection almost exclusively focus on this distinction. These methods are discussed in the next section.

4.3 Economic Methods of Cartel Detection

The economic literature on cartel detection distinguishes between so-called *structural* and *behavioral* methods of detection.²⁷ The structural approach aims to identify what type of industries are conducive to anticompetitive conduct. The behavioral approach focuses instead on suspect firm behavior. The two approaches are not mutually exclusive and, in order to enhance the effectiveness of cartel detection, may very well be used in combination. The structural approach can be viewed a first phase in which a (type of) market with a high probability of being a “potential crime scene” is deter-

²⁴See, for instance, Van Cayseele and Furth (2001) which shows that with a buy-out option firms can reach the monopoly outcome without an explicit cartel agreement

²⁵See Van Cayseele and Furth (1996a).

²⁶Christie and Schultz (1994) states that “market makers interact frequently and over longer periods of time with the same population of other market makers.”

²⁷See Grout and Sonderegger (2005) for a discussion of structural factors that make markets prone to collusion. The behavioral approach is discussed in detail in Harrington (2007).

mined. A second step is then to analyze firm behavior in this particular industry, i.e., the behavioral approach can be used for “potential crime scene” investigations.

This section provides an overview of economic methods of cartel detection. It is convenient to organize the detection literature in such a way that it closely follows the traditional structure-conduct-performance paradigm.²⁸ In its simplest form, this paradigm expresses the view that market structure affects firm conduct, which in turn determines market performance. Arguably, some industries are more prone to conspiracy than others and the structural approach helps to identify these industries. Detection methods that focus on firm behavior might then be used to delineate competition and collusion. Finally, economic performance can be determined by testing if observed behavior is more consistent with collusion or instead with (imperfect) competition. In addition, one could test for robustness of results by analyzing alternative specifications.

4.3.1 Market Structure

In the absence of warnings that an industry may be “infected” by cartel activity, it is *a priori* unclear *where* the antitrust agency should start its search for cartels. Screening all sectors in the economy is, given budget as well as time constraints, simply impossible.²⁹ Also, based on the premise that the majority of markets is reasonably competitive, selecting industries randomly as subjects of closer scrutiny is unlikely to be fruitful. Consequently, some preliminary information is required.

Since the start of Industrial Organization as a distinct field of economic research, many contributions have enhanced our understanding of what type of markets are conducive to collusive conduct.³⁰ Market characteristics, especially those that are relatively easy to observe, can be used as a first screen to determine the likelihood of collusion in a particular industry. One must keep in mind, however, that such a market structure analysis can never be a conclusive exercise. That is to say, some markets that seem highly vulnerable to anticompetitive behavior may very well not be cartelized. Likewise, cartels can be present in markets that, at first sight, appear quite competitive. Yet, identifying market structures that are sensitive to anticompetitive conduct, *ceteris paribus*, increases the probability that behavioral analyses, which are typically more involved, will be successful.

One particular advantage of detection methods that focus on the structure of the market is that market characteristics are to a large extent beyond the control of firms. Arguably, sophisticated cartels will consciously try to ‘fly under the radar’, but manipulating the market environment in which the cartel operates is often difficult or even impossible. As we discuss below in more detail, cartels are, *ceteris paribus*,

²⁸This paradigm dominated Industrial Organization until the 1970s. Joe Bain is by many viewed to be the founding father. He took this approach in several, mainly descriptive, papers, around 1950. Most of the results have been summarized in his famous work: *Barriers to New Competition* (1956).

²⁹Notice that, from a social welfare perspective, screening all markets is, even without budget and time constraints, unlikely to be an optimal policy given that competition is sufficiently strong in most industries.

³⁰For an overview see, for example, Motta (2004), Chapter 4.

more likely to emerge in industries characterized by a limited number of firms, stable demand patterns and many small buyers. It is difficult to see how a cartel could significantly manipulate these factors so as to give a more competitive appearance.

One important limitation of this type of analysis, however, is that it typically does not allow the investigator to discriminate between tacit and overt collusive arrangements. This is partly due to the fact that the structural approach relies quite heavily on theoretical analyses carried out in an infinitely repeated game setting. The infinitely repeated game framework stems from the *noncooperative* game theory literature in which explicit communication between players is commonly assumed to be absent. In other words, the stylized market characteristics listed below might equally well be used to select industries in which ‘conscious parallelism’ is a possibility.

Infinitely repeated games have already been discussed in the previous two chapters. Recall that in this setting collusive agreements are typically sustained by means of threats. Assuming grim-trigger strategies, a cartel member will find it profitable to abide by the cartel contract when the following incentive compatibility constraint is satisfied,

$$\sum_{t=t_0}^{\infty} \delta^{t-1} \pi^c \geq \delta^{t_0-1} \pi^d + \sum_{t=t_0+1}^{\infty} \delta^{t-1} \pi^N. \quad (4.1)$$

Like before, individual cartel profits, profits of defection and competitive profits are respectively denoted, π^c , π^d and π^N , with $\pi^d > \pi^c > \pi^N$. The common discount factor is given by δ and t_0 indicates the date of defection. The factors listed below affect (4.1) in a variety of ways.

Number of Firms

Ever since Chamberlin (1933), economists are well aware of the negative correlation between the number of firms in a market and the likelihood of collusion in that market. There are at least two reasons why the number of firms matters. On the one hand, it is easier to reach an agreement when there are only a few parties involved, all else equal. For example, negotiations are typically more difficult and therefore increasingly costly with a larger number of parties gathered around the table. On the other hand, a cartel agreement between only a few firms might be easier to sustain. To illustrate, consider the situation of an all-inclusive cartel agreement between firms that compete in price and produce a homogeneous product. The number of participants affects (4.1) in the following sense. First, π^c will be higher the smaller is the number of participants. Second, the short-run gain from defection increases with the number of participants. Both effects make that (4.1) is more likely to hold when the market consists of only a few undertakings, all else equal.³¹ It must be noted, however, that in other settings a larger number of cartel participants may result in a more severe punishment, which helps to stabilize the cartel agreement.

³¹A notable exception to this theory are markets in which there exist strong ties between firms, for example, through membership of a trade association. Here the trade association may function as a coordinating device. Historically, a significant number of cartels were organized by means of such a ‘head-quarter’ that sometimes even determined the prices and quantities to be set by all members.

Contestability of the Market

Entry barriers are probably among the most widely accepted structural factors to make markets conducive to anticompetitive behavior. Supra-normal profits generated by a cartel potentially attracts new producers and these newcomers may well undermine the stability of the cartel. The cartel is protected against new competition when entry barriers are significant. To put it differently, potential newcomers might find it too costly to start a new profitable business when the “entry-fee” is sufficiently high.

To see how entry barriers affect the likelihood of collusion, we may distinguish between two scenarios. Either the cartel anticipates entry or it does not. If it does not expect new producers to enter the market, but entry nevertheless occurs, the result will be a loss in market share, a price war or the newcomer (tacitly) joins the cartel agreement. Clearly, entry is never beneficial for cartel members, because in all three cases the per-firm cartel profit decreases. This, in turn, makes (4.1) less likely to hold. Hence, compared to a situation in which entry is not profitable, a cartel agreement is, *ceteris paribus*, less likely to be sustainable.

Alternatively, the cartel rightly anticipates the possibility of entry and *ex ante* decides either to accommodate or to deter entry. If the optimal cartel strategy is to accommodate entry, entry will occur and per-firm cartel profits will decrease. If the cartel, in contrast, finds it optimal to deter entry it has to put a limit on the per period profit levels earned, i.e., it has to implement some form of limit pricing. Either way, it will be, *ceteris paribus*, less attractive to take part in a cartel agreement the easier it is for new firms to enter the market.³²

Nature of Demand

The possibility for firms to collude depends on the stability, the frequency and the elasticity of demand. If demand is unstable and firms imperfectly monitor each others actions, it is difficult to maintain the cartel contract. The reason is that a sudden drop in sales caused by a demand shock may be interpreted by a firm as a signal that another cartel member deviated, which might cause a price war. That cartels tend to break down when slumps are sufficiently large has been argued by Green and Porter (1984). Rotemberg and Saloner (1986), in contrast, argues that cheating occurs more likely during booms even when firms have complete information. Either way, both contributions support the view that cartels are, *ceteris paribus*, more stable the lower the impact of both slumps and booms.

Likewise, infrequent demand makes it more difficult for firms to form a cartel that is sustainable. The less frequent is demand, the more attractive it becomes to defect, all else equal. The reason is that the punishment phase following the period of defection is not materialized until the next period in which demand is positive and, given that the discount factor is strictly lower than one, is therefore less of a threat. In a similar fashion, markets in which firms frequently interact are more prone to collusion. A low frequency of interaction implies that it takes longer before the punishment phase can become effective, which makes deviating more attractive under the assumption that cartel members are less than infinitely patient.

³²See Baumol *et al.* (1982), for an extensive discussion of theories of contestable markets.

Finally, collusion is only an attractive strategy if demand is sufficiently inelastic. Clearly, if demand is perfectly elastic no price increase is profitable and consequently there are no incentives to collude. Hence, there exists a negative correlation between the incentive to collude and the elasticity of demand.³³

Market Transparency

In many cases, cartel members will have to monitor each other in order to ensure compliance with the cartel contract.³⁴ To see that market transparency is an important factor, note that in the context of repeated games the incentive to deviate is negatively correlated with the level of punishment. In the extreme, if fellow members will not observe any change in the event of cheating, deviating will go unpunished, which implies collusion cannot be sustained. More generally, information lags make the future “more distant” and this implies that the punishment is lower, which enhances the incentive to deviate from the agreement. Collusion is therefore, *ceteris paribus*, more likely to occur in markets in which information on, for instance, prices and sales is easily accessible.

Symmetry

Intuitively, more or less identical firms will find it easier to reach an agreement than firms which differ substantially. To illustrate the idea, consider a situation in which potential cartel members have different cost levels. This has immediate implications for what cartel price these firms would find optimal or, to put it less strict, acceptable. Tacit collusion, for example, would be quite a challenge in absence of a natural focal price. However, also in the case of overt collusion asymmetry among parties is unlikely to help cartel formation. Negotiations on what price to set, for instance, will be more difficult and therefore costly, which, in turn, makes joining the cartel less attractive. There exists some theoretical support for the intuition that symmetry among firms indeed facilitates the sustainability of a cartel.³⁵

Number of Buyers

Many dispersed clients that individually do not affect total demand is ideal for cartels. Perhaps the best way to explain why is by taking the other extreme in which all firms are competing for one and the same customer. In this case, the trust level must be extremely high. Per period, only one firm has positive sales. This may imply the cartel will have to use a scheme of side-payments, which makes the cartel organization complex. Alternatively, the cartel can designate every member a certain period in which it “wins the customer”. Yet, this is a risky strategy for members that have to wait long, because those who already made their sale have only a weak incentive to adhere to the cartel agreement.

In addition, a few large buyers may have a stronger incentive to claim antitrust damages compared to many small dispersed clients. Large customers may also be

³³See, for instance, Pindyck (1979).

³⁴Recent evidence from European hard-core cartel cases suggests that monitoring is indeed an important issue for many cartels. See Harrington (2006).

³⁵See Chapter 2 of this thesis.

better informed about the illegality of cartel arrangements. As a result, the chance that an antitrust agency receives signals of antitrust violations via the complaint procedure may be higher in markets in which there is significant buyer power.³⁶ In short, with only a few large buyers the incentive to defect are strong and consequently forming a cartel agreement may be too high of a goal.

Also, markets in which trade is organized through auctions are more vulnerable to collusive conduct the larger the number of contracts or objects offered in a given period. Pesendorfer (2000), for example, studies the bidding for school milk contracts in Texas and Florida during the 1980s. In this period, incomplete bidding rings were active in both states. One ring used a market-sharing scheme, while the other used a scheme of side-payments. Pesendorfer establishes theoretically that both forms of cartel organization almost maximize expected cartel profits provided that the number of contracts is sufficiently large. Consequently, he concludes that a buyer should choose to offer a single contract instead of many small contracts when there exist a substantial risk of collusion. On the one hand, a single or small number of contracts increases the gains from defection, while, on the other hand, it lowers the expected cartel profits for cartels that do not use side-payments, i.e., so-called weak cartels.

4.3.1.1 Ambiguous Structural Factors

The list presented above is not exhaustive and in most studies that take the structural approach more factors are listed. However, these factors typically have an ambiguous effect and form well-known puzzles in the industrial economics literature. They have in common that the effect on the incentive to deviate and the level of punishment is the same in the sense that both are either positively or negatively affected. Consequently, the net effect is unclear and very sensitive to parameter specifications. Three of the most prominent ambiguous structural factors are briefly discussed below.

Type of Product

The type of product is traditionally viewed to be an important determinant of the likelihood of collusion. In general, the argument goes that competitive pressure is positively correlated with the incentive to collude. This competitive pressure is maximal when customers base their purchasing decisions solely on price, i.e., when products offered by firms are from a consumer perspective perfect substitutes. In the other extreme case, products are, from a customer perspective, completely different, which means that sellers do not actually compete. Hence, when products are independent there is no need for firms to reduce competitive pressure. Traditionally, therefore, it has been argued that one is, *ceteris paribus*, more likely to find cartels in homogeneous industries.³⁷

Yet, the degree of differentiation also affects the stability of the cartel. On the one hand, the higher the substitutability of products and services, the more severe is the

³⁶I thank Jan Tuinstra for pointing this out.

³⁷From a more practical perspective, we may argue that differentiation makes it more difficult to define strategies that are acceptable to all parties.

punishment, because competitive profits are lower.³⁸ This is in line with the traditional argument that homogeneous industries are more conducive to cartel activity. On the other hand, however, the more differentiated are products the lower the gain from deviating, because lowering the price slightly only attracts some extra customers and not the entire market. It is therefore *a priori* unclear which of these antagonistic forces dominates.³⁹

Excess Capacity and Inventories

Firms can commit themselves to the cartel agreement by limiting production capacity in such a way that the agreed upon output level is the maximum that can be produced in a given period. In such a setting, no incentives to deviate exist, because even though a price-cut will attract additional demand, it cannot be satisfied by the chiseling firm. From this perspective, limiting capacity makes cartels more stable, because it eliminates the possibility of profitable deviations. By contrast, however, the net effect on (4.1) is unclear in light of excess capacity and/or inventories. On the one hand, excess capacity and inventories create an incentive to cut the price and increase production, but, on the other hand, it allows fellow members to credibly punish a defecting member.⁴⁰

Multi-market Contact

Multi-market contact means that firms meet each other in more than one market. Traditionally, it has been argued that multi-market contact facilitates collusion, because punishment is potentially more severe. That is, price wars that are triggered by defection might occur in all markets at the same time, which makes deviating less attractive. However, gains from defection also go up if a firm deviates in all markets simultaneously. Again, therefore it is unclear, which of these effects dominates.⁴¹

4.3.2 Cartel Conduct

Once a “potential crime scene” has been identified, the next step to take is to screen this particular industry in search for signals of cartel activity, i.e., to do some “crime scene investigations”. At this stage, the focus is primarily on firm conduct. At the heart of the behavioral approach in cartel detection lies the idea that firms under a cartel regime behave differently than they do in competition. The key challenge is then to identify these differences. Cartel members often leave, though possibly subtle, traces that can be understood with economics. For instance, Porter and Zona (1999) analyzes procurement auctions for school milk contracts in Ohio. One of their findings is that a

³⁸In the extreme situation in which competition is ‘perfect’ the difference between cartel profits π^c and competitive profits is maximal, because $\pi^N = 0$.

³⁹See Chang (1991), Rothschild (1992) and Ross (1992) for formal analyses of this issue.

⁴⁰For theoretical analyses of the effect of excess capacity on the likelihood of collusion see Benoit and Krishna (1987) and Davidson and Deneckere (1990).

⁴¹For a theoretical analysis of the effect of multi-market contact on collusion see Bernheim and Whinston (1990). In particular, they show that multi-market contact might only have an effect on the stability of collusion in light of asymmetries among firms.

group of firms is suspect, because they submit higher bids on contracts located close by than on contracts located farther away. The competitive model logically predicts a positive correlation between the distance to the school and the level of the bid, which was confirmed by the bid-patterns of non-conspirators.

In order to determine if observed firm behavior is suspect one necessarily needs a point of reference. A first step is therefore to define a sound competitive benchmark.

4.3.2.1 What does Competition look like?

There basically exist three approaches to get an intuition of what competition in a particular industry looks like. If one has information about when the cartel started to operate or when it (temporarily) collapsed, one in principle could take the firm behavior before and after the collusive period as a proxy for competitive conduct. Notice, however, that this approach is only useful if one indeed has information about the date of birth and longevity of the cartel organization under consideration. Therefore, this approach at best is helpful for *ex post* analyses and will only be of limited use in *ex ante* search for cartel activities.⁴² Alternatively, one might analyze similar type of markets to formulate a reasonable competitive benchmark.⁴³

A third method to define a competitive benchmark is by using industry specific information. For example, descriptive studies and industry surveys often contain, albeit sometimes hidden, signals about how firms in that particular market compete. In addition, one could try to collect information by interviewing people that work in the field. Getting in close contact with a particular industry can potentially yield quite some valuable information. For example, one might discover what firms are considered major competitors (e.g., whether or not there is a price-leader, total number of competitors). Also, one may receive a better idea of the main inputs of a product, which might be helpful in assessing unit production costs. One further could obtain information about the importance of transportation and service costs and whether or not these form a significant part of total cost. Together, this (type of) information might be quite helpful in determining what price levels can be expected in competition.

Using questionnaires and similar techniques, however, may be risky, because it potentially could be perceived as a signal that the industry is under close scrutiny by the authorities. Also, the information may not always be reliable. Arguably, however, it is the most promising way to define a reasonable competitive benchmark.

4.3.2.2 What does Collusion look like?

One in principle could argue that in order to spot cartel activity it is sufficient to have a thorough understanding of what competition looks like in a particular industry. All observed conduct that for longer periods substantially differs from the competitive benchmark should then be considered suspect. Such a detection strategy, however, is risky and easily may lead to wrong conclusions, because we often lack confidence about

⁴²It may, however, be very useful to monitor industries in which cartels have been active in the past. Recidivism is a prevalent phenomenon in antitrust practice.

⁴³See, for instance, Posner (1970) who remarks that comparing similar regions may be a practical method to extend the scope of economic evidence in antitrust procedures.

the validity of the competitive model. That is, it is difficult to develop a competitive model that perfectly predicts competitive behavior in a particular market. We can significantly enhance the strength of cartel detection by forming hypotheses about what collusive behavior would look like. A cartel detection test then reduces to determining if either the competitive model or the collusive model better explains observed firm behavior.⁴⁴ This clearly does not tackle the problem of potential misspecification of the models used, but it allows the investigators to draw conclusions with far greater confidence.

In order to understand what collusion looks like, we have to be clear about what exactly we are looking for. Given the secret nature of cartel organizations one of the few things we have at our disposal is *observable* market conduct. The choices made by the cartel are at least partly reflected in market data. Investigators in principle have access to price and output patterns. Also, it may not be too difficult to observe investment decisions and entry and exit in the market, *etcetera*. This sort of information may contain hints that point to potential anticompetitive behavior.

Prices are a prime candidate to be used in the search for cartels, because these are affected by a hardcore-cartel directly and, in addition, are relatively easily accessible. Cartels are, *ceteris paribus*, believed to have higher average prices and a lower price variance. Bolotova *et al.* (2008) have tested this prediction by analyzing price data of the *citric acid* and the *lysine* cartel. The hypotheses on the first two moments of the price distribution is confirmed in the *lysine* case. However, the *citric acid* cartel appears to have had a higher variance in price during the collusive phase. The authors conjecture that the latter, counterintuitive, result may be caused by the relatively long duration of the cartel. If the cartel operates for a longer period it may have to deal more often with chiseling firms and if cheating is not sufficient to make the cartel collapse entirely the variance in price will increase. Also, part of the result may be due to the fact that the data set containing prices charged during the pre-cartel period and post-cartel period was limited, which may have caused a somewhat biased competitive benchmark.

Simultaneously, one may take a dynamic approach and use available price data to identify potential structural breaks in pricing behavior over time. Consider again the *citric acid* and *lysine* cartel. The following two graphs plot the prices charged in both cartelized markets for the relevant period.

Although both cartels differ in many respects, the price patterns in both markets exhibit remarkable similarities. The similarities are summarized in Figure 4.3, which is a stylized figure of a cartel price path.⁴⁵

In both cases, the pre-cartel period is characterized by a significant price volatility.⁴⁶ In particular, average industry prices drop substantially during the pre-cartel period.

⁴⁴See, for example, Bajari (2001) and Bajari and Summers (2002).

⁴⁵This figure has been presented by Joe Harrington in his keynote lecture at the EARIE IO conference, Amsterdam, 2006. For further detailed analyses on cartel pricing dynamics the reader is referred to Harrington (2004, 2005).

⁴⁶The *citric acid* and *lysine* cartel are used as illustrative examples. However, a similar price pattern could be observed for other cartels as well. The *graphite electrodes* cartel forms another example, see Levenstein and Suslow (2001).

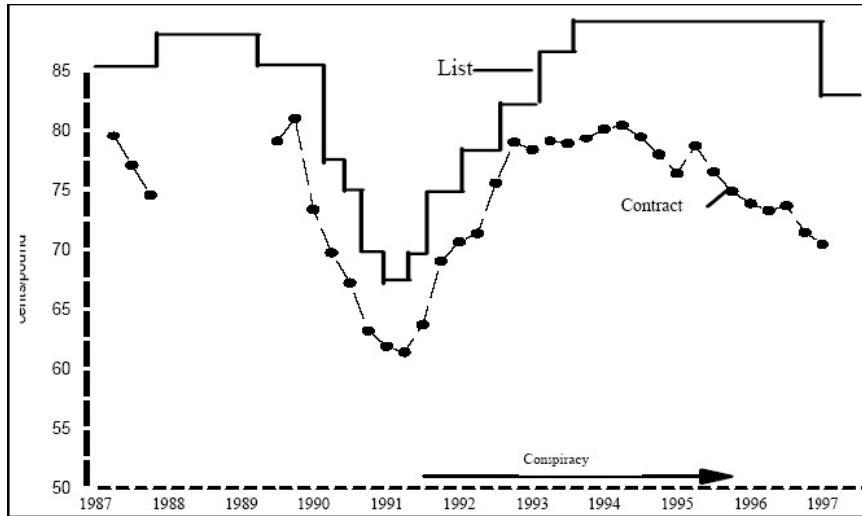


Figure 4.1 Price pattern of the citric acid market 1987-1997. Source: Connor (2001).

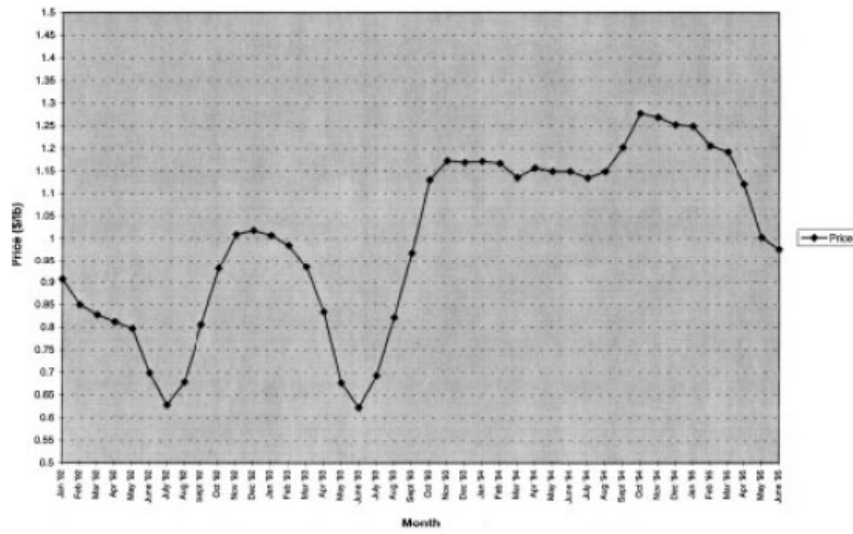


Figure 4.2 Price pattern of the lysine market 1992-1995. Source: Connor (2001).

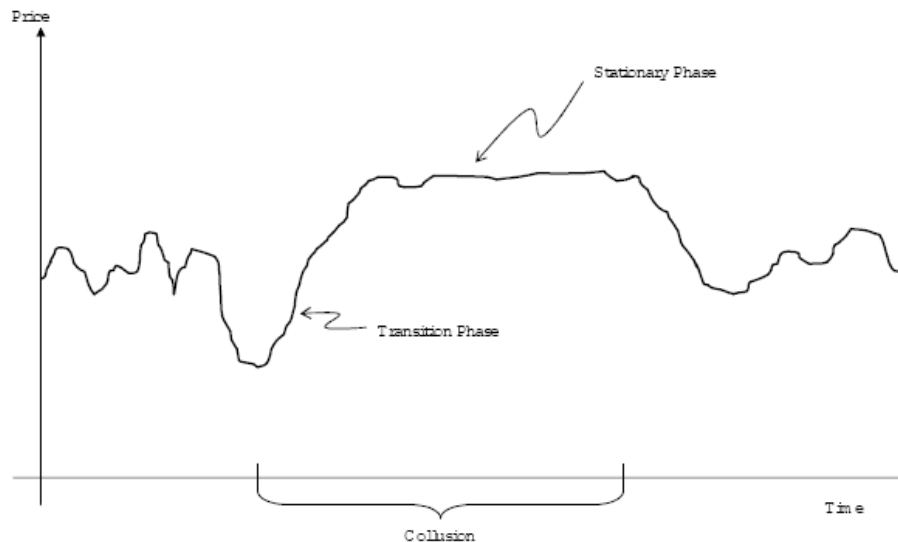


Figure 4.3 Canonical cartel price path. Source: Harrington (2006).

The cartels started to operate when industry prices were at a very low level. After the cartel was formed, however, industry prices increased only gradually. In other words, the cartels took their time to reach the intended cartel price. In Figure 4.3, this is referred to as the transition phase. The transition phase ends when the cartel has reached the preferred price level. Prices remain then remarkably stable for a substantial amount of time. This period is denoted ‘the stationary phase’ in Figure 4.3. The stationary phase ends when the cartel breaks down, which results in a decrease of average prices and prices become more volatile again. This completes the cartel price path. Note that the *lysine* cartel had a temporary break-down and that the cartel price path can be observed twice in Figure 4.2.

The intuition underlying this canonical cartel price path is as follows. The incentives to form a cartel are highest when competition is fierce. Figure 4.1 and 4.2 clearly indicate that the average price level decreased substantially before the cartel was formed. Once the cartel was formed industry prices were rising gradually, perhaps because the cartel did not want to attract too much attention from antitrust agencies and buyers.⁴⁷ We can imagine that a sudden significant price hike would indeed be considered strange in industries for which edgeworth price cycles are uncommon. The stationary stage reflects the conventional wisdom that cartels dislike change. Substantial price fluctuations undermine the stability of the cartel, for example, because an unantici-

⁴⁷See Harrington and Chen (2006).

pated price drop by one of the members may be interpreted as an attempt to defect from the agreement.

It must be noted, however, that merely studying visualized price patterns may not be sufficient and could even be misleading. For example, the shape of the price path is partly influenced by scaling, i.e., increments used on the axis. Ideally, therefore, one would like to analyze price data more formally and identify the various cartel phases by determining structural breaks in price patterns. To test for structural breaks in price series, one could apply a Chow test. Yet, a Chow test will only be useful as a screening device if one has an idea about candidate break points. Hence, we need some understanding of when the cartel started to operate or the date of its demise. An example of information that may be helpful in determining candidate break points are the formation of trade associations, which in the *lysine* cartel was formed to cover cartel meetings.⁴⁸

Descriptive cartel studies occasionally reveal behavioral aspects common to cartels. Harrington (2006) studies about twenty recent European hard-core cartel cases and derives some behavioral ‘stylized facts’ that may be indicative of collusion. One aspect is that cartels typically create more ‘uniformity’. For example, as a result of collusion, products tend to become increasingly standardized and there is an increased uniformity of prices, quality of products and services provided. Also, there is a reduced variation in prices offered to customers and market shares tend to be very stable over time. These factors all support the intuition that cartels prefer an easy life and that too much variety and changes in the market make collusion more difficult or even impossible.

4.3.3 Market Performance

The final stage in the application of economic methods of cartel detection is testing hypotheses. Cartel detection yields evidence of collusion when either the competitive model does not very well explain observed behavior or when firm conduct is better explained with some collusive model. It is important to note, however, that the economic approach at best yields circumstantial evidence of collusion. That is to say, economic methods of detection are not designed to find physical evidence of cartel activity and therefore at best will yield indirect evidence of anticompetitive conduct. Yet, in light of the goal of antitrust enforcement, collecting indirect evidence is arguably important. Circumstantial evidence may guide antitrust agencies in their decisions on what (type of) industries need closer scrutiny. When taking into account ‘tell-tale signs’ of collusion, antitrust authorities may increase the probability that in-depth investigations, e.g., search for more direct evidence via ‘dawn raids’, will be successful. As such, cartel detection contributes to the goal of desistance. Cartel detection is also likely to enhance the deterrent effect of antitrust enforcement. The reason being that it increases the probability of being caught, which makes collusion more costly in expected terms. Although difficult to verify, we may conjecture that cartel detection prevents some

⁴⁸See Harrington (2008). Connor (2001) reports that the Amino Acid Manufacturers International Association was formed by members of the *lysine* cartel.

firms from engaging in anticompetitive conspiracies. In addition, a more active detection policy may also render collusion less attractive by reducing (expected) mark-ups of potential cartel members.⁴⁹

The success of a certain detection method depends on a lot of factors and a significant number of potential problems have to be dealt with before one can successfully apply the method and be sufficiently confident about the results. A discussion of these potential pitfalls will be postponed to the next section. For now we may ask what method is the right one, if any? There exists no conclusive answer to this question. That is, there exists currently no manual that tells us when to apply what method. It will ultimately depend on the (type of) industry under consideration, on the available data, on (economic) intuition and even on pure luck. Very often, however, one may increase the chance of rightly determine market performance by using a combination of the various indicators and methods currently available. The following two studies form an illustrative example of how a combination of these factors may be used to discriminate between collusion and competition.

An interesting method of detection has been developed by Bajari and Ye (2003). They propose a model of competitive bidding for procurement contracts with bidders that may have different cost levels, e.g., some firms may be located closer to the project locations than others. Intuitively, bids are ‘conditionally independent’ and ‘exchangeable’ in competition. The first condition means that, after controlling for those costs that are publicly available (e.g., cost of traveling the distance between the factory and the project site), the bids must be independent. The second condition means that costs alone, independent of the identity of competitors, will determine the bid. By contrast, in situations of bid-rigging we may expect correlated bids when members consciously submit phony bids to give competitive appearance. Also, we might find that cartel members do not bid against each other as aggressively as a control group of non-participating firms. Bajari and Ye apply their test to procurement auctions for seal coating. Their test not only suggests when there is collusive behavior, i.e., some firms behaved contrary to the competitive model, but also identifies which firms are likely to be in the cartel.

Baldwin *et al.* (1997) uses a structural model to compare collusive and competitive bidding for Forest Service timber contracts. One of their main goals is to find out if bidding rings were active at the Forest Service auctions in the period 1975-1981. In particular, they also consider the possibility that relatively low winning bids could be caused by the positive supply shocks that occurred at the time. The authors construct five empirical models and find that the winning bids could best be explained by the presence of bidding rings.

4.4 Potential Pitfalls in Cartel Detection

Developing an economic detection method is a challenging exercise. A great many obstacles have to be dealt with before one can be reasonably confident about the

⁴⁹See, for instance, Block *et al.* (1981).

effectiveness of a detection tool. This section discusses some of the major problems that an investigator might encounter in the design and implementation of an economic detection technique.

4.4.1 *No Result, is a Result*

A detection method that rightly identifies a cartel, first and foremost, proves that antitrust enforcement is not fully deterrent. One therefore should be reluctant to judge the quality of a detection technique solely on the basis of the number of cartels that it was able to unveil. Indeed, abstracting from the level of antitrust penalties as well as the time and budget constraints of an antitrust authority, the perfect detection method is one that does not detect any cartel. The main aim of cartel detection is to make collusion more costly in expected terms, thereby contributing to the prime goal of antitrust enforcement; deterrence. This does not imply that a detection method is ineffective if it does not cause full deterrence. In principle, it is effective if it narrows the set of preferred collusive equilibria. Broadly speaking, this means that cartels must put more effort in avoiding discovery. In turn, this makes it, *ceteris paribus*, less attractive for firms to take part in a conspiracy.

In a related vein, a detection method that can be beaten is not ineffective *per se*. In principle, all detection techniques can be beaten. To illustrate how a cartel may avoid detection, consider the method proposed by Bajari and Ye (2003) that was discussed in Section 4.3.3. Suppose bidders had agreed on adding a fixed amount to their competitive bids. This would not affect the ‘conditional independence’ criterion, because adding a fixed amount does not affect the correlation of bids. The criterion of ‘exchangeability’ is also not violated, because the higher bids are related to the true cost levels. A bidding ring that implements this strategy could therefore beat this test. However, adding a fixed amount to competitive bids may not be the optimal collusive scheme in absence of cartel detection.

4.4.2 *Descriptive Flaws, Empirical Limitations and Theoretical Complications*

There exist quite a few descriptive cartel studies. Many studies use a data set of antitrust cases that is analyzed with the aim to test theoretical predictions.⁵⁰ Most of the structural factors described above find support in these descriptive studies. For example, concentration has been shown to be consistently and positively related to collusive success.⁵¹ Yet, from a methodological point of view, all these studies suffer from one major, fundamental problem.⁵² The data sets used are likely to describe only a subset of all cartels that are operational in a given period. The conclusions drawn

⁵⁰Examples include Hay and Kelly (1974), Arsch and Seneca (1975), Fraas and Greer (1977), Griffin (1989), Levenstein and Suslow (2006).

⁵¹Levenstein and Suslow (2004). Typically, a key role is played by industry associations and national governments for those cartels that comprise say 15 or more participants.

⁵²Of course, descriptive studies have more problems common to all inductive studies, like, for example, the reliability of the data that are used.

on the basis of these cases can therefore only be generalized safely when there exist no significant sample bias. However, because the total pool of cartels is fundamentally unknown we can neither be sure if such a sample bias is present, nor about the magnitude of this bias.⁵³ Descriptive studies of cartels form an important contribution and, at a minimum, provide an intuition about how cartels may operate, the problems they encounter and how they dealt with them. However, not denying the value of descriptive studies, there logically seems to be no way around the sample bias problem. This means we must be careful in designing cartel detection methods that rely heavily on individual cartel cases.

In order to screen markets or verify hypotheses about the likelihood of collusion in a particular market, one typically needs to make use of (econometric) techniques that require data. There are at least two potential limitations in this type of research. First, data may simply not be available or be very costly to collect. Also, even if data are available, they possibly cannot be used for detection purposes due to legal reasons (e.g., because of privacy protection). The latter can be particularly problematic when the strength of a detection technique depends heavily on micro data of individual firms. Alternatively, sometimes use can be made of more aggregate data. Second, it is typically hard to judge the quality of the data set. Clearly, if the data are of poor quality potential results are likely to be biased and may lead to wrong conclusions. Yet, it does not follow that econometric approaches are useless. Many econometric detection studies were, at least in retrospect, able to confirm cartel activity. Examples include Porter and Zona (1993, 1999) and Bajari and Ye (2003). Note that it should come as no surprise that all these studies make use of data that can be observed relatively easy, e.g., quantities, market shares, bids for procurement contracts.

Unlike the previous two approaches, theoretical research is not bound by data limitations. In order to delineate competition and collusion we must be able to predict what the competitive and the collusive situation looks like. We have already discussed the difficulties related to the competitive benchmark. In addition, we are confronted with the problem that the set of collusive equilibria is typically, in the words of Jean Tirole, an ‘embarrassment of riches’.⁵⁴ This problem is particularly severe when solutions of the model belong to the set of competitive and the set of collusive equilibria, i.e., when both sets overlap. If this is the case, then we might be able to understand why firms coordinate on a certain equilibrium, but we cannot be sure if that particular equilibrium is the result of (imperfect) competition, overt or tacit collusion.⁵⁵

⁵³Arguably, the sample bias will be lower for cartel cases in times where antitrust law enforcement was less strict or even absent. Yet, cartels that were legal, probably behave in different ways too, which implies the sample is probably biased in a different way. This certainly holds for the Netherlands, where cartels were sometimes announced in newspapers or otherwise known to almost everyone. It has been argued that in many markets this was due to ‘a culture of cooperation’. The recent cartel in the Dutch construction sector is a good example.

⁵⁴Tirole (1988)

⁵⁵See Harrington (2008) for an extensive discussion of problems related to cartel detection theories.

4.4.3 *On the Problem of Defining a Workable Benchmark*

Determining the competitive benchmark is a delicate matter. To what extent the three approaches that were discussed in Section 3 can be used partly depends on the available information and the idiosyncrasies of the case at hand. For example, one can use the pre- and post-cartel periods as a proxy for what competition looks like in a particular market. However, if one aims to search for cartel activity *ex ante*, the required information may often not be available. As remarked by Posner (1970), it is potentially very useful to look for similar markets in order to get a feeling for what competition may look like in the industry that is under scrutiny. Yet, one always should be very careful when comparing markets for the simple reason that no two markets will ever be exactly the same.

To illustrate that *a priori* obvious cases may still lead to false allegations, consider the *Bakers of Washington* case. From the mid 1950s through 1964, the retail price of bread in Seattle was about 15 percent higher than the U.S. average price, which was used as the competitive benchmark. In 1964, the Federal Trade Commission found the bakers guilty of price fixing mainly based on the peculiar price patterns that seemingly left no room for doubt that these bakers were indeed guilty of fixing prices. Yet, Newmark (1988) provides an alternative explanation, which indicates that the high pricing may not have been caused by a conspiracy. First, he shows that prices were higher in the West than in the rest of the United States due to higher retail margins, higher labour costs and a higher normal rate of return. Second, he provides evidence that the sharp price decline was not so much due to the order by the Federal Trade Commission to cease price fixing, but instead because some cheap low-quality brands of bread were sold in that period, i.e., new low-cost entry in the market.

4.5 Detecting Incomplete Cartels

In this section, we discuss how economic theory can be used to detect incomplete cartels. So far, we did not make a distinction between the detection of an all-inclusive cartel and the detection of an incomplete cartel. From a cartel detection perspective, however, taking account of (expected) cartel size can be of particular importance. In fact, when an investigator suspects a cartel to be less than all-inclusive, it may not be needed to formulate a competitive benchmark in ways described above. The reason is that the behavior of cartel members presumably differs from that of fringe members and, if so, the competitive fringe can be used as benchmark. As remarked by Porter (2005, p. 155),

“...the non-inclusive nature of the cartel may lead to evidence of its existence. A non-inclusive cartel can be easier to detect, as outsiders can serve as a standard of comparison.”

Blair and Romano (1989) is a good example of how economic theory can be used to detect an incomplete cartel. The authors describe a real-world case in which they, B&R Associates, had to provide an argument that would prove the innocence of one of five firms that were accused of fixing prices. The proposed test is based on a simple,

but sound result in economic theory on mergers and cartels. The central idea is that if the firm they represented was indeed innocent it most likely would have increased its output in response to the output reduction by the other four.⁵⁶ The firm was found ‘not guilty’, because its production increased after the date the cartel became effective.

This detection test is very attractive due to its simplicity and it probably can be applied successfully in quite a few instances. However, the test also suffers from some weaknesses. For example, suppose that there is a positive demand shock around the same time. This may very well lead to an increase in cartel production, which means the cartel could quite naturally pass the test. A potential solution for this problem would be to look at relative changes in market shares instead of changes in firm production, because these are neutral to demand shocks. Another, more severe problem with this test, is that it is often unclear when exactly the cartel started its operations. This information is crucial, because output levels of all firms tend to fluctuate over time. As a result, this test probably is effective as detection tool only in the presence of some other signals of cartel activity.

More generally, detection methods that work on the premise that a cartel is incomplete are effective only when the behavior of insiders and outsiders is *significantly* different. The next subsection illustrates that this might not necessarily be the case.

4.5.1 A Variance Screen for Collusion: an Example (1)

In ‘A variance screen for collusion’, Abrantes-Metz *et al.* (2006) propose a test to screen markets for the presence of hard-core cartel activities. The method aims to identify suspect firm behavior by comparing prices set by individual firms with average market pricing behavior in a given period. It is conjectured that during collusive regimes the average price of cartel members is significantly higher and price volatility is substantially lower in contrast with price levels of non-participants or compared to periods in which competition is not restricted. The intuition for this hypothesis is that a cartel at best only partly responds to exogenous changes in the market, because renegotiations are typically costly and risky. Meetings in ‘smoke-filled rooms’ are therefore likely to be organized only when new circumstances (e.g., cost shocks) imply substantial losses or threaten the stability of the coalition. Two main advantages of the variance screen are that no cost data are required to implement the method and that the test is hard to beat, because it presumably will be difficult for the cartel to coordinate on price patterns with a sufficiently high variance.

Yet, when Abrantes-Metz *et al.* apply their test to the Louisville retail gasoline market for the period 1996-2002, an industry that is known to be sensitive to anti-competitive behavior, no (cluster of) suspect petrol stations could be identified. The authors conclude that the difference in the coefficient of variation, i.e., the standard deviation normalized by its mean, is not big enough across retail gasoline stations to indicate the existence of a conspiracy. Obviously, this somewhat surprising result may simply mean that no cartels were active in this industry from 1996 until 2002. There exists however an alternative explanation, which supports the view that the

⁵⁶See Chapter 2 of this thesis.

Louisville retail gasoline market in fact may well have been infected, but quite naturally escaped detection by the variance screen. The reason is that the best response of non-participants to the pricing policy of the cartel may be such that their mean prices and price volatility closely follow that of cartel members.

To analyze how the pricing policy of an incomplete cartel may affect prices set by outsiders consider a market with n firms with identical cost structures of the form $TC(q_i) = cq_i + F$ that face the following Shubik-Levitan demand function for differentiated products,

$$q_i = a - (1 + \gamma(1 - \frac{1}{n}))p_i + (\frac{\gamma}{n}) \sum_{j \neq i} p_j, \quad (4.2)$$

with $\gamma \in [0, \infty)$ being a parameter that indicates the degree of product differentiation. When $\gamma = 0$, products are independent, while as $\gamma \rightarrow \infty$ products are standardized and, from a consumer perspective, close substitutes. Firms maximize the following profit function,

$$\max_{p_i} \pi_i = (p_i - c) q_i = (p_i - c) \left[a - (1 + \gamma(1 - \frac{1}{n}))p_i + (\frac{\gamma}{n}) \sum_{j \neq i} p_j \right].$$

Taking the first-order condition equal to zero yields the competitive equilibrium price, which is the same for all firms and equal to,

$$p^N = \frac{a + (1 + \gamma(1 - \frac{1}{n}))c}{2 + \gamma(1 - \frac{1}{n})}. \quad (4.3)$$

Now suppose a subset of k firms collude in price and that $f = n - k$ firms form a competitive fringe of independent outsiders. Let K denote the set of cartel members. Obviously, such an incomplete cartel is only reasonable for intermediate values of γ . The cartel maximizes the following profit function,

$$\max_{p^c} \pi^c = (p^c - c) q^c = (p^c - c) \left[a - (1 + \gamma(1 - \frac{k}{n}))p^c + (\frac{\gamma}{n}) \sum_{j \notin K} p_j \right].$$

Maximizing with respect to p^c and rearranging terms yields the best-response function of the cartel,

$$p^c = \frac{a + (1 + \gamma(1 - \frac{k}{n}))c + (\frac{\gamma}{n}) \sum_{j \notin K} p_j}{2(1 + \gamma(1 - \frac{k}{n}))}.$$

Independent outsiders maximize the following profit function,

$$\max_{p_i^o} \pi_i^o = (p_i^o - c) q_i^o = (p_i^o - c) \left[a - (1 + \gamma(1 - \frac{1}{n}))p_i^o + \gamma \frac{k}{n} p^c + (\frac{\gamma}{n}) \sum_{\substack{j \neq i \\ j \notin K}} p_j \right].$$

Taking the first-order condition and leveling to zero yields the best-response function of an individual outsider,

$$p_i^o = \frac{a + (1 + \gamma(1 - \frac{1}{n}))c + \gamma \frac{k}{n} p^c + (\frac{\gamma}{n}) \sum_{\substack{j \neq i \\ j \notin K}} p_j}{2(1 + \gamma(1 - \frac{1}{n}))}.$$

The equilibrium prices of fringe firms and the cartel can now be determined by combining the best-response functions. In equilibrium, the cartel and the fringe respectively set the following prices,

$$p^c = \frac{a(2n + \gamma(2n - 1)) + c(2n + \gamma(4n - 2k - 1)) + \gamma^2(\frac{f}{n})(2n + k - 2)}{4n + 2\gamma(3n - k - 1) + \gamma^2(\frac{f}{n})(2n + k - 2)}, \quad (4.4)$$

and

$$p^o = \frac{a(2n + \gamma(2n - k)) + c(2n + \gamma(4n - k - 2)) + \gamma^2(\frac{f}{n})(2n + k - 2)}{4n + 2\gamma(3n - k - 1) + \gamma^2(\frac{f}{n})(2n + k - 2)}. \quad (4.5)$$

Note that $p^c > p^o > p^N$ for $k \geq 2$, which holds trivially. That is, the cartel sets a higher price than in competition, but this creates a positive externality for fringe firms, which optimal response is to increase their price levels too, although by a lower amount. Notice that if the price increase of the competitive fringe is sufficiently high, applying the variance screen may not lead to the conclusion that cartel members have ‘unusual high means’. Yet, even if cartel members have an unusually high mean, this may be due to other circumstances. Indeed, as the authors rightly explain, one has to be careful drawing conclusions based solely on relatively high mean prices of a cluster of firms, because other factors may cause these effects. For example, if gasoline stations take a strategic position in the market (e.g., a main highway) it seems natural that, *ceteris paribus*, average prices will be higher than average prices of stations located in less advantageous positions.⁵⁷ An unusual price-variance may therefore be a better ‘tell-tale sign’ of anticompetitive behavior.

The variance in prices is determined by both the frequency and magnitude of price changes in a given period. Arguably, the cartel will modify its price level less frequently and when it does these changes are likely to be significant. In the gasoline market one may think of oil shocks or substantial changes in the exchange rate. Within the setting above, the effect of a change in marginal costs on the competitive price level is given by,

$$\frac{\partial p^N}{\partial c} = \frac{1 + \gamma(1 - \frac{1}{n})}{2 + \gamma(1 - \frac{1}{n})} > 0, \quad (4.6)$$

⁵⁷In fact, Abrantes-Metz *et al.* (2006) finds unusual high means on major cross roads and for gasoline stations that, in most cases, do not have a close competitor. The remaining part of the high means are located in the center of Louisville.

while for the cartel and the fringe firms the effect of a cost change on the price level is respectively,

$$\frac{\partial p^c}{\partial c} = \frac{2n + \gamma(4n - 2k - 1) + \gamma^2(\frac{f}{n})(2n + k - 2)}{4n + 2\gamma(3n - k - 1) + \gamma^2(\frac{f}{n})(2n + k - 2)} > 0, \quad (4.7)$$

and

$$\frac{\partial p^o}{\partial c} = \frac{2n + \gamma(4n - k - 2) + \gamma^2(\frac{f}{n})(2n + k - 2)}{4n + 2\gamma(3n - k - 1) + \gamma^2(\frac{f}{n})(2n + k - 2)} > 0. \quad (4.8)$$

Note that $\frac{\partial p^N}{\partial c} > \frac{\partial p^o}{\partial c} > \frac{\partial p^c}{\partial c}$ for $k \geq 2$, which always holds. That is, cost changes have the lowest impact on the cartel price, but the fringe firms also pass on a lower amount to consumers than they would have done in competition. Hence, in the presence of an incomplete cartel the variance in prices of non-participants will be, *ceteris paribus*, lower than in competition.

When the competitive fringe ignores the pricing behavior of the cartel, the variance screen is likely to be an effective method of detection. However, the pricing behavior of the competitive fringe is typically positively correlated with that of the cartel. If the variance screen is able to discriminate between cartel members and the competitive fringe therefore essentially depends on the magnitude of the effects described above.

4.6 The Need for Industry Specific Detection Tests

A great many problems have to be dealt with before the results of a cartel detection technique can be convincingly used as indirect evidence of collusion. However, at least part of the problems mentioned so far can be overcome by the development of industry specific detection methods. Not every market is the same and detection methods that potentially work well in one industry cannot automatically be applied to other industries.⁵⁸ In this respect, cartel detection could perfectly fit the school of New Empirical Industrial Organization (NEIO), which, in many ways, is a response to the classic structure-conduct-performance approach. Simply put, within this relatively new paradigm, the focus is on a single or small collection of industries, which are analyzed with the help of game theoretical and econometric techniques.

Taking the NEIO approach in cartel detection has several advantages. First and foremost, it allows for a better understanding of a particular industry. This, for example, will make it easier to define a reasonable competitive benchmark. Industry studies often contain detailed information about how a particular industry functions. Also, people who are working in the field (e.g., managers, employees, customers) potentially can provide valuable information about what competition normally looks like. In addition, it is possibly easier to understand what collusive behavior would look like within a particular environment. In many cases, focusing on one (type of) market

⁵⁸That economic tests must take account of the specific characteristics of an industry is, for example, argued by Baker and Bresnahan (2007).

allows for a more precise model and consequently makes it easier to take account of the specificities of that particular industry.

A good example of this approach is the study on the American automobile industry by Bresnahan (1987). He uses data from the American Automobile Industry in the years 1954, 1955 and 1956. The paper concludes that the developed competitive model better explains the year 1955, which was characterized by an extensive supply expansion, while the collusive variant of the model better fits the years 1954 and 1956. The model that has been used is one of product differentiation and the central hypothesis is that noncompeting firms are likely to have a different effect on comparative statics of price and quantity with respect to demand elasticities. The idea is that whether your closest neighbor belongs to you or another firm does not matter so much under a collusive regime, while in competition it matters a great deal if your closest substitutes belong to you or not. Therefore, if there exists a sufficient number of firms in the market (in that enough substitutes are available) competitive behavior will be significantly different from collusive behavior. The way to test this is by keeping structural hypotheses about cost and demand constant, while changing the behavioral hypothesis from collusive to the competitive equilibrium and then apply a likelihood model to determine which one is best able to explain the situation.

Another example is the study conducted by Porter (1983). He uses aggregate time series price and quantity data for the period 1880-1886 in order to examine if switches took place between collusive and noncooperative behavior. The cartel under consideration is known as the Joint Executive Committee, which is an anticompetitive agreement between railroad companies formed in 1879. The author finds evidence that the cartel broke down a couple of times (1881, 1884 and 1885), i.e., a reversion to noncooperative behavior. He conjectures that these reversions were caused by an unanticipated change in demand reflected by an unusually low market share for at least one firm.⁵⁹ These reversions, in turn, can be used to identify periods of collusive conduct. In particular, Porter finds that periods of cooperative behavior are best described by Cournot behavior, which essentially means that the cartel successfully restricted competition, but at the same time failed to set joint profit-maximizing prices.⁶⁰

We can imagine that methods like these can be applied by an antitrust authority *ex ante* given that sufficient data are available. Admittedly, however, such approaches are potentially quite sensitive to model specifications as is illustrated by the contributions of Porter (1983) and Ellison (1994). Yet, even though both these studies yield significantly different results, they have in common that collusive periods could be identified. As such, the design of economic detection tests that are tailored to certain (type of) industries is promising.

The following example illustrates the potential problem of a cartel detection method that fails to take into account the special characteristics of the market to which it is applied.

⁵⁹For a theoretical foundation of this behavior, see Green and Porter (1984).

⁶⁰Ellison (1994) analyzed the same data set, but with a slightly different model. In particular, he takes into account possible serial correlation in the error term of the supply equation. He finds evidence that the cartel approximately set prices that maximized joint profits during collusive periods.

4.6.1 A Variance Screen for Collusion: an Example (2)

The effectiveness of a detection device very much depends on the idiosyncrasies of the industry to which it is applied. Dynamic pricing in retail gasoline markets often exhibit a cyclical pattern similar to that of so-called Edgeworth cycles as described by Maskin and Tirole (1988).⁶¹ Such a cycle typically starts with a relatively high market price after which firms alternately undercut each others prices slightly, which implies slowly decreasing average market prices. Note that an important feature of Edgeworth cycles is that firms set their prices sequentially, which implies there exists a system of leaders and followers. At the cycle bottom firms face a coordination problem, because even though there is an incentive to hike the price, nobody is willing to be the first to relent. This coordination problem is more severe when more parties are involved. One way to escape from this prisoner's dilemma is by installing a system of collusive price-leadership in which the incomplete cartel announces a new high price that may or may not be followed by fringe firms.⁶² If the attempt to hike prices is successful, i.e., followed by non-participants, a new Edgeworth cycle starts.⁶³ This system of cyclical pricing patterns and collusive price-leadership has certain implications for the price variance of both the cartel as well as the fringe firms.

Competitive firms tend to adapt their prices also in case of minor cost changes, which, *ceteris paribus*, will yield a higher variance in prices. Note, however, that when the incomplete cartel operates as price-leader in markets that exhibit cyclical price patterns there exists another important price effect in particular when market prices are at the bottom of a cycle.⁶⁴ Consider the same scenario as described above and suppose the cartel announces a new high price. Such a price-hike is successful only when all outsiders relatively quickly respond by increasing their own prices. However, followers typically raise their prices by a lower amount, which means that the price-variance of cartel members *increases* compared to that of non-participants. The impact on the difference in price variance is even stronger when the attempt to hike the price is unsuccessful.⁶⁵ Indeed, when followers do not react or do not respond fast enough they will force the cartel back to the original price level. Again, this makes prices set by cartel members *more* volatile than that of the competitive fringe. The net effect on the price-variance of a cartel acting as price-leader in markets with cyclical price patterns is therefore at best ambiguous.

⁶¹See, for example, Castanias and Johnson (1993) who recognize that price cycles observed in Los Angeles in the late 1960s and early 1970s are closely resembled by Edgeworth price cycles. See also Eckert (2003) and Noel (2007).

⁶²Collusive price-leadership in gasoline markets is confirmed by Scherer and Ross (1990). A real-world example of collusive-price leadership in gasoline markets is analyzed by Wang (2005).

⁶³Note that high prices may be sustained for a while, because station operators recognize that reducing their own prices induces price cutting at rival stations and, moreover, once the undercutting process starts it may take some time to return to higher margins. See Borenstein and Shepard (1996).

⁶⁴Note that during the 'price war phase' everybody is undercutting everybody so that huge differences in price variance between firms are likely to be absent.

⁶⁵Attempts to hike the price tend to fail frequently as is illustrated by Wang (2005).

4.7 Discussion

The economics of cartel detection is an emerging field. Currently, most developments occur in academics, but an increasing number of antitrust agencies is starting-up an active detection policy. With firms becoming increasingly aware of antitrust enforcement and the risk of taking part in a cartel agreement we cannot rely on finding too many “smoking guns”, i.e., direct evidence of cartel meetings, in the future. Fortunately, it is typically beyond the power of firms to hide the results of the agreement and a substantial part of these effects is likely to be present in economic data and market indicators. Historically, economic theory has proven to be very valuable in identifying markets that are conducive to collusion, but it will in all likelihood become increasingly important in the assessment of firm behavior in those markets in the near future.

At the same time, it must be admitted that the number of cartel detection methods that is currently available to carry out such “crime scene investigations” is at best modest. In addition, this chapter reveals that the development of a workable method of detection is quite a challenge. Yet, as has been stressed already several times, it is arguably worth the effort in the long run. At this point, it is important to re-emphasize that part of the reward will naturally remain invisible. Taking part in a conspiracy becomes, *ceteris paribus*, less attractive, the higher is the chance of being caught and there are good reasons to believe that firms do take these (expected) costs into account. For example, we know that in procurement auctions some firms consciously submit phony bids or refrain from bidding to make sure that the ring remains unnoticed. Furthermore, the studies by Christie and Schultz (1994) and Christie *et al.* (1994) suggest that market players collectively defect from the agreement once “the heat is around the corner”. In summary, cartel detection helps to achieve one of the prime goals in antitrust enforcement; deterrence.

Part of the benefits of cartel detection that in principle can be assessed in money terms should stem from cartel cases that were the result of a successful application of one or a combination of detection techniques. To the best of our knowledge, so far none of the detection techniques discussed have been successfully applied in antitrust practice *ex ante*.⁶⁶ Consequently, only time will tell if the methods currently available are successful in the pro-active search for anticompetitive conduct. However, it is important to keep in mind that if a certain cartel detection method appears to have led to wrong conclusions in a particular case this not automatically implies that the method is ineffective. One can never be one hundred percent sure if some cartel activity takes place, i.e., cartel detection is utterly probabilistic. Hence, no matter how solid the cartel detection technique that is used is, there will always be room for an alternative explanation for observed firm conduct.

To what extent cartel detection can help to achieve the goal of desistance will not only depend on the quality of detection methods. In part it will also depend on legal developments. One key issue, which has been ignored in this chapter, is if and to what extent judges will accept circumstantial evidence of collusion to constitute ‘reasonable

⁶⁶To be clear, by detection techniques we mean here those methods that focus on cartel behavior.

doubt' that legitimizes further in-depth investigations. Currently, there seems to be no conclusive answer to this matter, although in all likelihood econometric evidence alone will not be sufficient.⁶⁷ In light of future research, a comprehensive legal analysis that relates potential economic detection results to the amount and quality of evidence that is needed to file a case would be a very welcome contribution. In this respect, a better understanding of what distinguishes overt from tacit collusion will prove to be important. That is, in-depth investigations typically imply a violation of some basic rights, which is easier to defend when there exists a significant chance of finding direct physical evidence. Economic research should therefore aim to find ways in which market data and other economic indicators could be used to delineate implicit and explicit cartel agreements.

Next to that, the development of economic detection methods that can be applied to more regular markets should be encouraged. This chapter shows that already quite some tests have been developed to detect bidding rings. This is understandable, because the data that are needed to implement the test are relatively easily accessible. Also, cartels are a prevalent phenomenon in auctions. Yet, a substantial part of cartels that have been discovered were operating in industries that were not organized by means of auctions. More generally, potentially a lot of progress can be made through the development of detection devices that are specifically designed for a certain (type of) industry. An example of such a detection method is presented in the next chapter.

⁶⁷See, for example, Froeb and Shor (2005) who claim that "Econometric evidence alone is unlikely to meet the burden for criminal prosecutions, though it may form a substantial part of evidence in a civil trial."

5

Tracing the Base: A Topographic Test for Collusive Basing-Point Pricing

“If he be Mr. Hyde. . . I shall be Mr. Seek” - Robert Stevenson¹

5.1 Introduction

One of the responsibilities of antitrust authorities is to discover cartels.² In the past, collusive agreements have been brought to light by disgruntled employees, complaining rivals, customers seeking antitrust damages, and remorseful cartel members applying for leniency.³ It is likely, however, that cartels increasingly succeed in preventing such leaks. Conspiring managers are smart to involve employees as little as possible, for example, and to assure that their closest competitors and direct purchasers benefit rather than suffer from the existence of the cartel.⁴ Cartels may further find ways to dissuade their members to apply for leniency, for example by each putting up collateral that falls to the other cartel members in the event of one of them defecting from the collusive agreement. For these reasons, it is essential that public enforcement produces a sufficiently high probability of discovery across the board through active cartel detection.

¹See Stevenson, R.L. (1886, 2003), “*The Strange Case of Dr. Jekyll and Mr. Hyde*,” Signet Classic, p. 49.

²This chapter is joint work with Maarten Pieter Schinkel and published as Bos and Schinkel (2008). Excellent research assistance has been provided by Eelko Ubels

³See McAnney (1991) or Levenstein and Suslow (2006). In the past decade, cartel enforcement in both the US and the European Union has relied heavily on applications for leniency.

⁴For a mechanism to tacitly imply direct purchasers in an upstream cartel, see Schinkel *et al.* (2008).

Economic theory can advise in the design of detection mechanisms by identifying ‘tell-tale signs’ of collusion. An emerging research area of cartel detection, recently surveyed in Porter (2005) and Harrington (2008), puts emphasis on revealed characteristics of cartel behavior for this purpose.⁵ The approach is to distinguish collusive from competitive patterns in common observables. Significant output reductions, structural breaks and reduced variations in time-series of prices, or in prices across producers and customers can be indicative of collusion. Likewise, sudden changes in conditions of sales or quality, unusually low entry and exit frequencies, persistent excess capacity, or strongly correlated market shares and stock price values may raise suspicion.

When antitrust authorities monitor markets for suspicious behavior, clever conspirators devise ways to dodge detection. Even if only a limited number of active cartels is so discovered—after all, every test can be beaten—tests for collusion will often make internal coordination amongst prospective cartel members more difficult and dangerous. The need to try to fly under the radar is likely to depress cartel profits. In this sense, active detection can make collusion less attractive. It is essential, therefore, that the antitrust authorities stay on top of their game of hide-and-seek with professional colluders, and arm themselves with the latest detection technology.

One particularly sophisticated way in which cartels can design and cover up their price cartel is by abusing the basing-point system. Basing-point pricing can develop under delivered pricing, which is pricing for product delivered inclusive of transportation costs. When producers are somewhat dispersed, they can each calculate their bids from locations that are not necessarily their own mill sites. As a result, buyers receive the same price quote from (at least two) sellers regardless of the sellers location, even if their transportation costs to the locations are different. Delivered pricing is typically observed in industries that produce a homogeneous bulky product that requires specialized transport to project sites, which forms a substantial part of total costs. Examples are gravel and liquid concrete in construction, oversized steel beams in shipbuilding and toxic chemicals in industry.

In competition, producers can each use the location of the rival that is nearest to the order location to calculate their freight-inclusive bid and so gain a modest economic profit, depending on the amount of geographical product differentiation. The basing-point pricing system’s potential for anticompetitive abuse lies therein that a geographically isolated cluster of firms can conspire and agree to all use one or more distant base locations instead. These collusive bases could be the mill location of a distant outsider. It can also be a fictitious point far away from every cartel member. Calculating prices from such ‘phantom bases’, all cartel members charge transportation costs that are not actually made. In this way, the firm that gets the order can reap cartel profits in the form of phantom-freight charges. With a regular stream of more or less evenly placed orders, all cartel members gain from this agreement.

Basing-point pricing came under antitrust scrutiny in the U.S. in the first half of the 20th century. In 1924, the Federal Trade Commission (FTC) ordered the United States Steel Corporation to stop using Pittsburgh as the basing point for steel pro-

⁵See also Porter and Zona (1993, 1999), Bajari and Ye (2003) and Abrantes-Metz *et al.* (2006).

duced and sold in the Chicago area.⁶ In *Cement Institute* in 1948 the U.S. Supreme Court found that basing point pricing facilitates collusion, after which the practice was long treated as a *per se* antitrust violation.⁷ After that, the FTC brought only few basing-point pricing cases, however, and ultimately lost, after it had stepped up enforcement again in the late 1970's, in *Boise Cascade* (1980) and *DuPont* (1984).⁸ *Boise Cascade* concerned the softwood-plywood industry, in which Portland served as a single base location from which incumbent producers in the coastal areas of the Pacific Northwest together with new entrants in southern states like Louisiana and New Orleans charged freight on deliveries nationwide.⁹ In *DuPont*, uniform delivered pricing was alleged to have been one of several devices that facilitated collusion in the sales of ethyl.¹⁰ Suspicious patterns in delivered pricing have also been found in Europe, Japan and Great Britain.¹¹ Nevertheless, courts have increasingly held that without explicit evidence of an agreement, plaintiffs must be able to show anticompetitive intent, or a lack of legitimate business reasons for the challenged practices.¹² In the past quarter century, there has been little or no enforcement of basing-point pricing and the debate about it seems to have withered.

Delivered pricing with basing points is quite a neat collusive mechanism, however. It has a number of benefits for a cartel that manages to abuse the system, including high profits and a low risk of discovery. Once the principle pricing rule is commonly understood—which is relatively straightforward, since it uses the actual transportation costs formula—basing-point pricing requires little to no communication between the cartel members. Essentially, all that is needed to fix the prices of orders of all different shapes and sizes are the geographical coordinates of the commonly used collusive base point location(s). These could be just two numbers— x degrees latitude, y degrees longitude—communicated publicly, without raising any suspicion. There is no need for executive meetings in which to allocate heterogeneous project or negotiate complex menus of prices and volumes, which leave evidence of collusion. Furthermore, it is often hard to distinguish between collusion and competition purely on the basis of observed transaction prices. The practice has the appearance of genuine competition and may thus escape common screening methods and damages assessments.

In this paper, we develop a method to test for the presence of collusive basing-point pricing. Our test uses transaction data (volumes and prices) and information on customer locations to recover the base locations used to calculate the bid by a suspected cluster of firms. The traced bases are compared to the mill locations. If the found base locations are far away from all firms, there may have been collusion. If they coincide with various firm locations, bidding is consistent with competition. On this

⁶Matter of the United States Steel Corporation, 8 F.T.C. 1 (1924). See Commons (1924).

⁷*Cement Institute vs. FTC*, 333 U.S. 683 (1948). See Kaysen (1949) and Loescher (1959).

⁸*Boise Cascade Corp. v. FTC* 637 F.2d 573 (9th Cir. 1980) and *E.I. duPont de Nemours & Co. v. FTC* 729 F.2d 128 (Second Cir., 1984).

⁹See Carlton (1983).

¹⁰See Hay (1999).

¹¹See Greenhut (1981) and Haddock (1982).

¹²See Loescher (1980) and Cabou *et al.* (1992).

principle, we develop a measure for the likelihood of collusion, as well as a software that can be used to screen large amounts of data for possible collusive bases.

Our method for tracing basing points is conceptually akin to a forensic technique known as geographic profiling, which is used to find serial offenders, such as murderers or arsonists.¹³ When several bodies are found at different crime scene locations, which are suspected to have all been the work of one and the same serial killer, this method uses the location coordinates to triangulate the area where the serial killer is likely to reside. The same principle can be applied to the various locations of fires in cases of arson. Geographical profiling uses the principle that the modus operandi of serial offenders typically include that such criminal acts are committed at locations that form a distance pattern around the perpetrator's base. The area may be further shaped by geographical restrictions on ease of access, such as lakes, forests, mountains and industrial areas. Such constraints are incorporated in software of detailed 3D maps and used to locate jeopardy areas for repeat offenders. The success of these computer programs in supporting law enforcement has been such that a number of them are endorsed by the U.S. Department of Justice.¹⁴

The remainder of this paper is organized as follows. The next section explains the practice of basing-point pricing in detail and summarizes the debate on its competitive nature. Section 5.3 develops a model of spatial competition between a given and isolated cluster of firms. We introduce conditions that suffice to fully characterize competitive and collusive basing-point pricing in this model. Section 5.4 illustrates how in this setting collusive basing-point pricing may escape some of the known methods of detection. Section 5.5 develops the principle of our topographic test. In Section 5.6, we operationalize the test in an algorithm. Section 5.7 concludes. A software developed for tracing bases is introduced in Appendix A.

5.2 Basing-Point Pricing

Under delivered pricing, shipping to customers is arranged by each seller at a price per order that includes production costs, transportation and freight insurance. Industries that apply delivered pricing often produce homogenous bulky products for manufacturers or wholesale intermediaries in geographically clustered production plants.¹⁵ The nature of these types of products typically means that they require special means of transportation, to guard quality, for example in the case of liquid concrete or asphalt, or to assure safety, such as in oversized loads or toxic chemicals. Although such specialized transportation could in principle be arranged by the buyer and offered 'free-on-board' at the mill, delivered pricing is often practical and efficient for such

¹³See Rossmo (1999) and Canter (2003).

¹⁴These softwares include Rigel[®], Dragnet and CrimeStat—see <http://www.ojp.usdoj.gov/nij/maps/gp.html>. Most of these programs originated in academic research. Several of them have also been commercially developed for use in law enforcement.

¹⁵See Soper *et al.* (1991) for examples of industries that applied delivered pricing.

products. Producers have specially modified trucks and equipment, for example, and can combine deliveries.¹⁶

Delivered pricing allows for basing-point pricing when firms are geographically (somewhat) dispersed. The principle mechanism of the basing-point pricing system is illustrated in Figure 5.1, which gives a bird's eye view of a local market with three firms and three customers. Consider this cluster as a regional market that is isolated from distant competitors. Production of a homogeneous commodity takes place at three competing mills, located at y_j , $j = 1, 2, 3$. These locations are connected by the dark-lined triangle in the figure. Customer projects are situated around the mills at x_i , $i = 1, 2, 3$. At each project location, there is a fixed demand for the commodity at a sufficiently high willingness to pay. The cost of production at all plant locations are identical.

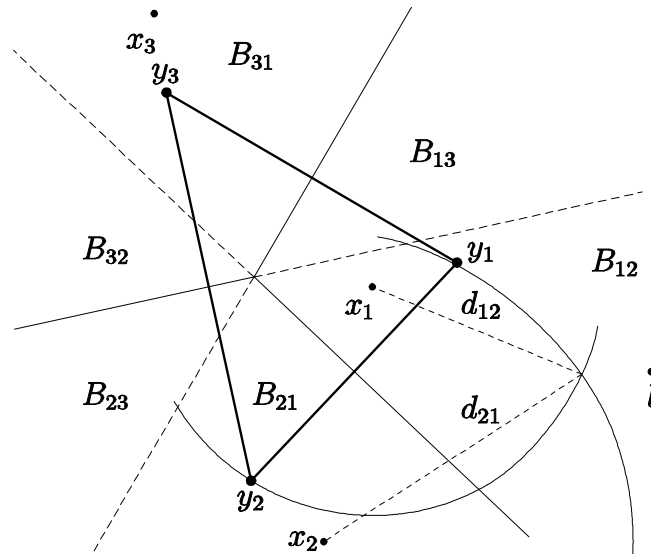


Figure 5.1 Bird's eye view of basing-point pricing in a regionally isolated market.

The geographical locations of plants and projects give each producer a natural home market of customers that are closest to it. As long as no two plants are exactly at the same location, producers are in imperfect competition. When mills all ship at the same transportation costs that increase in the distance of shipping, the boundaries of each producer's home market are halfway between the various production locations. The solid outgoing line-grid in Figure 5.1, which is constructed by placing lines orthogonal to the lines that connect the firm locations, halfway between each two plants, represents these boundaries.

¹⁶A part of the literature is concerned with the question under which conditions free-on-board pricing is the optimal competitive strategy, rather than delivered pricing. See Espinosa (1992).

When mills compete for customers, each producer can make a positive economic profit on those customers that are located within its home market as a result of geographical product differentiation. Under competitive basing-point pricing, sellers can perfectly price discriminate and charge each customer a total delivered price that is equal to the production costs for the volume plus the transportation costs that the nearest competitor would need to make to serve this customer. Since the seller itself is closer to the customer, it makes an economic profit that is equal to the difference between the price charged and the actual shipping costs, so-called ‘phantom freight’. The nearest competitor cannot undercut this offer, since that would imply a loss, for it does need to make the actual transportation costs charged.¹⁷ At best, the nearest competitor could meet the bid, so that the customer receives two identical price offers from these firms, even though their distances to the project site differ.

The extended dashed-line grid in Figure 5.1 divides the geographical market further into customer location categories. In the following, area B_{jv} is referred to as the ‘base area’ in the home market of firm j in which it charges customers from the location of firm v , where v indicates the nearest competitor. The geographical characteristics of the market thus define a multiple basing-point pricing pattern in competition.

If transportation costs as a function of distance differ between the firms, for example because a firm benefits from combined shipping or a more efficient means of transportation, the natural market division is more complicated and firms may deliver profitably in markets closer to other firm’s production locations. This practice is known as ‘cross-hauling’.¹⁸

Producers can also decide to conspire to exploit market circumstances and all use a remote base point location in calculating their offers. This eliminates competition on transportation costs between them. The collusive bases would be one or more arbitrarily agreed locations, or natural focal points, such as a common port location, as long as it is sufficiently far away from customers. Alternatively, the location of a distant rival that is not part of the cartel but within reach of its customers is a natural candidate base location. This choice assures that the rival cannot undercut the local cartel and enter into the market of any of its members.

By charging the going mill rate plus freight from the collusive base(s), cartel members can reap anticompetitive phantom freight. Such an abuse of the basing-point system can sustain a cartel if customers and mills are located relatively closely together, and producers can agree on a collusive base point that is sufficiently far removed from their cluster. In such circumstances, applying the collusive base returns higher (expected) phantom freight for all cartel members. Whether a market can profitably be cartelized in this way depends on the distribution of demand, relative to the locations of the mill sites. It is difficult to formulate an acceptable set of general conditions under which it always is. In constructed cases—for example with a cluster of demand very far away from a cluster of mill locations, around a mine in a thinly populated area, at comparable distance to which there is a rival mill as well—a local basing-point cartel

¹⁷If this does happen, it is referred to as ‘freight absorption’.

¹⁸See Haddock (1982).

may not work. At the same time, however, it is straightforward to construct generic examples in which collusive basing-point pricing is a very lucrative proposition.

In Figure 5.1, point l is such a collusive phantom base location when transportation costs increase linearly in distance, so that there is a positive cartel profit for each member as long as the distance from l to any of the customers x_i is larger than the distance from x_i to any cartel member's nearest competitor y_v . That is, outside the circle with radius d_{12} between customer 1 and firm 2, from which firm 1 charges customer 1 transportation costs in competition, firm 1 earns a net cartel profit. Analogously, firm 2 makes a cartel profit on its customer 2, since l is outside the radius d_{21} from x_2 to y_1 . Finally, location l is obviously profitable for firm 3 to use on its customer 3 (boundary not drawn). In general, there are many profitable single collusive base point candidates, as well as more complex collusive systems that use multiple basing points, also when transportation costs are not linear in distance.

As emphasized in the introductory section, collusive basing-point pricing has a number of benefits for a cartel in this type of markets. The principle is relatively easy to implement, and once established, there is no need for extensive communication between the cartel members. This helps to sustain the cartel while creating minimal proof of a conspiracy. Orders of all shapes and sizes are all individually overcharged on the basis of agreement on just two geographical coordinates. Moreover, the bids that a customer receives from different firms are identical, without the cartel needing to coordinate each case separately. The system furthermore returns a wide bid spread across customers, as project sites are each located differently relative to the collusive base. This price pattern can be similar to one that would emerge under competition, when firms that are far removed from a project site would refrain from bidding, so that customers only receive the identical bids of the (minimally two) closest rivals. This makes that the cartel can pass for a competitive industry without raising suspicion with customers or the antitrust authorities, even when bidding information is publicly available. In transaction price data only, collusive basing-point pricing need not leave any obvious traces.

At the same time, the system allows cartel members to straightforwardly monitor their coconspirators for cheating.¹⁹ Each cartel member is able to simply calculate what should have been the collusive offers of its coconspirators, to see if the others may be undercutting the agreement. Furthermore, collusive basing-point pricing includes a natural credible punishment strategy to discipline discovered cheaters. If a cartel member is found out defecting from the collusive basing-point strategy, its production plant can be made a 'punitive base', that is, the base location used by the other cartel members in their subsequent price offers.²⁰ This punishment is straightforward to implement. While costly for the defector, the punishment is relatively inexpensive for the loyal cartel members—somewhat depending on their distances to the defector relative to their customer bases. Finally, collusive basing-point pricing is a particularly suitable strategy for a cluster of closely located producers that seeks to jointly prevent entry from distant rivals into their local cluster's home market. By using the mill

¹⁹See Benson *et al.* (1990).

²⁰Loescher (1959) reports on this punishment strategy having been used in the cement industry.

locations of those rivals as collusive base locations, the local cartel deters entry by limit pricing, while all cartel members can make a positive cartel profit.²¹

The competitive nature of delivered pricing has been much debated. Machlup (1949) and Stigler (1949) were among the first to argue that basing-point pricing can facilitate collusion. Thisse and Vives (1988, 1992), Benson, *et al.* (1990) and Levy and Reitzes (1993) have shown that it is unlikely that the basing-point system develops noncooperatively in competition without explicit communication. Others, including Clark (1943; 1949), Carlton (1983) and Haddock (1982, 1990), have defended basing-point pricing as an efficient form of spatial competition and an unlikely system to be adopted by a monopolist or a cartel.

A number of weaknesses of basing-point pricing as a mechanism of collusion have been pointed out in this literature. There need not always be a natural candidate collusive base location, and since individual cartel profits increase in the distance of the collusive base location to its own customers, each cartel member would prefer the collusive base to be far away from its home market. In its most basic form, the cartel arrangement does not include an explicit division of the market if all cartel members offer the same high price to each customer.²² When buyers choose more or less randomly with which producer to order, each cartel member risks receiving little or no orders for considerable periods of time, even if orders may work out more or less evenly over all firms in the long run. As a result, the cartel may need to apply an end-of-year compensation scheme—and a risky bookkeeping that comes with it. This would add to the complexity of the arrangement and to the risk of discovery.

Uncertainty in the distribution of orders, combined with quoted prices often being private information, could also make the detection of chiseling more, rather than less difficult compared to more explicit cartel agreements. Also, firms would regularly need to deliver in other firms' home markets, thus transporting inefficiently and suffering from freight absorption.²³ In addition, delivery outside the home market would rarely be observed among competing firms, unless transportation costs differ strongly between different producers, so that if it is observed, it would create suspicion of an illegal arrangement in and of itself. Taken together, these drawbacks would make basing-point pricing an unlikely choice for coordinating price fixing. Instead, where delivered pricing and cross-hauling are applied, these pricing strategies would typically be efficient.²⁴

Most of these suggested problems with collusive basing-point pricing can, however, be quite easily overcome by small modifications of the basic form of abuse outlined above. By using multiple collusive bases and/or rotating them at regular intervals, for example, the cartel can distribute its profits evenly over time. The cartel could further agree on simple distance related market sharing rules, such as a fixed surcharge to

²¹In *Addyson Pipe and Steel*, a group of southern and Mid-western iron pipe manufacturers (the Addyston Group) collectively tried to block entry into its home market by establishing a collusive basing-point at the nearest cluster of competitors in the east. See *U.S. v. Addyston Pipe and Steel*, 175 U.S. 211, 1899. See Levy and Reitzes (1993) for a theoretical analysis.

²²See Carlton (1983).

²³See Smithies (1941; 1942).

²⁴See Haddock (1982).

tender request from outside the natural home market. This would imply an efficient division of the market along the lines of home markets, while only slightly increasing communication. Observe also that customers are indifferent which cartel member to order from at equal delivered price quotes, so that they are quite likely to obey natural market boundaries and select the firm closest to them. If lasting trade relationships do not assure this, slight discounts will, so that the cartel can easily prevent suspicious and inefficient cross-hauling.²⁵ These additions to the basic agreement do increase the complexity and risk of discovery of the cartel somewhat.

Empirical studies have remained inconclusive about the question whether basing-point pricing is competitive or collusive. Machlup (1949) and Loescher (1959) describe collusive patterns in the cement industry. Karlson (1990) revisits the industry data over the period 1910-1940 to find little support for competitive explanations for the practice and suggestion of collusion in the 1930's. Gilligan (1992) concludes that single basing-point pricing is unlikely to have facilitated collusion in the softwood-plywood industry. The study finds that while the FTC's interference reduced the phantom freight rents for the southern producers, it left the markups of the northern plywood producers unaltered, which is interpreted as suggestive of imperfect competition among the non-base-site producers, and not of collusion amongst the base-site producers.

5.3 A Model of Basing-Point Pricing

Consider a world represented by an unbounded Cartesian plane and a single homogeneous product. Customer projects are distributed over the plane. At location x_i , customer i has a perfectly inelastic demand for the product of order size q_i , for which it has a sufficiently high willingness to pay. Let $f(x)$ be the probability density function of a normalized single unit order from project location x . Total expected demand from a bounded region $\Omega \subset \mathbb{R}^2$ is then given by

$$\int_{\Omega} f(x) dx. \quad (5.1)$$

We assume that demand is distributed over the world so that there are some projects everywhere, as follows.

Assumption 1 The project demand probability density function $f(x)$ is continuously differentiable and satisfies,

- (i) $\int_{\Omega} f(x) dx = 1$;
- (ii) $f(x) > 0$ for all $x \in \Omega$.

Each customer orders with the firm offering the lowest price.

There is a fixed and finite number of production mills J that all produce at identical and constant marginal cost per unit of volume c , a fixed cost per order F , and without capacity constraints. Mills are located at geographically dispersed locations y_j , $j =$

²⁵See Loescher (1959).

$1, \dots, J$. We assume that plant locations are given and there is no entry. As a result, there is a given amount of spatial product differentiation. It is possible that two or more firms are very close together, and profits are zero in competition when they are exactly in the same location. Production plants can in principle be located anywhere in this world, yet the nature of products involved—raw materials in bulk with lower input transportation costs than output transportation costs—can dictate certain regional constraints—such as mining area's and locations that have easy access to water ways. Whenever possible, however, plants are likely to be located where demand is high, due to a substantial transportation cost component. In the following, we consider the cluster of J producers in isolation and further abstract from any outside distant rivals.

Let the physical distance between two locations, α at (a_α, b_α) and β at (a_β, b_β) , be indicated by $d_{\alpha\beta}$ and defined by the Euclidian distance

$$d_{\alpha\beta} = \sqrt{(a_\alpha - a_\beta)^2 + (b_\alpha - b_\beta)^2}. \quad (5.2)$$

Hence, $d_{y_j x_i}$ is the actual distance between firm j at y_j and customer i at x_i .

Using distances, let the natural home market of firm j be defined as the set of customers H_j for which firm j is closest in terms of physical distance. That is,

$$H_j = \left\{ \text{all customers } i \text{ for which } d_{y_j x_i} = \min_k (d_{y_k x_i}) \text{ for } k = 1, \dots, J \right\}.$$

We assume that each customer exclusively belongs to one home market.

Each natural home market can be divided into one or more so-called 'base areas'. A base area in the home market of firm j is denoted B_{jv} and defined as,

$$B_{jv} = \left\{ \text{all customers } i \in H_j \text{ for which } d_{y_v x_i} = \min_{k \neq j} (d_{y_k x_i}) \text{ for } k = 1, \dots, J \right\},$$

in which the distance $d_{y_v x_i}$ is between customer i and the producer v at y_v that is located closest to that consumer, except for producer j . Each firm j has one or more base area's, depending on the number of direct competitors, *i.e.*, competitors whose home markets border at H_j . The collection of base areas with the same second subscript v is referred to as a 'base group' G_v . Let B_j be the set of base areas of firm j —that is, collections of all base areas with the same first subscript. We denote the total number of elements in B_j by V_j .

Given the geographical spread of all players, to deliver an order to one of its customers, firm j incurs transportation costs. Let the transportation costs of firm j for transporting volume q_i from its plant at y_j to a customer project located at x_i be of the general form $T_j(q_i, d_{y_j x_i})$, which are positive, continuous and differentiable functions. We assume that firms have identical transport cost structures, and that total cost of transport are increasing in distance and the volume of the order.

Assumption 2 $T_j(q_i, d_{y_j x_i})$ satisfies for all $j = 1, \dots, J$,

- (i) $T_j(q_i, d_{y_j x_i}) = T(q_i, d_{y_j x_i})$;
- (ii) $\frac{\partial T(q_i, d_{y_j x_i})}{\partial d_{y_j x_i}} > 0$ and $\frac{\partial T(q_i, d_{y_j x_i})}{\partial q_i} \geq 0$.

Note that by assuming that production and transportation costs are identical across production plants, we rule out efficient cross-hauling. This implies that when cross-hauling is observed, it would in itself be indicative of collusion. We therefore also exclude the possibility of cross-hauling under collusion by assuming a natural market division.

Assumption 3 If a customer is indifferent between price offers from different producers, shipping will be from the mill that is closest to the project.

In competition, this assumption is innocuous. In case of cartel pricing, it implies an efficient market division mechanism, which further enhances internal monitoring by the cartel, as observed cross-hauling exposes a deviating member. Under these assumptions, it is not possible to detect collusive basing-point pricing by simply checking for cross-hauling.

In delivered-pricing systems, prices and profits are determined by distances. The system thus allows firms to price-discriminate between individual customers. Consider d_{lx_i} , the distance between customer i and a given geographical location l —be it a firm or a distant location. Customer i receives an all-inclusive price offer P_{ji} from firm j that is constituted as follows,

$$P_{ji} = cq_i + F + T(q_i, d_{lx_i}), \quad (5.3)$$

in which F may include any fixed component in transportation costs.

The profit firm j makes on an order delivered to customer i is,

$$\pi_{ji} = P_{ji} - cq_i - F - T(q_i, d_{y_jx_i}). \quad (5.4)$$

Substituting (5.3) in (5.4) yields,

$$\pi_{ji} = T(q_i, d_{lx_i}) - T(q_i, d_{y_jx_i}), \quad (5.5)$$

which is the difference between actual transportation cost and the transportation cost charged to the consumer.

Obviously, $\pi_{ji} = 0$ if $d_{lx_i} = d_{y_jx_i}$, which would be the case if producer j calculates freight cost from its own mill location. Profits are positive for orders for which the actual freight costs, $T(q_i, d_{y_jx_i})$, are smaller than the freight costs calculated to the customer, $T(q_i, d_{lx_i})$ by the amount of phantom freight over $d_{lx_i} - d_{y_jx_i}$. Profits per project increase in phantom freight, since the actual transportation costs over $d_{y_jx_i}$ are fixed by locations x_i and y_j , so that $\frac{d\pi_{ji}}{dd_{lx_i}} = \frac{\partial T(q_i, d_{lx_i})}{\partial d_{lx_i}} > 0$ for all $j = 1, \dots, J$ by Assumption 2. Furthermore, the difference between real and calculated freight charges for a given volume may increase in the volume of the order q_i , depending on the functional form of $T(\cdot)$.²⁶

²⁶Note that in the basing-point price structure, the frequency of orders over time is important for profits, and the frequency of trips is likely to be inversely related to the volume of sales per order if customers, for example, save up demand to combine several orders into one. By assuming the order-frequency and the volume to be independent of price, we ignore these substitution effects.

5.3.1 Competitive Basing-Point Pricing

In competition, firms play a one-shot game in which plants $j = 1, \dots, J$ at y_j simultaneously chose an action P_{ji} , depending on a chosen location l , in response to realized demand in the form of a tender request for q_i from project site x_i , leading to a pay-off $\pi_{ji} = T(q_i, d_{lx_i}) - T(q_i, d_{y_jx_i})$. The following result establishes that in competition delivered pricing allows each mill to charge competitive phantom freight from the location of their nearest competitor to each project in its home market.

Proposition 5.1 *In competition, firm j uses mill location y_v as a base for all customers $i \in B_{jv}$ and y_j for all $i \notin H_j$.*

Proof. Consider a customer $i \in B_{jv}$ and, without loss of generality, suppose that the mills j and v use a basing point located at l such that $d_{lx_i} > d_{y_jx_i}$, $d_{lx_i} > d_{y_vx_i}$ and $d_{lx_i} < d_{y_kx_i}$, $\forall k \neq j, v$. Hence, $T(q_i, d_{lx_i}) - T(q_i, d_{y_kx_i}) < 0$, $\forall k \neq j, v$, so that we solely concentrate on mills j and v . Calculating the bid from l yields, according to (5.3), $P_{ji} = P_{vi}$, which by Assumption 3, gives mill j the trade, so that $\pi_{ji} > 0$ and $\pi_{vi} = 0$. This, however, cannot be optimal for v , because lowering P_{vi} slightly to $P_{vi}^* = P_{ji} - \varepsilon$, with ε small but positive, shifts the trade to mill v to generate $\pi_{vi} > 0$ by equation (5.5). The best response for mill j then is to always match the bid of v , as long as $P_{ji}^* = P_{vi}^* > cq_i + F + T(q_i, d_{y_jx_i})$. Hence, mill v 's (weakly) optimal bid equals $P_{vi}^{**} = cq_i + F + T(q_i, d_{y_vx_i})$, because any further lowering of the bid implies a loss. Mill j therefore optimally submits $P_{ji}^{**} = P_{vi}^{**}$, yielding concomitant profits $\pi_{vi} = 0$ and $\pi_{ji} = T(q_i, d_{y_vx_i}) - T(q_i, d_{y_jx_i}) > 0$. Note that, with respect to consumer i , all other mills k are in the same position as mill v . Therefore, all mills, but the mill that is located closest to the customer project, will use their own plant location as basing point. ■

Serving its natural home market yields firm j a total profit equal to the sum of the profits it expects to earn on the customer projects in each of its V_j base area's B_{jv} . These expectations are based on $f(x)$, the distribution of single order unit demand, and therefore only depend on distances. Our assumptions on the distribution of customer projects assures that H_j is non-empty for all j . Hence, all mills are used as a base by at least one rival firm in competition. The expected profits earned by firm j under a competitive basing-point regime can therefore be written as

$$\pi_j^c = \sum_{v=1}^{V_j} \int_{B_{jv}} f(x) (T(d_{y_vx_i}) - T(d_{y_jx_i})) dx, \quad (5.6)$$

where $T(d_{y_vx_i})$ is the distance between project location x_i and the nearest rival that defines base area B_{jv} . The superscript c refers to 'competition'.

Clearly, firms ability to earn positive economic profits depends on their relative positions in the geographical market. When mills are few and far between, competitive profits can be substantial. When they are clustered relatively close together, there is little room to price above true transportation costs. If two mills are almost exactly in the same location—say, next door to each other—they will effectively calculate transportation costs from their own product site and make almost zero profits.

5.3.2 Collusive Basing-Point Pricing

The vulnerability of the basing-point pricing system to collusion stems from the fact that the cartel members only need to agree on one jointly used collusive base location at any given point in time.²⁷ They do this *ex ante*, based on the expected profits they can each generate over their home market population. That is, the expected cartel profit of firm j is equal to

$$\pi_j^a = \int_{H_j} f(x) (T(d_{lx_i}) - T(d_{y_jx_i})) dx, \quad (5.7)$$

where l is a common collusive base location. The superscript a stands for ‘anticompetitive’.

Agreeing on the location of the collusive base among all members of the coalition is not straightforward. Certainly, the cartel has a collective interest in choosing a base location that is sufficiently far away from the main customer projects and the mill locations of its members, so that $\pi_j^a \geq \pi_j^c$ for all j .²⁸ Yet, a distant collusive base location will always be closer to some customer projects than the serving mill is, so that expected losses are to be accepted on part of the remote projects. Typically, therefore, collusive base candidates are in the tail of the customer project distribution. Another problem for the cartel is that different distant collusive bases are most profitable for different cartel members. More specifically, each member would prefer the collusive base to be located in the farthest corner of a distant coconspirator’s home market.

As explained above, when there are distant rivals that could poach the cartel’s market, the location of one or more of them—in particular the closest—are natural collusive base candidates. When there is not such a natural focal point, the class of possible collusive base locations is large. We can restrict this class by the requirement that each cartel member individually increases its expected profits under the cartel regime. Making use of the fact that, since actual shipping distances and volumes remain the same, total costs in competition and collusion are the same, that is, each individual cartel member j , $j = 1, \dots, J$, blocks all candidate collusive bases l , for which

$$\int_{H_j} f(x) T(d_{lx_i}) dx \leq \sum_{v=1}^{V_j} \int_{B_{jv}} f(x) T(d_{y_vx_i}) dx, \quad (5.8)$$

where $d_{y_vx_i} = \min_{k \neq j} (d_{y_kx_i})$, $k = 1, \dots, J$, for all $v = 1, \dots, V_j$. Although quite natural, note that this blocking condition is an assumption in the sense that it rules out Pareto improving side-payments between the cartel members. In this way, it creates

²⁷A cartel abusing the basing-point pricing system may also use a number of different collusive base locations, for example to support a certain profit sharing rule or to avoid detection. In the following, we will analyse the use of a single collusive base at any point in time. We briefly discuss how our results transfer to collusive multiple basing-point pricing in Section 5.7.

²⁸In this paper, issues of cartel stability are further ignored. It is assumed that a cartel exists whenever all members benefit from collusion over competition. Given the existence of a cartel surplus, internal cartel stability can be assured by standard means. In addition, the basing-point pricing system offers typical punishment strategies, as discussed in Section 2.

a larger area of blocked candidate collusive bases than strictly required to increase expected profits.²⁹

The set of geographic locations for which condition (5.8) holds for firm j , we call j 's 'blocking zone' \mathcal{B}_j . We have the following result.

Proposition 5.2 *For each cartel member j , $j = 1, \dots, J$, $y_j \in \mathcal{B}_j$.*

Proof. Suppose that $l = y_j$, then $\pi_j^a = \int_{H_j} f(x) (T(d_{y_j x_i}) - T(d_{y_j x_i})) dx = 0$. Since $\pi_j^c = \sum_{v=1}^{V_j} \int_{\mathcal{B}_{jv}} f(x) T(d_{y_v x_i}) dx \geq 0$, according to condition (5.8), l will be blocked as a collusive base candidate by firm j . ■

The area in which always at least one cartel member will block candidate base points, we refer to as the 'blocking zone'. It is found as

$$\mathcal{B} = \cup_{j=1}^J \mathcal{B}_j.$$

It follows directly from Proposition 5.2 that \mathcal{B} contains all of the cartel's plant locations. Figure 5.2 illustrates the blocking areas and the blocking zone for the geographic market depicted in Figure 5.1 above.

The area of candidate collusive base locations blocked by at least one cartel member is amorphous, because while profits are positively related to the distance over which phantom freight is calculated, they need not be linear in distance, which in turn is not linear in the coordinates of locations in Euclidean space. We can further say very little about the blocking zone, beyond that it contains all plant locations of the cartel members. In Figure 5.2, \mathcal{B} comfortably covers the convex hull of firm locations. This is not a general property, however, and it can also not straightforwardly be guaranteed by placing specific regularity conditions on the distributions of plant and project locations.³⁰

The zone of collusive base point candidates does have natural outer boundaries: no cartel will be able to calculate its orders from the end of the universe. The distance at which the collusive base can be put from the market is restricted, for example, by the maximum willingness to pay of customers, the locations of distant rival firms that are not included in the local cartel, and new entrants being attracted to the area. We

²⁹The alternative more general blocking condition

$$\sum_{j=1}^J \sum_{v=1}^{V_j} \int_{\mathcal{B}_{jv}} f(x) T(d_{y_v x_i}) dx \leq \sum_{j=1}^J \int_{H_j} f(x) T(d_{l x_i}) dx$$

which requires that each candidate collusive base increases the sum total of cartel profits, is less restrictive and requires coordination on a separate profit sharing rule.

³⁰Even if $f(x)$ is unimodal with a mode location ϕ around which the firms are distributed at more or less equal distance, and sufficiently steeply bell-shaped with ever fewer customer projects out in the tail ring—which can be interpreted as a high concentration of customers living in a condensed area, and the further away in all directions from this center, the thinner customer projects are spread on the ground—it cannot obviously be assured that a cartel will never choose to locate a single collusive base in their midst, in particular when the circle around the center of the market is large and contains many plants that are at close distance from each other.

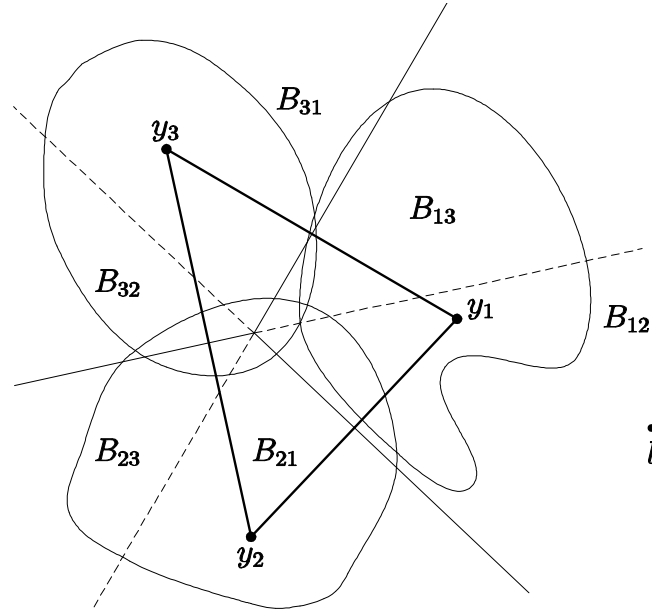


Figure 5.2 Cloud-shaped blocking zone in a regionally isolated market.

do not formalize these restrictions in this paper, since they are not directly relevant for our method of detection.

5.4 Detecting Collusive Basing-Point Pricing

The features of competitive and collusive basing-point pricing identified above allow us to discriminate between competitive and collusive basing-point pricing by a given cluster of firms. After a cartel has chosen its collusive base(s) and demand materializes, trades take place and generate a series of transaction data. To see the difficulty of detecting collusive basing-point pricing in such time-series, consider Table 5.1. It displays two transaction price patterns per unit of demand, in a regional market with 12 customers and 3 firms, located non-specifically at (460, 460), (440, 540) and (650, 440). Customer locations are indicated in the first column. Transportation costs increase multiplicatively in volume and distance, at a marginal cost of 0.15 per mile per unit.³¹ Marginal costs of production are 50 per unit. There is a fixed cost of 1500 per order. The percentage overcharge per project is given in the last column.³²

³¹That is, $T(q_i, d_{lx_i}) = td_{lx_i}q_i$ with $t = 0.15$.

³²The overcharge is measured as the percentage unit price increase relative to competitive prices, or $\frac{\text{Price/unit}_{\text{coll}} - \text{Price/unit}_{\text{comp}}}{\text{Price/unit}_{\text{comp}}} \times 100$.

Consumer	Volume	Price/unit _{comp}	Price/unit _{coll}	Overcharge (%)
(500,580)	980	70.50	91.15	29.3
(630,380)	1070	79.58	103.72	30.3
(490,520)	1160	61.36	97.33	58.6
(480,410)	1140	71.72	110.19	53.6
(520,460)	1260	68.16	100.46	47.4
(550,410)	820	67.49	104.73	55.2
(520,420)	900	71.40	105.72	48.1
(410,510)	1090	61.98	108.00	74.2
(600,430)	1100	72.84	98.17	34.8
(540,470)	930	68.72	97.84	42.4
(470,450)	830	66.04	107.18	62.3
(430,490)	1010	59.13	107.38	81.6

TABLE 5.1 Prices per unit and profits under competitive and collusive basing-point pricing.

Clearly, prices in the left-hand column are structurally lower than those in the right-hand column. The competitive price series has a mean of 68.2 per unit. The collusive base point prices were calculated from a single collusive base location at (727, 715), has a mean of 102.7. Facing just one series of observations, however, that is, without having the luxury of being able to tell the difference by comparison, it is difficult to see a suspicious pattern. The prices differ across consumers in both regimes. In fact, the example is constructed in such a way that the variance in the competitive price series is exactly equal to the variance in the collusive series—29.3. Hence, this industry is not likely to surface in a test based on price variance. The cartel overcharge per customer is nevertheless high. The cartel realizes substantial illegal profits over and above the profits under competition, with an average net overcharge of over fifty percent. In the following two subsections, we will show that cartels can indeed systematically exploit the basing-point system to avoid detection by known methods.

5.4.1 Variance Screens

Consider the application of a variance screen for cartels. It tests for local pockets of low variances in transaction prices on the insight that it is difficult for the cartel to create a natural price-variance, because regular price renegotiations, for example in response to cost shocks, are costly and risky. Therefore, prices are fixed for longer periods of time at a sufficiently large margin to cover intermediate cost increases. Several international cartels indeed in hindsight display periods of low variance. Harrington (2004) builds on this to develop a screen and study how cartels subsequently try to avoid it. Abrantes-Metz, *et al.* (2006) presents a mean-variance screen for clusters of firms colluding, surrounded by competitors. Suspect firms are then those which, within a certain period of time, ask prices with a relatively high mean and a low standard deviation.

Obviously, such suspicion can only arise in comparison to some benchmark, be it an earlier period known to have been competitive, a competitive fringe in the same market, or similar firms competing in a different market. In our model of full collusion in a local cluster, such a benchmark is not available. But suppose a threshold value of

low price variance somehow does exist.³³ Then the cartel can strategically locate the collusive base location so as to exactly mimic that variance level and not surface in the test. To see this, consider a world with discretely located customers, $i = 1, \dots, I$ and let p_i be transaction price per unit of demand under collusive basing-point pricing from base location l that customer i accepted. That is,

$$p_i = c + \frac{F + T(q_i, d_{lx_i})}{q_i}, \quad (5.9)$$

with the sample mean being

$$\hat{p} = \frac{1}{I} \sum_{i=1}^I \left(c + \frac{F + T(q_i, d_{lx_i})}{q_i} \right). \quad (5.10)$$

The sample variance of a given price series is given by,

$$\hat{\sigma}_{p_i}^2 = \frac{1}{I} \sum_{i=1}^I (p_i - \hat{p})^2 = \frac{1}{I} \sum_{i=1}^I \left(c + \frac{F + T(q_i, d_{lx_i})}{q_i} - \hat{p} \right)^2, \quad (5.11)$$

so that $\hat{\sigma}_{p_i}^2$ is a function of the collusive base location l . A proper choice of the collusive base location equates the price variance in collusion to that under competition, which is fixed by the locations of the players. It is not difficult to construct examples in which there is a continuum of collusive bases with the same price variance. Figure 5.3 illustrates such an iso-variance curve.

The figure is a bird's eye-view of the market that generated the data in Table 5.1. The customer locations are black dots, the triangular area in the middle is the convex hull that connects the three production plant locations. The union of the dashed lined areas around each of the plant locations delineates the blocking zone. A collusive phantom base location outside this zone assures positive net cartel profits for all cartel members. Point l is the collusive base location at (727, 715).

The ellipse through l is an iso-variance curve of base locations at which the variance in prices when that location is used to determine collusive offers is identical to the price variance under competition. The part of this iso-variance curve outside the blocking zone collects all collusive base candidates that would generate net expected cartel profits for all firms and escape a variance screen that uses the competitive price variance as a benchmark. On the basis of expected demand $f(x)$, a cartel can analogously determine a set of collusive base locations with an expected variance that would be consistent with competition.

5.4.2 Bid-distance Correlation

A different method for identifying collusive patterns in markets with a geographical dimension uses the distance between the bidding firm and the project location as a

³³Obviously, a known competitive benchmark value for the price mean would be impossible to mimic for the cartel while still making net cartel profits in the long-run.

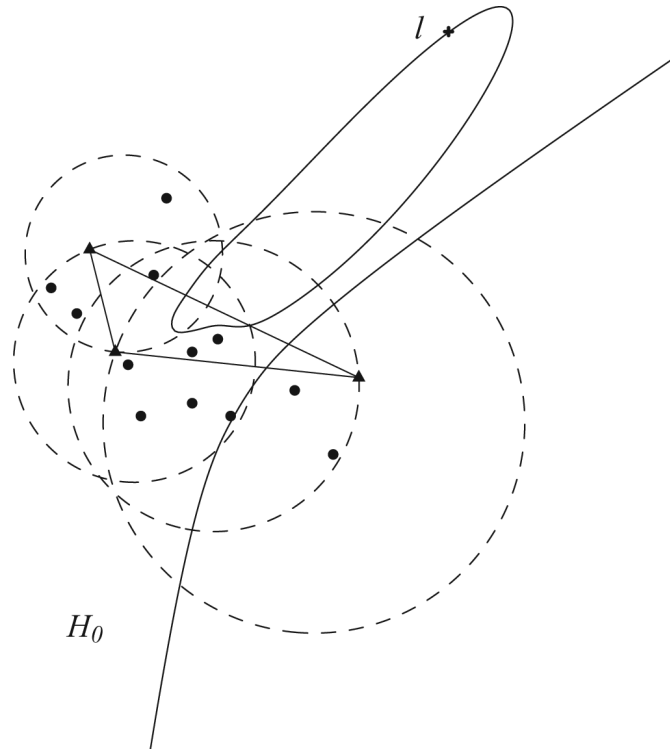


Figure 5.3 Bird's eye view of base locations in the collusive base zone resulting in a price variance that mimics competition.

proxy for production costs. In public procurement auctions for construction of maintenance contracts, one would expect firms to submit higher bids the farther away their plant is located from the project site. Porter and Zona (1993) confirms suspicion of a bidding ring for a Long Island highway construction project in this way. Porter and Zona (1999) reports on evidence of collusion in procurement auctions for school milk in various districts in Ohio, in which the bids of several firms decreased in the distance from their home market to the schools. In Bajari and Ye (2003) the approach is extended to a screen. In first-price sealed bid procurement auctions, the bid-cost ratio is regressed on various explanatory variables, amongst which distance to the project site. Applied to bids for highway maintenance contracts in the late 1990's in the Mid-West, several firms were found to bid suspiciously relative to their location.

In this paper, we focus on collusive schemes that only leave trails of transaction prices. The basing-point pricing system allows cartels to avoid detection methods that rely on patterns between distance and transaction prices by carefully choosing their collusive bases. To see this for the example of materialized demand above, consider a simple approach to the bid-distance screen, using an ordinary-least-squares test. That is, suppose that a market auditor estimates a relationship between unit transaction

prices, p_i , and the distances between the project and the nearest rival from which the bid would have been calculated in competition, d_{y_0, x_i} . That is,

$$p_i = \alpha + \frac{\beta}{q_i} + \gamma d_{y_0, x_i} + \varepsilon_i, \quad (5.12)$$

in which α captures marginal costs of production, β estimates fixed factors and $\varepsilon_i \sim N(0, \sigma_{\varepsilon_i}^2)$. In competition, one would expect γ to be significantly larger than zero, so that the bids on average increase in the distance between the delivering firm's competitive base and the project location. Therefore, the cartel needs to choose collusive base locations that also return an average positive relationship between these distances and bids. It can do this as follows.

The vector of estimated coefficients is

$$\begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \\ \hat{\gamma} \end{bmatrix} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{p}, \quad (5.13)$$

where \mathbf{X} contains 1's in the first column, values $\frac{1}{q_i}$ in the second column, and distances d_{y_0, x_i} in the third column. Vector \mathbf{p} contains observed prices. The estimator of the standard deviation is,

$$\hat{\sigma}_{\hat{\gamma}} = \sqrt{\frac{\sum e_i^2}{I-3}(\mathbf{X}'\mathbf{X})^{-1}}, \quad (5.14)$$

where e_i^2 are the residuals in the sample.

The cartel can avoid suspicion raised by this test, as long as it makes sure that the null hypothesis $H_0 : \gamma \geq 0$ is not rejected. H_0 will be rejected with 95% confidence if $\hat{\gamma} < t_{0.05} \times \hat{\sigma}_{\hat{\gamma}}$, where the relevant value for t is derived from $\hat{\gamma} \sim t(I-3)$ and equal to 1.833 in the example of Table 5.1 above. For all candidate collusive base locations to the left of the curved vertical line in Figure 5.3, this test cannot reject the null-hypothesis. Note that the area includes the iso-variance curve through l , so that the cartel in our example passes the screen based on transaction price variance, as well as the test for bid-distance correlation. As before, a cartel can ex ante set its collusive bases to avoid rejection of H_0 on expectation.³⁴

5.5 Tracing the Base

We have established that it is not obvious how to detect collusive basing-point pricing in historical prices, particularly so when cartels are clever in their choices of collusive base locations. Transaction prices in principle suffice, however, for our test based on the recovery of base locations. To see how this test works, we specialize the general structure of bids under basing-point pricing (5.3) to

$$T(q_i, d_{lx_i}) = t_d d_{lx_i} + t_q q_i + F_T, \quad (5.15)$$

³⁴If information on losing bids is available in addition, cartels can select bases to avoid patterns in those.

where t_d is the marginal cost of transportation per unit of distance and t_q per unit of volume, and F_T is a fixed component. Assuming a transportation cost structure that is linear in distance and volume implies that the profits from phantom freight per order are independent of the volume of the order. This assumption simplifies our analysis in this section for expositional purposes. Note, however, that our test can handle a variety of other bid-structures as well.³⁵

5.5.1 Base Recovery

For now, suppose that for a given sample of transaction data it is known that a single collusive base location was used. Using the definition of distance (5.2), prices are related to base locations, as follows:

$$P_{ji} = \overbrace{(c + t_q)}^{\tilde{c}} q_i + \overbrace{(F + F_T)}^{\tilde{F}} + t_d \sqrt{(a_l - a_{x_i})^2 + (b_l - b_{x_i})^2}. \quad (5.16)$$

Customer project locations (a_{x_i}, b_{x_i}) and transaction data (P_{ji}, q_i) are known, so that equation (5.16) has five relevant unknowns: $\tilde{c} = (c + t_q)$, $\tilde{F} = (F + F_T)$, t_d , plus the two coordinates of the implied base location (a_l, b_l) . In principle therefore five independent observations on price-quantity transactions of customers in one and the same base group suffice to determine the base location from which they were calculated. The observations have to be independent in the sense that the consumer locations in each tuple of five observations should not be perfectly aligned. If they are, there exists a mirror point to the applied base, so that the system is not uniquely determined.

In practice, the data on both transactions and project locations will typically be noisy as a result of calculation and measurement errors in production and transportation costs, and possibly missing factors. A least-squares estimation procedure can deal with this. We assume that

$$P_{ji} = \tilde{c} q_i + \tilde{F} + t_d \sqrt{(a_l - a_{x_i})^2 + (b_l - b_{x_i})^2} + \epsilon_i, \quad (5.17)$$

with $\epsilon_i \sim N(0, \sigma_{\epsilon_i}^2)$ on the total bid.

We cluster all transaction data per base group, so that all tuples of observations in each data subset correspond to one and the same base location, both in competition and collusion. Since mill and project locations are given, this sorting of the data is straightforward. Let I_{G_v} be the number of independent customer projects located in base group G_v . The base location of base group G_v is in principle recoverable if I_{G_v} is minimally equal to the number of unknowns in the bid structure.³⁶ We can now

³⁵This is discussed further in Section 5.7.

³⁶In the example given in Section 5.4, the base group that receives offers calculated from the firm at (460, 460) has five customers, that of the firm at (440, 540) four, and of the firm at (650, 440) three. Using the non-linear specification of transportation costs in footnote 31, which also implies five unknowns, it is only possible to trace two bases in the example: under a competitive regime, the mill locations of the firm at (460, 460) and a single collusive base in case of collusion. The system is underdetermined to find the other two competitive bases. Yet, we can just discriminate between competition and collusion on the basis of this information.

define the following criterion function

$$S = \sum_{i=1}^{I_{G_v}} e_i^2, \quad (5.18)$$

where e_i are the residuals in the observed transaction data, and determine our estimators to solve

$$\min_{\tilde{c}, \tilde{F}, \tilde{t}_d, a_l, b_l} \sum_{i=1}^{I_{G_v}} \left(P_{ji} - \tilde{c}q_i - \tilde{F} - t_d \sqrt{(a_l - a_{x_i})^2 + (b_l - b_{x_i})^2} \right)^2. \quad (5.19)$$

Figure 5.4 illustrates the principle of determining $\hat{l} = (\hat{a}_l, \hat{b}_l)$ in two dimensions, that is, for given values of \hat{c} , \hat{F} and \hat{t}_d . In this case, in the absence of noise in the data, three observations per base group suffice to uniquely determine the common base location from which all consumer prices were calculated. Each observation $(P_{ji}, q_i, a_{x_i}, b_{x_i})$ defines a circle of locations l around each consumer project location x_i . The intersection of these three circles is the common base. There is no such perfect intersection when the observations contain noise. Instead each candidate estimate of \hat{l} implies a difference e_i between the distance \hat{d}_i that corresponds with the observation on project i and the distance explained by \hat{l} . Our criterion now is to find \hat{l} so as to minimize the sum of the squared differences e_i . It is applied to simultaneously estimate \hat{c} , \hat{F} and \hat{t}_d as well.

Problem (5.19) does not allow for an explicit form for the estimators, nor does it need to have a unique global minimum.³⁷ We therefore have to resort to numerical methods of estimation. We use an iterative global search algorithm that is described in detail in Appendix A. It benefits from a number of constraints on the parameter spaces. The method is generally determinate and converges to a global minimum.

Each base group data subset that contains a number of independent observations equal to or larger than the number of unknowns (in this case five) returns one base location, so that J bases are recovered as long as the minimum data requirements are met. On the basis of our findings in Section 5.3, we can subsequently analyze the pattern of recovered bases to see if there is indication of collusive basing-point pricing. A pattern may be indicative of collusion if all firms have used one and the same distant base point. If instead various plant locations are found to have been bases, this would be consistent with competition. In order to further operationalize this basic distinction, in the next section we develop a screen for collusive basing-point pricing.

³⁷To see this, it suffices to note that the second-order own-derivate of S to a_l , which is

$$\frac{\partial S}{\partial a_l \partial a_l} = 2 \sum_{i=1}^{I_{G_v}} \frac{(a_i - a_l)^2 \left[\frac{1 + P_i - \tilde{c}q_i - \tilde{F} + d_{li}}{d_{li}} \right] - d_{li} (P_i - \tilde{c}q_i - \tilde{F} + d_{li})}{(a_i - a_l)^2 + (b_i - b_l)^2},$$

can be negative—e.g. when there are small differences in the a -coordinates and large differences in the b -coordinates. Hence, the 4×4 Hessian matrix is not positive semidefinite, so that S is not convex.

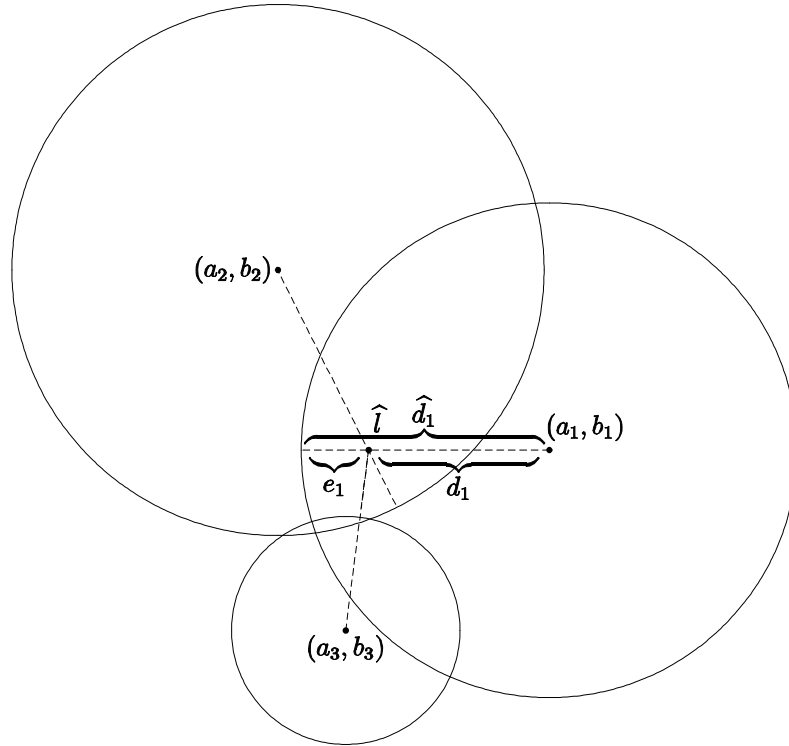


Figure 5.4 A least-squares point estimator of the base location.

5.6 A Likelihood Measure of Collusion

Let N be the number of basing points independently found. If producers colluded, their bids were calculated from one and the same base location, so that the bases found will be close together—and sufficiently far away from the cluster of mill locations. In contrast, a competitive world will yield a number of basing points of which the coordinates coincide with the coordinates of firm locations. We therefore take the sample average base location

$$\bar{a} = \frac{1}{N} \sum_{l=1}^N \hat{a}_l \text{ and } \bar{b} = \frac{1}{N} \sum_{l=1}^N \hat{b}_l,$$

and define the average distance to (\bar{a}, \bar{b}) as a measure of the geographical spread in the recovered bases.³⁸

$$\bar{\sigma} = \frac{1}{N} \sum_{l=1}^N \sqrt{((\hat{a}_l - \bar{a})^2 + (\hat{b}_l - \bar{b})^2)}.$$

The general distinct characteristics of competitive and collusive base locations derived above now allow us to discriminate between collusive and competitive basing-point pricing. First consider the following straightforward result on taking the average of recovered competitive basing points. Let \mathcal{C} denote the convex hull of mill locations, i.e., the smallest convex set containing all mills.

Lemma 5.1 *Without misspecification and noise in the data, under competitive basing-point pricing, $(\bar{a}, \bar{b}) \in \mathcal{C}$.*

Proof. If $\sigma_{\epsilon_i}^2 = 0$, by Proposition 1 all discovered bases (\hat{a}_l, \hat{b}_l) correspond to mill locations and therefore are in \mathcal{C} . Since \mathcal{C} is convex, the average of all discovered bases, (\bar{a}, \bar{b}) , is also located in \mathcal{C} . ■

In addition, since the average basing point under competition generally combines several bases, $\bar{\sigma} > 0$ under competition. When applied to a collusive world in which a single basing point has been used, the spread in the recovered bases is equal or very close to zero. In contrast, in case of collusion (\bar{a}, \bar{b}) will typically be located outside \mathcal{C} .

We use these distinct features to define a measure for the likelihood that there has been collusion in a market under consideration. Using only information on the locations of firms and customers, it is possible to construct the location (\bar{a}, \bar{b}) that would theoretically be produced under competition by applying the above described procedure. For that, it is required to determine which firms would be found by the test as a base for its nearest rival in competition—note that these are the firms indicated by the subscript of all base groups G_v that contain the minimally required number of independent observations to recover the base. Let this theoretical average competitive base location be l^* , at (\bar{a}^*, \bar{b}^*) with average spread $\bar{\sigma}^*$. We then define the following measure

$$\lambda = \frac{\max_j d_{l^* y_j}}{\bar{\sigma}^*}, \quad j = 1, \dots, J, \quad (5.20)$$

where the numerator is the longest of all distances between (\bar{a}^*, \bar{b}^*) and the firms on the boundary of \mathcal{C} .

Let $S^{\mathcal{C}}$ be the surface area of \mathcal{C} .³⁹ We construct a circle with radius $\lambda \times \bar{\sigma}$ around the mean recovered base location (\bar{a}, \bar{b}) and define the set

³⁸We use this measure rather than the standard deviation in distances to (\bar{a}, \bar{b}) , since the latter is a measure of symmetry, rather than spread. The standard deviation would be zero, for example, for recovered bases that are found evenly distributed on a circle around (\bar{a}, \bar{b}) , irrespective of the radius of that circle.

³⁹Note that firm locations should not all be aligned for \mathcal{C} to have a positive surface area. This implies that $J \geq 3$ for the test to work.

$$\mathcal{L} = \left\{ (a, b) : \sqrt{(a - \bar{a})^2 + (b - \bar{b})^2} \leq \lambda \times \bar{\sigma} \right\}.$$

Let $S^{\mathcal{L}}$ be the surface area of \mathcal{L} .

The likelihood of collusion is now related to the overlap between the convex hull of firm locations and the constructed circle around the recovered mean base location. We introduce the following measure:

$$LoC = 1 - \frac{S^{\mathcal{C}} \cap S^{\mathcal{L}}}{S^{\mathcal{C}}}. \quad (5.21)$$

The value of LoC is always between zero and one. Higher values would be indicative of collusive basing-point pricing. The following proposition defines boundary values.

Proposition 5.3 *Without misspecification and noise in the data, the LoC -measure discriminates perfectly between competitive and collusive basing-point pricing, with:*

(i) $LoC = 0$ in the case of competition; and

(ii) $LoC = 1$ in the case of collusion.

Proof. Suppose $\sigma_{\epsilon_i}^2 = 0$. (i) Under competitive basing-point pricing, by Proposition 1 and Lemma 5.1, $(\bar{a}, \bar{b}) = l^* \in \mathcal{C}$. Furthermore, $\bar{\sigma} = \bar{\sigma}^* > 0$, so that $S^{\mathcal{C}} \cap S^{\mathcal{L}} = S^{\mathcal{C}}$ by construction. Hence

$$LoC = 1 - \frac{S^{\mathcal{C}} \cap S^{\mathcal{L}}}{S^{\mathcal{C}}} = 1 - \frac{S^{\mathcal{C}}}{S^{\mathcal{C}}} = 0.$$

(ii) Under collusive basing-point pricing, all mills use the same collusive base $l = (a_l, b_l)$, which implies that $(\bar{a}, \bar{b}) = (a_l, b_l)$ and $\bar{\sigma} = 0$. Hence, $\lambda \times \bar{\sigma} = 0$, so that $S^{\mathcal{L}} = 0$, which yields

$$LoC = 1 - \frac{S^{\mathcal{C}} \cap S^{\mathcal{L}}}{S^{\mathcal{C}}} = 1 - \frac{0}{S^{\mathcal{C}}} = 1.$$

Note that typically l is a distant location. In a punishment phase, however, it is a single firm location. ■

Obviously, this perfect discrimination result only holds in the absence of noise in the data or misspecification errors. One way to apply the LoC -measure when there are such imperfections is to raise a red flag of likely collusion on an industry when the value of LoC is found to be above a certain absolute level, for example 0.5. Alternatively, the measure can be used to design a hypothesis test, also using the bid sample variance. To what extent measurement errors affect the usefulness of our test statistic depends on the structure and the size of the noise. Obviously, if the actual bid structure is different from the structure used to recover the bases, the error terms are not normally distributed and the test will be structurally off. A misspecification test with alternative bid structures could help to identify such problems.⁴⁰

Truly white noise on the bid structure, which may for example have resulted from miscalculations by the firms or measurement errors, need not affect the power of the

⁴⁰Possible directions in which to develop such specification tests are briefly discussed in Section 5.7.

LoC-test in practice, provided it is not too strong. White noise will dislocate (\bar{a}, \bar{b}) from l^* under competition, so that $S^{\mathcal{L}}$ generally no longer covers $S^{\mathcal{C}}$ and *LoC* values are larger than zero. Imperfect observations will also make that $\bar{\sigma}$ is positive under collusion. Multiplied by λ , this potentially results in a substantial surface area $S^{\mathcal{L}}$. When the collusive base is inside or not too far away from \mathcal{C} , this may result in partial overlap, leading to an *LoC* value that is smaller than 1, despite that fact that the mills are in fact colluding. Generally, however, $\bar{\sigma}^2$ converges to zero when the number of independent observations per base group becomes sufficiently large. In addition, a wider spread of customer locations and firm plants makes for a more accurate location of the collusive base, in particular when it is far away from the market. Also the blocking zone will often embed the hull and provide a safety belt against large overlap.

Using the method described in Appendix A, in Figure 5.5, we plot the value of *LoC* against σ_{ϵ_i} in equation (5.17) for two types of geographic markets: one in which the cartel locates the collusive base far away from all members, and one in which the cartel can profitably place the base inside the convex hull of firm locations.⁴¹ This analysis reveals that the *LoC*-measure remains a robust tool also when the size of the random noise in the price data increases.

In the upper two panels, 3 firms located at (460, 460), (440, 540) and (650, 440), as in the example above, produce with marginal costs of 5, fixed costs of 1500 and transportation costs equal to 1. They serve 25 customers per base group.⁴² There is some variability in demand at each project location, $q_i \sim N(1000, 50)$, capped at zero.⁴³ The upper-left graph provides a bird's eye view of the market. The upper-right panel plots the average *LoC* calculated over 10 drawings of price bids for each value for the standard deviation. Average profits under the cartel are substantial, ranging from 5 to 25 times competitive profits. The upper line shows the *LoC*-measure under collusion falling from 1, where the mean collusive overcharge was 7.31%.⁴⁴ The lower line going from 0 is the *LoC*-measure under competition. The figure shows that the measure can discriminate roughly for a standard deviation of up to 100 from an average collusive price bid of roughly 7000.

The lower two panels of Figure 5.5 are simulations with a constructed example of 7 firms that are located in a circle—as seen in the lower-left graph—in such a way that they can profitably collude with a base location in their midst.⁴⁵ All mills produce with a marginal costs of 5, fixed costs of 1500 and transportation costs of

⁴¹In these simulations, we have fixed $t_d = 1$, the true value of common transportation costs. See Appendix A.

⁴²The coordinates of these 75 consumer locations are drawn from two normal distributions, each with a mean equal to the sample mean firm coordinate and a variance equal to the sample variance in the firm locations.

⁴³The size of this standard error of demand is somewhat smaller but comparable to that in the example in Section 4, Table 1 above.

⁴⁴The collusive base was (927, 915) and transportation costs are linear in distance and volume in these simulations. From each base group there are 25 observations on transactions. The plotted lines in the figure are the average of 10 random price series under each regime. Consumer locations remained fixed over these runs. The mean transaction price in competition was 6655.94 and in collusion 7141.34.

⁴⁵These firm locations are (440, 630), (640, 890), (820, 610), (650, 440), (740, 480), (760, 800) and (460, 530).

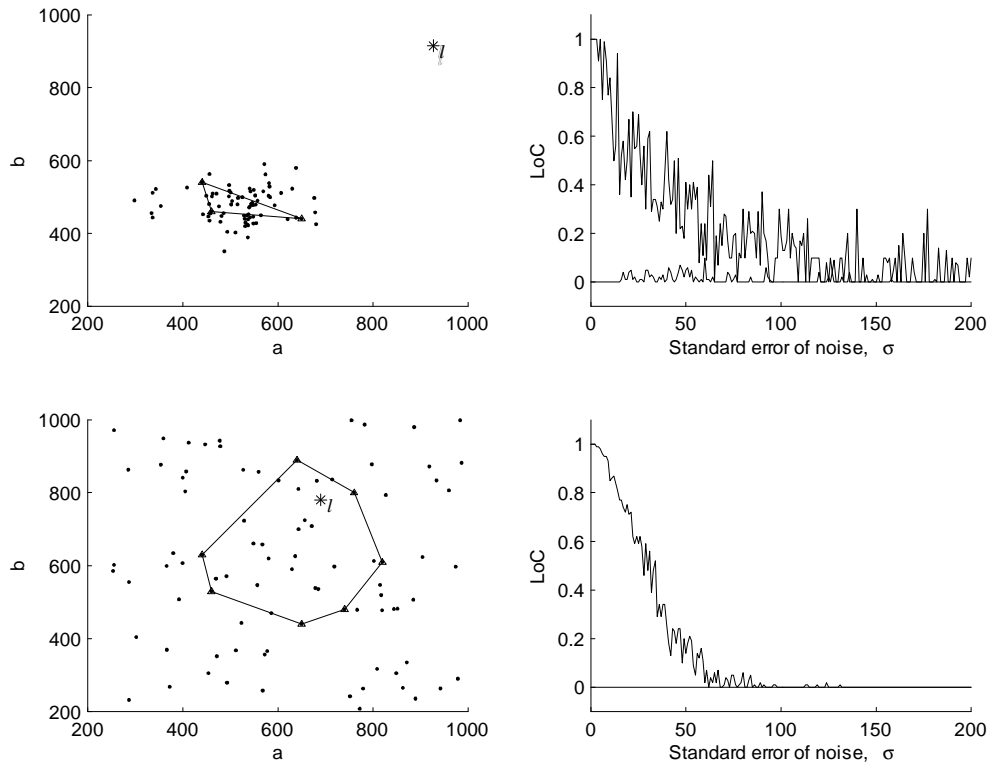


Figure 5.5 Robustness of the LoC-measure as an enforcement tool for two examples of collusive basing-point pricing.

1. The cartel profits of each firm are positive, ranging between a few hundred and a few thousand, but the mean cartel overcharge is modest, 1.97%—obviously, there are more attractive collusive base candidates outside the hull. Each of the 7 base groups contains 25 customer projects with variable demand, as in the previous example.⁴⁶ The *LoC*-measure under competition is a flat line. Under collusion, the *LoC*-value appears no longer to discriminate roughly at around $\sigma_{\epsilon_i} = 60$ from an average collusive price bid of about 7000.⁴⁷

These simulations have been generated with a computer program that uses a step grid search. Its structure is explained in Appendix A. The software can be operationalized to use the *LoC*-measure as a screen for anticompetitive behavior in large data sets—in particular, it would need to be extended to also independently estimate

⁴⁶The coordinates of these 175 consumer locations are drawn from two normal distributions, each with a mean equal to the sample mean firm coordinate and a variance equal to three times the sample variance in the firm locations to obtain a wider spread.

⁴⁷The collusive base here was (690, 780). The mean transaction price in competition was 6917.49 and in collusion 7053.56.

t_d . The program, BaseLocator[®], needs a structured data file with the individual mill and customer project locations, and for each project location the volume of trade and the price at which it was ordered. For a given structure of transportation costs—so far the linear specification (5.15), but alternative structures are possible as well—the program returns the likelihood of collusion. The software could be used as a screen in the sense that suspiciously high values of LoC could be candidates for targeted investigations. A deeper investigation can be entered, for example, when the null hypothesis $H_0 : LoC = 0$ is rejected at a set confidence level.

5.7 Concluding Remarks

Collusive basing-point pricing is an elusive form of price conspiracy. The system is clever, easy to implement and it leaves few traces. Therefore, collusive basing-point pricing is difficult to tell apart from genuine competition by conventional means. Where a cartel is suspected to have abused basing-point pricing, it may be difficult to prosecute for lack of hard evidence. We propose to test for collusive basing-point pricing by recovering the base(s) used in calculating prices, using order volumes, transaction prices and customer locations. By comparing the implied base location(s) with the locations of the production mills, the test discriminates between competitive and collusive basing-point pricing. The method can be applied as a screen in industries that traditionally use delivered pricing systems. It can help to detect cartels in a relatively simple and inexpensive way, that is both non-invasive for innocent firms and hard to beat for cartels. When a cartel was found to have abused the basing-point system, our method can be used to estimate antitrust damages. A software that operationalizes the proposed method is outlined.

The next step for us would be to apply our method to actual data from industries known or suspected to have used collusive basing-point pricing. Although the early antitrust cases involving basing-point pricing mentioned in the introduction were obviously more complex than our basic model, they appear to be broadly consistent with our analysis. We leave the identification and consistent collecting of suitable data for future research, but we can sketch some directions for analysis.⁴⁸

In the U.S. steel example, Pittsburgh-Plus was first used in 1880, when the dominant producer Carnegie Steel, located in Pittsburgh, commonly priced with several smaller producers in eastern Pennsylvania and New Jersey from Pittsburgh to all destinations.⁴⁹ For Carnegie Steel, which faced no competitors west of it, the common base assured that it could deliver steel profitably in the whole of the United States. Chicago mills, which had a limited capacity, benefitted from phantom freight. The system broke down at around 1920, when producers sprang up more western and some Chicago mills had increased their capacity and started to undercut the system. Our

⁴⁸In order to be able to confront our method with real transaction data, the numerical estimation method described in Appendix A would first need to be extended to determine \hat{t}_d as well. This slightly increases data and computational requirements.

⁴⁹See Commons (1924), Marengo (1955) and Greenhut (1987).

test would in principle be able to identify Pittsburgh as a common single base point and be useful in estimating realized cartel overcharges.

In *Cement Institute*, collusion was facilitated first by a single basing-point system. The cartel later changed to a complex multiple basing-point system after the “Cement Institute” was established in 1929 to publish and distribute ‘freight rate books’ from which cartel members were to calculate their bids.⁵⁰ Collusive bases would vary over time and often coincide with production locations. In 1937, for example, of the 165 mills in the cartel, 79 acted as a basing point, against 8 additional distant basing points.⁵¹ This may have been related to the fact that cement could not be shipped over such long distances as steel could, so that the Institute in fact administered a number of locally isolated cartel clusters. It has also been suggested that the complexity of using multiple basing points helped to conceal the collusive nature of the system.⁵² Our method could be employed to try to identify clusters of plants that used a common base for certain periods of time in local cartel rings.

In *Softwood-Plywood*, there appear to have indeed been two locally isolated clusters. Prior to 1963, plywood, a building material, could be produced almost exclusively in the coastal area of the Pacific Northwest. The development of lamination techniques in the early 1960’s made production possible in the south-eastern part of the United States. Beginning in 1963, Southern plywood production grew steadily and by the middle part of the 1970’s more than fifty Southern plants accounted for roughly 40 percent of the plywood produced in the U.S. The incumbent northern firms had long used Portland, Oregon as a common base in their midst from which they each charged near competitive delivered prices. Until the FTC issued a cease-and-desist order in 1977, and the industry adopted mill pricing, the southern entrants managed to maintain Portland as a single base. Consumers purchasing southern plywood, wherever located, paid delivered prices based on rail freight charges from Portland, regardless of the actual mode of transportation utilized or the origin of the plywood shipment. The southern cluster profited, as Gilligan (1992) establishes. We could use our method to establish the Portland base and estimate antitrust overcharges.

We offer some ideas for extensions of the analysis in the text. For illustrative purposes, in Section 5.5 we have assumed away any interaction in costs between distance and quantity. However, our test can in principle deal with a variety of transportation technologies, as long as they are continuous.⁵³ Note, however, that it would introduce a specification error if the true structure of transportation costs were to be different from the one assumed in the procedure. This will generally result in the identification of locations that are systematically off base, so that the test may no longer discriminate. It is straightforward to program a choice of transportation cost specifications. The software could also be set up to try a number of non-linear specifications for

⁵⁰See Loescher (1959).

⁵¹See Machlup (1949).

⁵²See Greenhut (1987).

⁵³Alternatively, freight could be organized in integer units of a given volume, say the cargo hull of a truck or airplane, so that the transportation costs of a total volume shipping are given in a multiple of this minimal volume. Note the transportation costs can in principle decrease in distance, as well. This violates Assumption 2, but would still allow base recovery.

the best fit. This could be done on the basis of a traditional LS-criterion, or using a criterion of lowest sample variance.

In order to bring our analysis to data successfully, we would probably need to extend it with differences in productive efficiency and capacity constraints, so as to capture all motives to apply basing-point pricing in some of these cases. Also, we have so far worked under the assumption that there is a known and isolated local cluster of mills involved in the cartel. As said, if outside competitors exist—be it individual firms or separate clusters that each formed their own local cartel ring—their plant locations are natural collusive bases. Obviously, for the above described basic version of our test to work properly, it must be applied to a correctly identified cartel cluster that does not include fringe competitors.⁵⁴ Should competitors mistakenly be included, the test would use too wide a convex hull of firm locations. Moreover, if one of the nearest distant rivals has been the collusive base for the cartel, the test in principle recovers its mill location as a single base. Since this base location is found with a small spread, as well as likely on the boundary of the hull, the software would in principle still return a high value of the *LoC*-measure.

Our basic test has to be applied with caution when there are multiple clusters of firms that each collude independently.⁵⁵ Such distant clusters that are mistakenly taken together as one cartel in the application of our test can lead to a high sample variance and to a low *LoC*-measure. Note, however, that it is really only the straightforward application of the *LoC*-measure which causes this problem. Suspicious multiple basing-points systems could still be spotted if instead the entire pattern of recovered bases would be studied by hand, rather than automatically processed in the proposed screen. Furthermore, a relatively straightforward extension of the algorithm could test recursively for the presence of separate collusive clusters in the transaction data. Starting from all firms, the method could in principle find the best fit of local clusters of firms with a common collusive base. Such an approach would determine the boundaries of (local) cartels with less than all competitors in the relevant market. We leave such collusive cluster analysis for further research.

Caution must also be applied in the interpretation of the *LoC*-values when the structure of the data does not result from the use of a single collusive basing point. If the data cover a long period of time, they may include bids formulated under collusive as well as competitive regimes. If the period covers a punitive phase (or phases), the vector of recovered bases includes the location (s) of firm(s) that were punished.⁵⁶ This will result in a large variance and test surface area, albeit around an average base location that is likely to be away from the plant locations. A similar measurement error would occur if the cartel formed during the recorded time-series. Apart from being quite coincidental, positive values of *LoC* would often still single these markets out for

⁵⁴Conceptually, including competitors in an alleged cartel cluster is comparable to mistakingly using unrelated crime scenes in the geographical profile of a serial offender.

⁵⁵See Levy and Reitzes (1993).

⁵⁶Greenhut (1987) warns for this where he points out that sellers can establish "... difficult to evaluate variants of multiple base point system." This leads the author to conclude against the common belief at the time that: "To assert ... that the Basing Point system is ... on its way out of use in the United States would ... be a stronger inference than one can soberly make." (*op.cit.*, pp.202-3).

closer inspection. In general, our test would better not be applied when the intervals between bids are large.

A related problem arises when cartels have applied different distant collusive bases over the sample period. If more than one collusive base, located at different quarters around the cluster of firms, have generated a sample that is treated as if generated from one base, the sample variance is likely to be erroneously large and the sample mean may be close to the center of the market. A cartel that would understand this effect on the test could deliberately rotate its collusive bases in an attempt to avoid discovery.⁵⁷ In addition, this could help to allocate cartel profits more evenly between the members. Transaction data could in principle be scrutinized for such avoidance patterns, thus forcing the cartel in ever more complex and costly agreements with associated difficulties of cheating and monitoring.

Cartels could also try to avoid a basic version of our test by using a somewhat altered collusive bid formula, in which each cartel member adds the collusive phantom freight based on the distance between the collusive base and the customer project to the phantom freight it would charge in competition, which is based on the customer's distance to the mill's nearest rival. This collusive bid structure is still relatively straightforward. It has as a benefit over the common rule discussed in the text that it retains the differences in quoted prices that would be observed under competition when also all distant firms bid their true transportation costs. The system furthermore creates a natural division of the market in home markets, as all competitive bids are marked up by the same amount. The added variability in the bids would make it somewhat more difficult to trace, yet the software can straightforwardly be extended to test for this specific alternative bid structure. In addition, new plant location choices and entry decisions may endogenize the presence of basing-point screens—which, in turn, may provide grounds for additional tests on location patterns typical for collusion. In any event, for industries that aspire to conspire, the need to try to avoid sophisticated detection methods will at a minimum make it more complex—and thus less attractive—to collude.

⁵⁷In reference to the geographical profiling methods discussed in Section 1, this would be similar to a serial killer that outsmarts the sheriff's software by well-placing the bodies of its victims around somebody else's address.

6

Summary and Conclusions

“The important thing is not to stop questioning.” – Albert Einstein¹

6.1 Introduction

Standard economic theory of collusion typically presumes cartels encompass all firms in the industry. In practice, however, many cartels did not control all industry supply. In this dissertation, we have analyzed such incomplete cartels from an economic theoretical perspective. We have examined the nature of incomplete cartels and explored ways in which economic theory can be used to detect (incomplete) cartels. In this concluding chapter, we summarize the main findings of this thesis and draw some implications for antitrust policy. Additionally, we outline avenues for future research.

6.2 Main Findings

The first main research question that has been addressed in this thesis is: What explains optimal cartel size to be less than maximal? Existing theories of incomplete cartels indicate that the full cartel might not be stable. In many oligopoly models, undertakings that do not take part in a cartel earn more profits than those which form the coalition. This is partly due to the fact that a cartel creates positive externalities for non-conspirators. The price increase and output reduction of cartel members provides an incentive for outsiders to increase production and to charge a higher selling

¹See www.quotationspage.com.

price. Fringe members therefore enjoy the benefit of higher industry prices, but do not incur the cost associated with a reduction in sales. However, no cartel will emerge if every firm expects its rivals to collude. As a result, quite a few theories predict the equilibrium cartel size to be the minimum cartel size that is required to sustain some collusion.

In this dissertation, we have provided an alternative explanation for the existence of incomplete cartels. In a price setting supergame in which firms differ in terms of production capacity it is shown that the full cartel is stable when collusive gains are allocated properly. Consequently, instability of the full cartel is no explanation for the existence of incomplete cartels in this setting. Yet, to establish this particular cartel arrangement typically requires explicit communication between firms, which implies that cartelizing is costly. Arguably, these (expected) costs of colluding are increasing in the number of participants. Under very general conditions, we have shown that the optimal cartel size is not all-inclusive when colluding is costly and the smallest firm in the industry is sufficiently small. Thus, when taking account of the cost of cartelizing, the most profitable anticompetitive agreement can be a cartel with less than one hundred percent market share.

The second major research topic concerns the question: What are the traits of those firms that form the cartel? Descriptive cartel studies suggest that large firms are more inclined to join a cartel. However, existing theories of incomplete cartels remain largely silent about this issue. This is essentially due to the fact that most contributions assume identical firms. In the model presented in this thesis, firms are characterized by their capacity stock, which is taken as a proxy for firm size. We have established a positive correlation between the incentive to collude and firm size. In particular, a sufficiently small seller lacks the incentive to take part in any conspiracy, while a sufficiently large firm always has an incentive to join a cartel. A cartel comprising the largest firms is shown to be a subgame perfect equilibrium of the game, but this equilibrium is not guaranteed to be unique. Additionally, it has been shown that, under certain conditions, firms have an incentive to form the cartel for which total profits are highest. Free-rider incentives might therefore not be as strong as in symmetric oligopoly models. As a consequence, the equilibrium cartel size can be larger than the minimum cartel size required to sustain some collusion.

As to the third main research question, conclusions are less clear-cut. We conjectured that certain industry structures are particularly conducive to the formation of incomplete cartels. The analysis suggests that in industries in which the size distribution of firms is asymmetric and in which one or more sellers are small, optimal cartel size is likely to be less than all-inclusive. However, the possibility of an incomplete cartel in industries with a more or less equal division of production capacity cannot be excluded. What cartel size can be considered optimal will in part depend on the magnitude of the cost of cartelizing, but these are largely unknown. We further explored the impact of changes in the size distribution of firms. Among other things, we have shown by example and by performing simulations that the strongest coordinated effects may come from mergers involving moderate-sized firms. It proved to be difficult to establish more definitive conclusions. A more in-depth analysis on the relationship between industry structure and optimal cartel size is left for future research.

In the second main part of this dissertation, our focus has been primarily on antitrust enforcement. The question that we have addressed is how economic theory can be used to detect (incomplete) cartels. We have made the case that antitrust agencies around the globe should take an increasingly active role in the search for cartels. First and foremost, this is likely to enhance the deterrent effect of antitrust policy. Additionally, it will make authorities less dependent on more passive instruments of enforcement such as the complaint procedure and the leniency program. Arguably, firms will become increasingly sophisticated and it is to be expected that more cartels will hold their meetings in secret and reduce communication and associated documentation to a minimum. A major advantage of economic methods of detection is that the focus is mainly on the consequences of a cartel arrangement. Even though cartel members are likely to find ways to hide their practices, they will have a hard time covering the impact of the cartel contract on the market. Obviously, a cartel might attempt not to attract too much attention, but implementing a cartel agreement while maintaining sufficient competitive appearance is complicated. Thus, a more active detection policy potentially lowers the incentive to join a cartel.

However, the discussion of economic methods currently available to detect cartels reveals that applying these techniques successfully is all but trivial. One potential complication is that detection tests typically require good quality data, which might not be available or difficult (and therefore costly) to obtain. Also, available data sometimes cannot be used due to legal restraints. An economic detection technique is therefore most promising if data are relatively easily accessible. An example of such a method is the ‘variance screen for collusion’ by Abrantes-Metz *et al.* (2006). Basically, this test only requires a sufficient amount of price data. We have argued that the effectiveness of this method in part depends on the (type of) industry to which it is applied. For instance, the variance screen may lead to wrong conclusions when it is applied to industries in which products are only marginally differentiated and in which price-cycles frequently occur. Building on this, we have made the case that detection tests should take account of the idiosyncrasies of the industry under consideration. It has been argued that substantial progress can be made by developing detection methods that are specifically designed for a particular (type of) market.

In this dissertation, we have developed a detection technique that is specifically designed for markets in which firms apply a basing-point system. The basing-point system is often labeled a ‘facilitating practice’ and there have been quite a few (incomplete) cartels that made use of this pricing method. Using a two-dimensional spatial model, we have shown that in competition firms either use their own plant location or the location of their nearest competitor as basing-point. By contrast, under a cartel regime cartel members often will use a basing-point relatively far from their production centres. Hence, basing-point locations potentially reveal something about the likelihood of collusion. Our detection test attempts to trace the basing-points used by firms. As an input, it requires transaction data and customer locations. Basing-points that are far from mill locations and located relatively close together can be marked a tell-tale sign of collusion. Instead, basing-points that are close to firm locations with a larger variance are compatible with competition. To capture this likelihood of collusion, we have introduced a measure that ranges from zero to one. Numbers close to one are indicative of collusion, while numbers close to zero correspond to competition.

In order to deal with large amounts of data as well as with noise in the data, a software has been developed. The application of this method to real-world cases is left for future research.

6.3 Implications for Antitrust Policy

The prime goal of antitrust law enforcement is to prevent firms from infringing rules of competition. In light of antitrust law enforcement, therefore, it is important to assess whether or not a policy measure has this intended effect. To illustrate how a certain policy measure might yield an unintended result, consider the study conducted by Albaek *et al.* (1997). This work examines the effect of a policy adopted by the Danish Competition Council in 1993. In an attempt to make markets more competitive, the Council decided that transaction prices should be made public. The idea behind this policy is to make buyers better informed, which then should result in a downward pressure on prices. The authors analyze the impact of this policy on the Danish concrete market. In this market, prices were made public for some types of concrete, but not for others. It is found that publicly known prices rose by 15-20%, while nontransparent prices only rose by 1-2%. According to the authors, there are no convincing arguments that can explain this difference, but the fact that price transparency fosters collusion. This analysis suggests that a government policy that makes markets more transparent may result in more collusive outcomes. Hence, economic theory potentially has an important role to play in shaping antitrust policy.

The research conducted in this dissertation has several implications for antitrust policy. We have argued that economic theory can advise antitrust authorities in the search for cartels. To determine the impact of cartel detection techniques on the incentive of firms to collude is difficult, but not impossible. For example, Posner (1970) proposes as a start to have a closer look at recidivist. In the beginning of modern antitrust policy, recidivist were not punished more severely than other defendants. Nowadays, they are. This provides an opportunity to assess the deterrent effect of cartel detection methods, albeit a crude estimation. However, to perform such an analysis requires detailed information and the data presented in antitrust cases is often quite limited. In particular, case descriptions often lack information on the market share of the cartel. This is problematic as we have shown that the applicability of a detection technique partly depends on the inclusivity of the cartel. In determining the effectiveness of antitrust law enforcement and in particular cartel detection methods it is important that antitrust authorities provide as much detailed information about a cartel as possible.

In particular, it is important to define market boundaries properly. In theories of cartels, industry size is, more often than not, taken as given. In antitrust practice, however, defining this so-called *relevant market* is all but a trivial exercise; the outcome

of which potentially has a major impact on the (potential) case under consideration.² It is common practice to define a “market” according to its geographic and product dimension. The well-known SSNIP test is a popular tool to assess which undertakings operate on the same market.³ SSNIP stands for ‘Small but Significant Non-transitory Increase in Prices’ and it works as follows. Consider some product x . According to the SSNIP test, there is a separate market for product x when it is profitable to increase its price with a certain percentage.⁴ If, however, following the price increase, many customers opt for product y , then product x and y are considered to be part of the same product market. In a similar fashion, the test may be used to assess how many producers belong to the same market. If a producer v would find a hypothetical price increase not profitable because it loses too many customers to producer z , then both suppliers operate on the same relevant market.

Obviously, the market definition is important to determine whether or not a cartel is likely to be all-inclusive. A cartel can truly be considered less than all-inclusive only when non-conspirators operate in the same relevant market. The evidence that was presented in this dissertation is in large part based on the definition of the relevant market as formulated in the various cases. In principle, therefore, there is room for an argument stating that incomplete cartels are not as common, because antitrust agencies around the world might have failed to define properly the market in which these cartels were operating. More precisely, there is a potential risk that the alleged incomplete cartel might well have been all-inclusive when the “true relevant market” was more narrow than was believed by the authorities.

Indeed, it is not so hard to believe that the relevant market is defined wrongly in quite a few cases. Yet, what is more difficult to believe is that antitrust authorities around the globe on a large and systematic scale defined the relevant market too broadly. On the contrary, it is arguably in their interest to make the case that the cartel under consideration caused substantial harm both for customers and the economy at large. Such claims, for example, may contribute to the willingness of the public to invest in antitrust enforcement, which in turn enhances the power of the antitrust agency. As a result, the bias in market definitions used in antitrust practice, if any, would be that the relevant markets are often defined too narrow. In other words, the known number of incomplete cartels in all likelihood forms a lower bound, not an upper bound.

Defining the relevant market properly is important to assess whether or not a cartel is likely to be all-inclusive. This, in turn, is important to determine if a competitive fringe is present and can be used as a benchmark, i.e., whether or not detection techniques that assume incomplete cartels are likely to be successful. The research conducted in this thesis further suggests that, when an incomplete cartel forms, small firms gain market share at the cost of larger undertakings that take part in the cartel. Hence, a decline in total market demand in combination with a more symmetric

²In some cases, however, the relevant market is relatively easy to define. One may think, for example, of bidding rings. It is often not too hard to find out what firms submit bids for well-defined contracts in a certain region.

³This test is also known under the header ‘Hypothetical Monopolist test’.

⁴In the U.S. this is typically 5%, while in the European Union it is commonly in the range 5-10%.

division of market shares might indicate collusion. Clearly, such patterns cannot be observed when the relevant market is defined too narrowly.

We have argued that a more active detection policy potentially increases the probability of discovery and therefore raises the expected costs of collusion. In its simplest form, the expected cost of cartelizing equals the monetary penalty times the probability of discovery. Arguably, therefore, an easier route to obtain the same effect is to increase the level of antitrust fines. Yet, it is questionable if fines can be raised sufficiently within the current legal framework. The new European fining policy is illustrative in this respect. In 2006, the European Commission released new fining guidelines ‘Guidelines on the method of setting fines imposed pursuant to Article 23(2)(a) of Regulation No 1/2003’.⁵ The new guidelines are meant to determine the level of antitrust fines in a more systematic way. In imposing the fine, the Commission must respect the legal maximum fine as specified in subparagraphs 2 and 3 of Article 23(2) of Regulation 1/2003. The fine cannot exceed 10% of the undertaking’s total turnover in the preceding business year. So far, this legal maximum has been binding in only a limited number of cases. However, as we have shown in Bos and Schinkel (2006), it is to be expected that the legal maximum will be far more restrictive with the new fining guidelines. That is, applying the method laid out in the new guidelines yields a fine that is higher than 10% of worldwide turnover in a wide array of situations. Consequently, even though the Commission announced deterrence as the main objective, it is questionable if this goal can be achieved by an increase in punishments alone.⁶

6.4 Future Research

In this concluding section, we discuss some research topics that in our opinion must have a prominent position on the future research agenda. In several chapters, we have already touched upon some issues for which additional research is warranted. The purpose of this section is not to provide a complete overview of these research topics. Instead, we limit ourselves to a brief discussion of two themes that, we believe, are of particular importance.

6.4.1 *Cartel Formation with Heterogeneous Firms*

In this dissertation, we have addressed the issue of coalition formation with positive externalities. Like most theories of collusion, the strand of literature concerned with cartel formation assumes identical firms. The outcome of the game under consideration is therefore mainly limited to a prediction of optimal cartel size. To say anything more, one necessarily has to give players an identity, which implies they have to differ

⁵The 2006 fining guidelines replace the ‘Guidelines on the method of setting fines imposed pursuant to Article 15(2)(a) of Regulation No 17 and Article 65(5) of the ECSC Treaty’ of 1998.

⁶The objective to prevent infringements of the European competition law is mentioned explicitly in recitals 4, 30 and 31 of the 2006 fining guidelines.

in at least one respect. As mentioned before, however, formation games with firm heterogeneity tend to be complex and easily render intractable. Nevertheless, the analysis carried out in Chapter 3 shows that potentially some progress can be made in this field.

The main problem with cartel formation games with firm heterogeneity that are analytically tractable is that it leaves a large class of equilibrium outcomes. As we have seen in Chapter 3, sufficiently large firms always join the cartel independent of the cartel formation game, while sufficiently small firms always lack the incentive to collude. Yet, moderate-sized firms may or may not join, which depends heavily on the cartel formation game. As a result, it remains difficult to predict what stable cartel is most likely to emerge. Nevertheless, we can imagine firms to have a preference ranking for the various stable coalitions.

To provide some preliminary intuition, consider the price setting supergame with capacity constraints developed in Chapter 3. Suppose there is one firm that is sufficiently large so that it has an incentive to join any coalition. Clearly, this firm has a preference for the stable cartel that controls the largest amount of industry capacity, because this is the most profitable cartel. It might be in the interest of this firm to take the lead in the cartel formation process in attempt to make the ‘right’ firms joining the coalition. For example, it can approach its rivals in a certain order so that the largest firms have an incentive to join the cartel, leaving the smallest competitors as outsiders. A better insight of what (type of) firms initiate a cartel arrangement is not only interesting in its own right, but may also yield important insights in light of antitrust enforcement. For instance, ringleaders typically lack the right to apply for leniency. An analysis of such a game has not been pursued in this thesis, but is potentially a very interesting extension of the analysis in Chapter 3.

6.4.2 *Disentangling Overt Collusion and Tacit Collusion*

An important theme, which we have touched upon in several chapters, is the distinction between explicit collusion and tacit collusion. So far, economic theory has mainly focused on the difference between (imperfect) competition and collusion. From an enforcement perspective, however, the key issue is to delineate overt collusion from other sorts of market conduct. The main difficulty is that, according to economic theory, any collusive behavior can be explained as being the result of a tacit coordination of actions. To put it differently, any collusive outcome that can be achieved with talking can also be achieved without talking. As a result, economic theories of collusion have not much to say about the impact of explicit communication between firms on the market outcome. In practice, there arguably is a substantial difference between explicit collusion and tacit collusion. After all, why would firms make the effort to gather together and incur associated costs if there is no additional benefit?

An illuminating discussion on this issue can be found in Whinston (2006). He hypothesizes that the likelihood of overt collusion increases in the additional gains due to talking. In order to clarify the discussion, we can write the net benefits of explicit collusion (EC) as,

$$EC = \pi(\text{talk}) - \pi(\text{do not talk}) - T.$$

In every market, firms can obtain a particular market result without explicit communication, which is denoted $\pi(\text{do not talk})$. Explicit communication between firms presumably leads to a better coordination of strategies and, as a result, yields additional profits equal to $\pi(\text{talk}) - \pi(\text{do not talk})$. Simultaneously, talking makes collusion costly.⁷ These (expected) costs are given by T . Clearly, under the assumption that firms are risk-neutral, ‘talking’ is preferred to ‘not talking’ whenever EC is positive. A major challenge for economic research is then to identify factors that positively affect EC .

We can distinguish three categories of factors. First, there are factors that influence the potential size of $\pi(\text{talk})$. When expected gains from collusion are small, then there is not much reason to talk. For instance, in industries characterized by many sellers and low barriers to entry the expected profits from collusion are likely to be negligible. Second, one can think of factors that have an effect on T . Examples include the probability of discovery and the level of punishment. Economists have a relatively good understanding of these first two categories. The third category is more problematic. These are factors that affect $\pi(\text{talk})$ more than $\pi(\text{do not talk})$, i.e., factors that determine the additional gain resulting from explicit communication. Here, the main problem is that factors that positively affect $\pi(\text{talk})$ typically also have a positive effect on $\pi(\text{do not talk})$, which makes the outcome often ambiguous. The key challenge is then to identify situations in which direct communication between firms has substantial added value.

Athey and Bagwell (2001) explains that explicit communication could be important in settings in which firms are heterogeneous and obtain private and imperfect signals of past play. In particular, private information typically yields more volatile market shares and in order to maintain cartel stability firms with a low market share today should be compensated in the future. This potentially provides an opportunity to discriminate between tacit collusion, overt collusion and imperfect competition. Alternating sizes of market shares may indicate overt collusion, stable market shares may be the result of tacit collusion, while less stable but not alternating market shares may indicate imperfect competition.

In principle, this problem could be circumvented by declaring illegal explicit communication between businessmen. Yet, business meetings can yield substantial benefits for society and preventing firms from gathering together is unlikely to be optimal from a social welfare perspective. Allowing firms to meet occasionally then comes with a price, which is the risk of collusion. Indeed, Adam Smith already recognized the risk associated with business meetings. He wrote:⁸

“People of the same trade seldom meet together, even for merriment and diversion, but the conversation ends in a conspiracy against the public, or in some contrivance to raise prices.”

⁷Note that we assume ‘false positives’ to be absent. If antitrust authorities would make a significant amount of so-called ‘Type 1 errors’ firms might have an additional incentive to collude explicitly. See Schinkel and Tuinstra (2006) for a discussion and formal analysis.

⁸Smith (1994), p. 148.

However, he also realized that prohibiting explicit communication would not be consistent with a free enterprise system. The next line reads,

“It is impossible indeed to prevent such meetings, by any law which either could be executed, or would be consistent with liberty and justice.”

That said, the main challenge for future research is to identify factors that affect $\pi(\text{talk})$ relatively more than $\pi(\text{do not talk})$, all else unchanged.

Appendix A

BaseLocator[®]

We have operationalized the cartel screen that has been developed in Chapter 5 in a simple topographic detection routine in Delphi 5[®].¹ Here we illustrate the logic of the algorithmic steps taken in the software. Below we give the main program with the key subroutines *TracingBP*, which traces the base using function *SumOfSquares*, and *EstimateLstatistic*, which calculates the value of *LoC*.²

A.1 Steps to Trace the Base

Input is a structured data file (in Notepad) that has a column of individual mill locations, and a column of individual customer project locations with the volume of trade and the transaction price per project. Base tracing consists of three main steps: data sorting, base tracing, and calculation of the *LoC*-measure.

In the first step, the data are sorted. All transaction data (P_i, q_i) are grouped by base group, using the information on project site locations. Those combinations of project locations that are aligned are disregarded as not independent—but this is a rare occasion. All base groups with less independent observations than the number of unknowns, which is four in the system developed in the text, are ignored—this small sample problem would normally not need to appear. What remains is N sets of independent observations, $N \leq J$.

¹Excellent programming assistance has been provided by Eelko Ubels.

²The other subroutines called in the program are less insightful and lengthy. They are available upon request from the author.

In the second step, each constructed set of observations $l = 1, \dots, N$ is used to recover \tilde{c} , \tilde{F} and base location used for the base group considered, (\hat{a}_l, \hat{b}_l) . For this, the specification of $T(q_i, d_{li})$ is crucial. The software would in principle allow for a variety of specifications of $T(q_i, d_{li})$ —for the user to choose from, or for the program to find the best fit amongst. In the present version, transportation costs are linear in distance and volume, and including a fixed component, as in equation (5.15) in the text.

In a bounded area—that we determine as the size of the convex hull area of customer projects locations, extended with twice that size in all directions—the program searches in a grid for the specification of $(\tilde{c}, \tilde{F}, (\hat{a}_l, \hat{b}_l))$ that returns the lowest S value. To be computationally efficient, we first use the information of the partial first-order conditions to problem (5.19):

$$\begin{aligned}\frac{\partial S}{\partial \tilde{c}} &= -2 \sum_{i=1}^{I_{G_v}} q_i \left(P_i - \tilde{c}q_i - \tilde{F} - \sqrt{(a_l - a_{x_i})^2 + (b_l - b_{x_i})^2} \right) = 0, \text{ and} \\ \frac{\partial S}{\partial \tilde{F}} &= -2 \sum_{i=1}^{I_{G_v}} \left(P_i - \tilde{c}q_i - \tilde{F} - \sqrt{(a_l - a_{x_i})^2 + (b_l - b_{x_i})^2} \right) = 0,\end{aligned}$$

to obtain

$$\begin{aligned}\tilde{c} &= \frac{\sum_{i=1}^{I_{G_v}} q_i (P_i - \tilde{F} - d_{li})}{\sum_{i=1}^{I_{G_v}} q_i^2}, \text{ and} \\ \tilde{F} &= \frac{\sum_{i=1}^{I_{G_v}} (P_i - \tilde{c}q_i - d_{li})}{I_{G_v}}.\end{aligned}$$

where $d_{li} = \sqrt{(a_l - a_{x_i})^2 + (b_l - b_{x_i})^2}$.

Using the averages $\bar{P} = \frac{\sum_{i=1}^{I_{G_v}} P_i}{I_{G_v}}$, $\bar{q} = \frac{\sum_{i=1}^{I_{G_v}} q_i}{I_{G_v}}$ and $\bar{d}_l = \frac{\sum_{i=1}^{I_{G_v}} d_{li}}{I_{G_v}}$, some manipulation yields

$$\begin{aligned}\tilde{c} &= \frac{\frac{1}{I_{G_v}} \sum_{i=1}^{I_{G_v}} q_i (P_i - d_i) - \bar{q} (\bar{P} - \bar{d}_l)}{\frac{1}{I_{G_v}} \sum_{i=1}^{I_{G_v}} q_i^2 - \bar{q}^2}, \text{ and} \\ \tilde{F} &= \bar{P} - \left(\frac{\frac{1}{I_{G_v}} \sum_{i=1}^{I_{G_v}} q_i (P_i - d_i) - \bar{q} (\bar{P} - \bar{d}_l)}{\frac{1}{I_{G_v}} \sum_{i=1}^{I_{G_v}} q_i^2 - \bar{q}^2} \right) \bar{q} - \bar{d}_l.\end{aligned}$$

Plugging these expressions for \tilde{c} and \tilde{F} into the criterion function (5.19), we obtain

$$S = \sum_{k=1}^{I_{G_v}} \left(P_k - \bar{P} + \bar{d}_l - d_{lk} + [\bar{q} - q_k] \left(\frac{\frac{1}{I_{G_v}} \sum_{i=1}^{I_{G_v}} q_i (P_i - d_i) - \bar{q} (\bar{P} - \bar{d}_l)}{\frac{1}{I_{G_v}} \sum_{i=1}^{I_{G_v}} q_i^2 - \bar{q}^2} \right) \right)^2,$$

for which we are to find the value(s) for (\hat{a}_l, \hat{b}_l) that return(s) the lowest S -value.

In the search area, the value of S is determined for each combination of (a, b) . The program stores the S -value and overwrites it when further grid-point yields a lower value. This is our candidate basing point.

In the third step, the base locations found are translated into the LoC -measure. The program determines the convex hull of firm locations and its surface. It determines λ on the basis of the theoretical competitive mean base point and variance, using project and mill locations only. Since the data are sorted per base group, in competition each mill would be found only once. The theoretical competitive mean base point therefore is the unweighted mean of the mill locations. The program subsequently calculates the mean recovered base location and the ‘distance spread-circle’ around it. This returns the surface $S^{\mathcal{L}}$. The intersection of these two areas gives the LoC -measure, a number between zero and one.

As output, the program returns the name of data set used, the location l^* , which is referred to as the center of the convex hull for reference, the value of λ , the sample mean base, the sample mean variance, the parameters of the bid structure estimated (normalized on $t_d = 1$), and the value of the LoC -measure. High values of LoC are indicative of collusion, in particular when supported by a small sample variance.

A.2 Kernel of the Software

```
function SumOfSquares(const p,q:Vector; const x:Matrix; const ag,bg:integer):Vector;

    {uses criterion function to calculate sum of squares, marginal cost and fixed
cost
for given basepoint candidate}
var s_i,s_j,s_jt,s_jn,pm,qm,dm,h:extended;
    i,j,len:integer;
    d,bp,s:Vector;
begin
    len:=Length(p);
    s_i:=0;
    pm:=Mean(p);
    qm:=Mean(q);
    SetLength(d,len);
    SetLength(bp,2);
    bp[0]:=ag; bp[1]:=bg;
    for i:=0 to len-1 do begin
        h:=0;
        for j:=0 to 1 do
            h:=h+Sqr(bp[j]-x[i][j]);
        d[i]:=Sqrt(h);
    end;
    dm:=Mean(d);
```

```

s_jt:=0; s_jn:=0;
for j:=0 to len-1 do begin
s_jt:=s_jt+q[j]*(p[j]-d[j])-qm*(pm-dm);
s_jn:=s_jn+Sqr(q[j])-Sqr(qm);
end;
s_j:=s_jt/s_jn;
for i:=0 to len-1 do begin
s_i:=s_i+Sqr(p[i]-pm+dm-d[i]+(qm-q[i])*s_j)
end;
SetLength(s,3);
s[0]:=s_i;
s[1]:=s_j;
s[2]:=pm-(s[1]*qm)-dm;
SumOfSquares:=s;
end; {SumOfSquares}

procedure TracingBP(const base:integer; const x:Matrix; const p,q:Vector; var
BP:Matrix;
const i0,j1,i1,j0:integer);
{determines location with lowest value of sum of squares}
var step,len_BP,intm,a,b,a0,a1,b0,b1:integer;
sum,c,fc:extended;
sos:Vector;

begin {TracingBP}
step:=50;
len_BP:=Length(BP);
SetLength(BP,len_BP+1);
SetLength(BP[len_BP],4);
intm:=Max(j1-j0,i1-i0);
a0:=i0-0*intm;
a1:=i1+0*intm;
b0:=j0-0*intm;
b1:=j1+0*intm;
sum:=SumOfSquares(p,q,x,a0,b0-1)[0];
c:=SumOfSquares(p,q,x,a0,b0-1)[1];
fc:=SumOfSquares(p,q,x,a0,b0-1)[2];
SetLength(sos,3);
for a:=a0 to a1 do begin
if a mod step=0 then begin
for b:=b0 to b1 do begin
if b mod step=0 then begin
sos:=SumOfSquares(p,q,x,a,b);
if sos[0]<sum then begin
sum:=sos[0];

```

```

BP[len_BP][0]:=a;
BP[len_BP][1]:=b;
c:=sos[1];
fc:=sos[2];
BP[len_BP][2]:=c;
BP[len_BP][3]:=fc;
end;
end;
end;
end;
end;

end; {TracingBP}

procedure EstimateLStatistic(const h:Matrix; const sq,v:extended; const m:Vector;
out Ls:extended);
{determines overlap of circle in hull}
var i,j,i0,i1,j1,j0,ot,tot:integer;
    p:Vector;
begin
  SetLength(p,2);
  MinimalRectangle(h,i0,i1,j1,j0);
  ot:=0;
  tot:=0;
  for i:=0 to 100 do begin
    for j:=0 to 100 do begin
      p[0]:=i0+i*(i1-i0)/100;
      p[1]:=j0+j*(j1-j0)/100;
      if PointInHull(h,p,sq) then begin
        tot:=tot+1;
        if PointInCircle(m,p,r) then begin
          ot:=ot+1;
        end;
      end;
    end;
  end;
  Ls:=1-ot/tot;
end; {EstimateLStatistic}

begin {main}
  LoadFileName(s);
  s_out:='LoCw1.txt';
  AssignFile(f,s_out);
  Rewrite(f);

```

```

SetLength(res,2);
for i:=0 to 1 do begin
SetLength(res[i],201);
end;
for j:=0 to 200 do begin //read 201 files with standard errors on price
for i:=0 to 1 do begin
SetLength(res[i][j],10);
end;
for k:=0 to 9 do begin //averageing over 10 files with same error
SetLength(arr_BP,0);
WriteLn(k,Chr(9),j);
s1:=s+'-'+IntToStr(k)+'-'+IntToStr(j)+'-comp_n.txt';

LoadData(s1,BP,sd,int_c,int_F,int_I,int_J,arr_x,arr_y,arr_p,arr_q,all_x,all_p,all_q);

thMean:=MeanVector(arr_y);
ConvexHull(arr_y,arr_h,sqHull);
len_x:=Length(arr_x);
MinimalRectangle(all_x,west,east,north,south);
for i:=0 to len_x-1 do begin
if Length(arr_x[i])>3 then begin

TracingBP(i,arr_x[i],arr_p[i],arr_q[i],arr_BP,west,east,north,south);
end;
end;
thSE:=SigmaBar(arr_y);
ConvexHull(arr_y,arr_h,sqHull);
lambda:=Labda(Radius(thMean,arr_h),thSE);
r:=lambda*SigmaBar(arr_BP);
SetLength(mC,2);
mC:=MeanVector(arr_BP);
EstimateLStatistic(arr_h,sqHull,r,mC,LStat);
res[0][j][k]:=LStat;

SetLength(arr_BP,0);
s1:=s+'-'+IntToStr(k)+'-'+IntToStr(j)+'-col_n.txt';

LoadData(s1,BP,sd,int_c,int_F,int_I,int_J,arr_x,arr_y,arr_p,arr_q,all_x,all_p,all_q);

len_x:=Length(arr_x);
MinimalRectangle(all_x,west,east,north,south);
for i:=0 to len_x-1 do begin
if Length(arr_x[i])>3 then begin

TracingBP(i,arr_x[i],arr_p[i],arr_q[i],arr_BP,west,east,north,south);

```



```

end;
end;
r:=lambda*SigmaBar(arr_BP);
SetLength(mC,2);
mC:=MeanVector(arr_BP);
EstimateLStatistic(arr_h,sqHull,r,mC,LStat);
res[1][j][k]:=LStat;
end;Writeln(f,Mean(res[0][j]):4:2,Chr(9),Mean(res[1][j]):4:2,Chr(9),sd:4:0);

end;
Writeln('the end');

GetProfits(all_x,arr_y,BP,int_I,int_J,all_p,all_q,int_c,int_F,pi_col,pi_comp);

for j:=0 to int_J-1 do begin
Writeln(f,'firm ',j+1,Chr(9),pi_comp[j]:4:2,Chr(9),Chr(9),Chr(9),pi_col[j]:4:2);

end;
CloseFile(f);
Readln;
end. {main}

```


References

- [1] Abrantes-Metz, R.M., L.M. Froeb, J.F. Geweke and C.T. Taylor (2006), "A Variance Screen for Collusion," *International Journal of Industrial Organization*, 24(3), 467-486;
- [2] Abreu, D. (1986), "Extremal Equilibria of Oligopolistic Supergames," *Journal of Economic Theory*, 39(1), 191-225;
- [3] Abreu, D. (1988), "On the Theory of Infinitely Repeated Games with Discounting," *Econometrica*, 56, 383-396;
- [4] Albaek, S., P. Mollgaard and P. Overgaard (1997), "Government Assisted Oligopoly Coordination? A Concrete Case," *Journal of Industrial Economics*, 45, 429-443;
- [5] Alexander, B. (1997), "Failed Cooperation in Heterogeneous Industries under the National Recovery Administration," *Journal of Economic History*, 57(2), 322-344;
- [6] Asch, P. and J.J. Seneca (1975), "Characteristics of Collusive Firms," *Journal of Industrial Economics*, 23(3), 223-237;
- [7] Athey, S. and K. Bagwell (2001), "Optimal Collusion with Private Information," *RAND Journal of Economics*, 32(3), 428-465;
- [8] Aubert, C., P. Rey and W.E. Kovacic (2006), "The Impact of Leniency and Whistle-blowing Programs on Cartels," *International Journal of Industrial Organization*, 24(6), 1241-1266;

- [9] Aumann, R. and J. Drèze (1974), "Cooperative Games with Coalition Structures," *International Journal of Game Theory*, 3, 217-237;
- [10] Bain, J.S. (1956), "*Barriers to New Competition*," Cambridge, MA: Harvard University Press;
- [11] Bajari, P. (2001), "Comparing Competition and Collusion: A Numerical Approach," *Economic Theory*, 18, 187-205;
- [12] Bajari, P. and G. Summers (2002), "Detecting Collusion in Procurement Auctions," *Antitrust Law Journal*, 70, 143-170;
- [13] Bajari, P. and L. Ye (2003), "Deciding between Competition and Collusion," *Review of Economics and Statistics*, 85, 971-989;
- [14] Baker, J.B. (2003), "The Case for Antitrust Enforcement," *Journal of Economic Perspectives*, 17(4), 27-50;
- [15] Baker, J.B. and T.F. Bresnahan (2007), "Economic Evidence in Antitrust: Defining Markets and Measuring Market Power," in: *Handbook in Antitrust Economics*, Buccirossi, P., ed., MIT Press, Cambridge, MA;
- [16] Baldwin, L.H., R.C. Marshall and J-F. Richard (1997), "Bidder Collusion at Forest Service Timber Auctions," *Journal of Political Economy*, 105, 657-699;
- [17] Baumol, W.J., J.C., Panzar and R.D. Willig (1982), "*Contestable Markets and the Theory of Industry Structure*," New York: Harcourt Brace Jovanovich;
- [18] Belleflamme, P. (2000), "Coalition Structures with Open Membership and Asymmetric Firms," *Games and Economic Behavior*, 30, 1-21;
- [19] Benoit, J.P. and V. Krishna (1987), "Dynamic Duopoly: Prices and Quantities," *Review of Economic Studies*, 54, 23-36;
- [20] Benson, L., M.L. Greenhut and G. Norman (1990), "On the Basing-Point System," *American Economic Review*, 80, 584-588;
- [21] Bernheim, B.D. and M.D. Whinston (1990), "Multimarket Contact and Collusive Behavior," *RAND Journal of Economics*, 21, 1-26;
- [22] Blair, R.D. and R.E. Romano (1989), "Proof of Nonparticipation in a Price Fixing Conspiracy," *Review of Industrial Organization*, 4, 101-117;
- [23] Bloch, F. (1995), "Endogenous Structures of Association in Oligopolies," *RAND Journal of Economics*, 26(3), 537-556;
- [24] Bloch, F. (1996), "Sequential Formation of Coalitions in Games with Externalities and Fixed Payoff Division," *Games and Economic Behavior*, 14, 90-123;
- [25] Bloch, F. (2002), "Coalitions and Networks in Industrial Organization," *Manchester Studies*, 70(1), 36-55;

- [26] Bloch, F. (2005), "Group and Network Formation in Industrial Organization: A Survey," in: *Group Formation in Economics: Networks, Clubs and Coalitions*, Demange, G. and M. Wooders (eds.), Cambridge University Press;
- [27] Bolotova, Y. (2006), "Cartel Overcharges: an Empirical Analysis," *Working paper*, University of Idaho;
- [28] Bolotova, Y., J.M. Connor and D. Miller (2008), "The Impact of Collusion on Price Behavior: Empirical Results from Two Recent Cases," *International Journal of Industrial Organization*, 26(6), 1290-1307;
- [29] Borenstein, S. and A. Shepard (1996), "Dynamic Pricing in Retail Gasoline Markets," *RAND Journal of Economics*, 27(3), 429-451;
- [30] Bos, I. (2007), "Can we catch a Thief with an Economist?," *AENORM*, 57, 61-65;
- [31] Bos, I. and J.E. Harrington (2008), "Endogeneous Cartel Formation with Heterogeneous Firms," Economics Working Paper Archive, 544, Johns Hopkins University;
- [32] Bos, I. and M.P. Schinkel (2006), "On the Scope for the European Commission's 2006 Fining Guidelines under the Legal Maximum Fine," *Journal of Competition Law and Economics*, 2(4), 673-682;
- [33] Bos, I. and M.P. Schinkel (2008), "Tracing the Base: A Topographic Test for Collusive Basing-Point Pricing," ACLE working paper, No. 2008-07;
- [34] Bresnahan, T.F. (1987), "Competition and Collusion in the American Automobile Industry: The 1955 Price War," *Journal of Industrial Economics*, 35(4), 457-482;
- [35] Brock, W.A. and J.A. Scheinkman (1985), "Price Setting Supergames with Capacity Constraints," *Review of Economic Studies*, 52(3), 371-382;
- [36] Brusco, S. and G. Lopomo (2002), "Collusion via Signaling in Simultaneous Ascending Bid Auctions with Heterogeneous Objects, with and without Complementarities," *Review of Economic Studies*, 69(2), 407-463;
- [37] Cabou, C.G., D.D. Haddock and M.H. Thorne (1992), "An Analytic History of Delivered Price Litigation: Do Courts Properly Distinguish Rivalrous from Collusive Instances," *Economic Inquiry*, 30, 307-321;
- [38] Canter, D. (2003), *Mapping Murder: The Secrets of Geographical Profiling*, Virgin Publishing;
- [39] Carlton, D.W. (1983), "A Reexamination of Delivered Pricing," *Journal of Law and Economics*, 26, 51-70;
- [40] Castanias, R. and H. Johnson (1993), "Gas Wars: Retail Gasoline Price Fluctuations," *Review of Economics and Statistics*, 75, 171-174;

- [41] Chamberlin, E.H. (1933), "*The Theory of Monopolistic Competition*," Cambridge, MA: Harvard University Press;
- [42] Chang, M-H. (1991), "The Effects of Product Differentiation on Collusive Pricing," *International Journal of Industrial Organization*, 9(3), 453-469;
- [43] Christie, W.G. and P.H. Schultz (1994), "Why do NASDAQ Market Makers Avoid Odd-Eighth Quotes?," *Journal of Finance*, 49(5), 1813-1840;
- [44] Christie, W.G., J.H. Harris and P.H. Schultz (1994), "Why Did NASDAQ Market Makers Stop Avoiding Odd-Eighth Quotes?," *Journal of Finance*, 49(5), 1841-1860;
- [45] Clarke, I., O. Sandberg, B. Wiley and T.W. Hong (2001), "Freenet: A Distributed Anonymous Information Storage and Retrieval System," in: *Designing Privacy Enhancing Technologies*, Heidelberg: Springer Berlin, 46-66;
- [46] Clark, J.M. (1943), "Imperfect Competition Theory and Basing-Point Problems," *American Economic Review*, 33(2), 283-300;
- [47] Clark, J.M. (1949), "The Law and Economics of Basing Points: Appraisal and Proposals," *American Economic Review*, 39(2), 430-447;
- [48] Clay, K. and W. Troesken (2002), "Strategic Behavior and Market Structure in Whiskey Distilling; 1887-1895," *Journal of Economic History*, 62(4), 999-1023;
- [49] Coase, R.H. (1991), "The Institutional Structure of Production, Nobel Memorial Lecture," in: Coase, R.H., *Essays on Economics and Economists*, University of Chicago Press, Chicago;
- [50] Commons, J.R. (1924), "The Delivered Price Practice in the Steel Market," *American Economic Review*, 14, 505-519;
- [51] Compte, O. (1998), "Communication in Repeated Games with Imperfect Private Monitoring," *Econometrica*, 66(3), 597-626;
- [52] Compte, O., F. Jenny and P. Rey (2002), "Capacity constraints, mergers and collusion," *European Economic Review*, 46(1), 1-29;
- [53] Connor, J.M. (2001), "*Global Price Fixing: Our Customers are the Enemy*," Boston: Kluwer Academic Publishers;
- [54] Connor, J.M. (2003), "Private International Cartels: Effectiveness, Welfare, and Anticartel Enforcement," *Working paper 03-12*, Department of Agricultural Economics, Purdue University;
- [55] Cyert, R., P. Kumar and J. Williams (1995), "Impact of Organizational Structure on Oligopolistic Pricing," *Journal of Economic Behavior and Organization*, 26, 1-15;

- [56] Cramton, P.C., and T.R. Palfrey (1990), "Cartel Enforcement with Uncertainty about Costs," *International Economic Review*, 31(1), 17-47;
- [57] D'Aspremont, C., A. Jacquemin, J.J. Gabszewicz and J.A. Weymark (1983), "On the Stability of Collusive Price Leadership," *Canadian Journal of Economics*, 16(1), 17-25;
- [58] Davidson, C. and R. Deneckere (1990), "Excess Capacity and Collusion," *International Economic Review*, 31, 521-541;
- [59] Deneckere, R.J. and C. Davidson (1985), "Incentives to form Coalitions with Bertrand Competition," *RAND Journal of Economics*, 16(4), 473-486;
- [60] Deneckere, R.J. and D. Kovenock (1992), "Price Leadership," *Review of Economic Studies*, 59(1), 143-162;
- [61] DeRoos, N. (2001), "Collusion with a Competitive Fringe: An Application to Vitamine C," Manuscript, Yale University;
- [62] Diamantoudi, E. (2005), "Stable Cartels Revisited," *Economic Theory*, 26, 907-921;
- [63] Dick, A.R. (2004), "If Cartels Were Legal, When Would Firms Fix Prices?," in: *How Cartels Endure and How They Fail*, Grossman, P.Z. (ed.), Edward Elgar;
- [64] Donsimoni, M.-P. (1985), "Stable Heterogeneous Cartels," *International Journal of Industrial Organization*, 3(4), 451-467;
- [65] Donsimoni, M.-P., N.S. Economides and H.M. Polemarchakis (1986), "Stable Cartels," *International Economic Review*, 27(2), 317-327;
- [66] Eaton, C. and M. Eswaran (1998), "Endogenous Cartel Formation," *Australian Economic Papers*, 37(1), 1-13;
- [67] Eckbo, P.L. (1976), *The Future of World Oil*, Cambridge, MA: Ballinger;
- [68] Eckert, A. (2003), "Retail Price Cycles and the Presence of Small Firms," *International Journal of Industrial Organization*, 21, 151-170;
- [69] Ellison, G. (1994), "Theories of Cartel Stability and the Joint Executive Committee," *RAND Journal of Economics*, 25(1), 37-57;
- [70] ESCRIHUELA-VILLAR, M. (2003), *Partial Cartels*, Doctoral Thesis, University of Alicante;
- [71] ESCRIHUELA-VILLAR, M. (2004), "Cartel Sustainability and Cartel Stability," Working Paper;
- [72] ESCRIHUELA-VILLAR, M. (2008a), "On endogenous cartel size under tacit collusion," *working paper*;

- [73] Escriuhela-Villar, M. (2008b), "Partial coordination and mergers among quantity-setting firms," *International Journal of Industrial Organization*, 26(3), 803-810;
- [74] Espinosa, M.P. (1992), "Delivered Pricing, FOB Pricing, and Collusion in Spatial Markets," *RAND Journal of Economics*, 23(1), 64-85;
- [75] Farrell, J. and C. Shapiro (1990), "Horizontal Mergers: an Equilibrium Analysis," *American Economic Review*, 80, 107-126;
- [76] Farrell, J. and M. Rabin (1996), "Cheap Talk," *Journal of Economic Perspectives*, 10(3), 103-118;
- [77] Federal Trade Commission, "Findings, Opinions and Orders in the Matter of Boise Cascade Corporation, et al.," *Federal Trade Commission Decisions*, January 1978, 91, 1-111;
- [78] Feinstein, J.S., M.K. Block and F.C. Nold (1985), "Asymmetric Information and Collusive Behavior in Auction Markets," *American Economic Review*, 75(3), 441-460;
- [79] Feldenkirchen, W. (1992), "Concentration in German Industry, 1870-1939," in: *The Rise of Big Business, International Library of Critical Writings in Business History*, Supple, B.E., Edward Elgar;
- [80] Fershtman, C. and A. Pakes (2000), "A Dynamic Game with Collusion and Price Wars," *RAND Journal of Economics*, 31(2), 207-236;
- [81] Fraas, A.G. and D. Greer (1977), "Market Structure and Price Collusion: An Empirical Analysis," *Journal of Industrial Economics*, 26(1), 21-44;
- [82] Friederiszick, H.W. and F.P. Maier-Rigaud (2007), "The Role of Economics in Cartel Detection in Europe," in: *The More Economic Approach in European Competition Law*, Schmidtchen, D., M. Albert and S. Voigt (eds.), Tübingen: Mohr Siebeck;
- [83] Friederiszick, H.W. and F.P. Maier-Rigaud (2008), "Triggering Inspections *Ex Officio*: Moving Beyond a Passive EU Cartel Policy," *Journal of Competition Law and Economics*, 4(1), 89-113;
- [84] Friedman, J.W. (1971), "A Non-cooperative Equilibrium for Supergames," *Review of Economic Studies*, 38(1), 1-12;
- [85] Froeb, L.M. (1988), "Auctions and Antitrust," *Economic Analysis Group Discussion paper*, EAG 88-8, U.S. Department of Justice;
- [86] Froeb, L.M. and M. Shor (2005), "Auctions, Evidence, and Antitrust," in: Harkrider, J. (ed.), *The Use of Econometrics in Antitrust*, American Bar Association;

- [87] Genesove, D. and W.P. Mullin (2001), "Rules, Communication, and Collusion: Narrative Evidence from the Sugar Institute Case," *American Economic Review*, 91(3), 379-398;
- [88] Gilligan, T.W. (1992), "Imperfect Competition and Basing-Point Pricing: Evidence from the Softwood Plywood Industry," *American Economic Review*, 82, 1106-1119;
- [89] Goppelsroeder, M., (2008), "Collusion and Entry: an Experimental Study," ACLE working paper, No. 2009-05;
- [90] Goppelsroeder, M., M.P. Schinkel and J. Tuinstra (2008), "Corporate Leniency programs in the Cartel Lifecycle: 'Cleaning out the Closet'," *ACLE working paper*;
- [91] Graham, D.A. and R.C. Marshall (1987), "Collusive Bidder Behavior at Single Object Second Price and English Auctions," *Journal of Political Economy*, 86, 1217-1239;
- [92] Graham, D.A., R.C. Marshall and J-F. Richard (1990), "Differential Payments within a Bidder Coalition and the Shapley Value," *American Economic Review*, 80(3), 493-510;
- [93] Granitz, E. and B. Klein (1996), "Monopolization by 'Raising Rivals' Costs': The Standard Oil Case," *Journal of Law and Economics*, 39(1), 1-47;
- [94] Green, E. and R.H. Porter (1984), "Noncooperative Collusion under Imperfect Price Information," *Econometrica*, 52, 87-100;
- [95] Greenhut, M.L. (1981), "Spatial Pricing in the USA, West Germany, and Japan," *Economica*, 48, 79-86;
- [96] Greenhut, M.L. (1987), "Basing point system," in: J. Eatwell, M. Milgate and P. Newman (eds.), *The New Palgrave: A Dictionary of Economics*, Vol. 1, Stockwell Press, New York;
- [97] Griffin, J.M. (1989), "Previous Cartel Experience: Any Lessons for OPEC?," in: Lawrence R. Klein and J. Marquez (eds.), *Economics in Theory and Practice: an Eclectic Approach*," Dordrecht: Kluwer Academic Publishers, 179-206;
- [98] Grout, P. and S. Sonderegger (2005), "Predicting Cartels", *Economic Discussion Paper OFT 773*, OFT, London;
- [99] Grunzel, J. (1902), "*Über kartelle*," Leipzig: Duncker & Humblot;
- [100] Gupta, B. (1997), "Collusion in the Indian Tea Industry in the Great Depression: an Analysis of Panel Data," *Explorations in Economic History*, 34, 155-173;
- [101] Haddock, D.D. (1982), "Basing-Point Pricing: Competitive vs. Collusive theories," *American Economic Review*, 72, 289-307;

- [102] Haddock, D.D. (1990), "On the Basing-Point System: Comment," *American Economic Review*, 80, 957-962;
- [103] Halliday, J. and P. Seabright (2001), "Networks good, cartels bad: but how could anyone tell the difference?," in: *Fighting Cartels*, report of the Swedish competition authority;
- [104] Harrington, J.E. (1989), "Collusion among asymmetric firms: The case of different discount factors," *International Journal of Industrial Organization*, 7(2), 289-307;
- [105] Harrington, J.E. (1991), "The determination of price and output quotas in a heterogeneous cartel," *International Economic Review*, 32(4), 767-792;
- [106] Harrington, J.E. (2004), "Cartel Pricing Dynamics in the Presence of an Antitrust Authority," *RAND Journal of Economics*, 35(4), 651-673;
- [107] Harrington, J.E. (2005), "Optimal Cartel Pricing in the Presence of an Antitrust Authority," *International Economic Review*, 46(1), 145-169;
- [108] Harrington, J.E. (2006), "How do Cartels Operate?," *Foundations and Trends in Microeconomics*, 2(1), 1-105;
- [109] Harrington, J.E. (2008), "Detecting Cartels," in: *Handbook in Antitrust Economics*, Buccirosi, P., ed., MIT Press, Cambridge, MA;
- [110] Harrington, J.E. and J. Chen (2006), "Cartel Pricing Dynamics with Cost Variability and Endogenous Buyer Detection," *International Journal of Industrial Organization*, 24, 1185-1212;
- [111] Hart, S. and M. Kurz (1983), "Endogenous Formation of Coalitions," *Econometrica*, 51(4), 1047-1064;
- [112] Hay, G. and D. Kelly (1974), "An Empirical Survey of Price Fixing Conspiracies," *Journal of Law and Economics*, 17, 13-38;
- [113] Hay, G.A. (1999), "Practices that Facilitate Cooperation: the Ethyl Case," in: J. Kwoka and L. White (eds.), *The Antitrust Revolution*, 3rd ed., Oxford University Press, New York;
- [114] Hinloopen, J. and A.R. Soetevent (2008), "Laboratory evidence on the effectiveness of corporate leniency programs," *RAND Journal of Economics*, 39(2), 607-616;
- [115] Hotelling, H. (1929), "Stability in Competition," *Economic Journal*, 39(153), 41-57;
- [116] Hviid, M. (1992), "Endogenous Cartel Formation with Private Information," *Canadian Journal of Economics*, 25(4), 972-982;

- [117] Kandori, M. and H. Matsushima (1998), "Private Observation, Communication and Collusion," *Econometrica*, 66(3), 627-652;
- [118] Karlson, S.H. (1990), "Competition and Cement Basing Points: F.O.B. Destination, Delivered from Where?," *Journal of Regional Science*, 30, 75-88;
- [119] Kaysen, C. (1949), "Basing Point Pricing and Public Policy," *Quarterly Journal of Economics*, 63, 289-314;
- [120] Kinghorn, J.R. and R. Nielsen (2004), "A Practice without Defenders: The Price Effects of Cartelization," in: *How Cartels Endure and How They Fail*, Grossman, P.Z. (ed.), Edward Elgar;
- [121] Koskoff, D.E. (1981), "*The Diamond World*," New York: Harper & Row;
- [122] Kühn, K.-U. and M. Motta (1999), "The Economics of Joint Dominance," Working paper, Florence: European University Institute;
- [123] Lambson, V.E. (1987), "Optimal Penal Codes in Price-Setting Supergames with Capacity Constraints," *Review of Economic Studies*, 54(3), 385-397;
- [124] Lambson, V.E. (1994), "Some Results on Optimal Penal Codes in Asymmetric Bertrand Supergames," *Journal of Economic Theory*, 62, 444-468;
- [125] Leibenstein, H. (1966), "Allocative Efficiency vs. X-Efficiency," *American Economic Review*, 56, 392-415;
- [126] Levenstein, M.C. (1995), "Mass Production Conquers the Pool: Firm Organization and the Nature of Competition in the Nineteenth Century," *Journal of Economic History*, 55(3), 575-611;
- [127] Levenstein, M.C. (1996), "Do Price Wars Facilitate Collusion? A Study of the Bromine Cartel before World War I," *Explorations in Economic History*, 33, 1107-1137;
- [128] Levenstein, M.C., V.Y. Suslow and L. Oswald (2003), "International Price-Fixing Cartels and Developing Countries: A Discussion of Effects and Policy Remedies," *William Davidson Working Paper*, 538;
- [129] Levenstein, M.C. and V.Y. Suslow (2004), "Studies of Cartel Stability: A Comparison of Methodological Approaches," in: Grossman, P.Z. (ed.), *How Cartels Endure and How They Fail*, Cheltenham, UK: Edward Elgar Publishing, 9-52;
- [130] Levenstein, M.C. and V.Y. Suslow (2006), "What Determines Cartel Success?," *Journal of Economic Literature*, 44(1), 43-95;
- [131] Levy, D.T. and J.D. Reitzes (1993), "Basing-Point Pricing and Incomplete Collusion," *Journal of Regional Science*, 33(1), 27-36;
- [132] Liefmann, R. (1977), "*Cartels, Concerns and Trusts*," New York: Arno Press;

- [133] Loescher, S.M. (1959), "*Imperfect Collusion in the Cement Industry*," Cambridge, MA: Harvard University Press;
- [134] Loescher, S.M. (1980), "Economic Collusion, Civil Conspiracy, and Treble Damage Deterrents: The Sherman Act Breakthrough with Southern Plywood," *Quarterly Review of Economics and Business*, 20, 6-35;
- [135] Machlup, F. (1949), "*The Basing-point System*," Philadelphia, The Blakiston Company;
- [136] MacKie-Mason, J.K. and R.S. Pindyck (1987), "Cartel Theory and Cartel Experience in International Mineral Markets," in: Gordon, Jakoby and Zimmerman (eds.), *Energy: Markets and Regulation*, MIT Press, Cambridge, Massachusetts, 187-214;
- [137] Mailath, G. and P. Zemsky (1991), "Collusion in Second-Price Auctions with Heterogeneous Bidders," *Games and Economic Behavior*, 3, 467-486;
- [138] Marengo, L.M. (1955), "The Basing Point Decisions and the Steel Industry," *American Economic Review*, 45(2), 509-522;
- [139] Markham, J.W. (1951), "The Nature and Significance of Price Leadership," *American Economic Review*, 41(5), 891-905;
- [140] Marshall, R.C., M.J. Meurer, J-F. Richard and W. Stromquist (1994), "Numerical Analysis of Asymmetric First Price Auctions," *Games and Economic Behavior*, 7, 193-220;
- [141] Marshall, R.C. and L.M. Marx (2007), "Bidder Collusion," *Journal of Economic Theory*, 133, 374-402;
- [142] Martin, S. (1990), "Fringe Size and Cartel Stability," *European University Institute Working Paper 90/16*;
- [143] Martin, S. (2002), "*Advanced Industrial Economics*," 2nd edition, Blackwell Publishing, UK;
- [144] Maskin, E. and J. Tirole (1988), "A Theory of Dynamic Oligopoly, 2: Price Competition, Kinked Demand Curves, and Edgeworth Cycles," *Econometrica*, 56(3), 571-599;
- [145] McAfee, P. and J. MacMillan (1992), "Bidding Rings," *American Economic Review*, 82, 579-599;
- [146] McAfee, P. (1994), "Endogenous Availability, Cartels, and Merger in an Equilibrium Price Dispersion," *Journal of Economic Theory*, 62, 24-47;
- [147] McAnney, J.W. (1991), "The Justice Department's Crusade against Price-Fixing: Initiative or Reaction?," *Antitrust Bulletin*, 521-542;

- [148] McCutcheon, B. (1997), "Do Meetings in Smoke-Filled Rooms Facilitate Collusion?," *Journal of Political Economy*, 105, 330-350;
- [149] Motta, M. and M. Polo (2003), "Leniency Programs and Cartel Prosecution," *International Journal of Industrial Organization*, 21(3), 347-379;
- [150] Motta, M. (2004), "*Competition Policy: Theory and Practice*," Cambridge: Cambridge University Press;
- [151] Neumann, M. (2001), "*Competition Policy: History, Theory and Practice*," Cheltenham: Edward Elgar;
- [152] Newmark, C.M. (1988), "Does Horizontal Price Fixing Raise Price? A Look at the Bakers of Washington Case," *Journal of Law and Economics*, 31(2), 469-484;
- [153] Noel, M.D. (2007), "Edgeworth Price Cycles: Evidence from the Toronto Retail Gasoline Market," *Journal of Industrial Economics*, 55(1), 69-92;
- [154] Perry, M. and R. Porter (1985), "Oligopoly and the Incentives for Horizontal Merger," *American Economic Review*, 75, 219-227;
- [155] Pesendorfer, M. (2000), "A Study of Collusion in First-Price Auctions," *Review of Economic Studies*, 67, 381-411;
- [156] Pindyck, R.S. (1979), "The Cartelization of World Commodity Markets," *American Economic Review*, 69(2), 154-158;
- [157] Pindyck, R.S. and D.L. Rubinfeld (2001), "*Microeconomics*," fifth edition, New Jersey: Prentice Hall;
- [158] Piotrowski, R. (1933), "*Cartels and Trusts: their origin and historical development from the economic and legal aspects*," Porcupine Press: Philadelphia;
- [159] Porter, R.H. (1983), "A Study of Cartel Stability: The Joint Executive Committee, 1880-1886," *Bell Journal of Economics*, 14, 301-314;
- [160] Porter, R.H. (2005), "Detecting Collusion," *Review of Industrial Organization*, 26(2), 147-167;
- [161] Porter, R.H. and J.D. Zona (1993), "Detection of Bid Rigging in Procurement Auctions," *Journal of Political Economy*, 101, 79-99;
- [162] Porter, R.H. and J.D. Zona (1999), "Ohio School Milk Markets: An Analysis of Bidding," *RAND Journal of Economics*, 30(2), 263-288;
- [163] Posada, P. (2001), "Leadership Cartels in Industries with Differentiated Products," *Warwick Economic Research Papers*, No. 590;
- [164] Posner, R.A. (1970), "A Statistical Study of Antitrust Law Enforcement," *Journal of Law and Economics*, 13, 365-419;

- [165] Prokop, J. (1999), "Process of Dominant-Cartel Formation," *International Journal of Industrial Organization*, 17(2), 241-257;
- [166] Ray, D. and R. Vohra (1999), "A Theory of Endogenous Coalition Structures," *Games and Economic Behavior*, 26, 286-336;
- [167] Rey, P. (2007), "On the Use of Economic Analysis for Cartel Detection," in: *European Competition Law Annual 2006: Enforcement of the Prohibition of Cartels*, C.-D. Ehlermann and I. Atanasiu (eds.), Oxford: Hart;
- [168] Ross, T.W. (1992), "Cartel Stability and Product Differentiation," *International Journal of Industrial Organization*, 10, 1-13;
- [169] Rossmo, D.K. (1999), "*Geographic Profiling*," CRC Press, Baton Rouge, FL;
- [170] Rotemberg, J.J. and G. Saloner (1986), "A supergame-theoretic model of price wars during booms," *American Economic Review*, 76(3), 390-407;
- [171] Rothschild, R. (1992), "On the Sustainability of Collusion in Differentiated Duopolies," *Economics Letters*, 40, 33-37;
- [172] Rothschild, R. (1999), "Cartel Stability when Costs are Heterogeneous," *International Journal of Industrial Organization*, 17(5), 717-734;
- [173] Salant, S.W. (1976), "Exhaustible Resources and Industrial Structure: A Nash-Cournot Approach to the World Oil Market," *Journal of Political Economy*, 84(5), 1079-1094;
- [174] Salant, S.W., S. Switzer and R.J. Reynolds (1983), "Losses from Horizontal Merger: The Effects of an Exogenous Change in Industry Structure on Cournot-Nash Equilibrium," *Quarterly Journal of Economics*, 98(2), 185-199;
- [175] Salop, S.C. (1979), "Monopolistic Competition with Outside Goods," *Bell Journal of Economics*, 10(1), 141-156;
- [176] Salop, S.C. and D.T. Scheffman (1983), "Raising Rivals' Costs," *American Economic Review*, 73(2), 267-271;
- [177] Salop, S.C. and D.T. Scheffman (1987), "Cost-Raising Strategies," *Journal of Industrial Economics*, 36(1), 19-34;
- [178] Scherer, F.M. (1980), "*Industrial Market Structure and Economic Performance*," 2nd edition, Houghton Mifflin Company, Boston;
- [179] Scherer, F.M. and D. Ross (1990), "*Industrial Market Structure and Economic Performance*," 3rd ed., Boston: Houghton Mifflin Company;
- [180] Schinkel, M.P. (2008), "Forensic Economics in Competition Law Enforcement," *Journal of Competition Law and Economics*, 4(1), 1-30;
- [181] Schinkel, M.P. and J. Tuinstra (2006), "Imperfect competition law enforcement," *International Journal of Industrial Organization*, 24(6), 1267-1297;

- [182] Schinkel, M.P., J. Tuinstra and J. Rüggeberg (2008), "Illinois Walls: How Barring Indirect Purchaser Suits Facilitates Collusion," *RAND Journal of Economics*, 39(3), 683-698;
- [183] Schumpeter, J. (1912), "*Theorie der Wirtschaftlichen Entwicklung*," Leipzig: Duncker & Humblot;
- [184] Selten, R. (1973), "A Simple Model of Imperfect Competition, where 4 are Few and 6 are Many," *International Journal of Game Theory*, 2(3), 141-201;
- [185] Shaffer, S. (1995), "Stable Cartels with a Cournot Fringe," *Southern Economic Journal*, 61(3), 744-754;
- [186] Shenoy, P. (1979), "On Coalition Formation: A Game Theoretic Approach," *International Journal of Game Theory*, 8, 133-164;
- [187] Slade, M.E. (1990), "Cheating on Collusive Agreements," *International Journal of Industrial Organization*, 8, 519-543;
- [188] Smith, A. (1994), "*The Wealth of Nations*," Edwin Cannan (ed.), New York: The Modern Library;
- [189] Smith, V.L. (2008), "*Rationality in Economics: Constructivist and Ecological Forms*," Cambridge: Cambridge University Press;
- [190] Smithies, A. (1941), "Monopolistic Price Policy in a Spatial Market," *Econometrica*, 9(1), 63-73;
- [191] Smithies, A. (1942), "Aspects of the Basing-Point System," *American Economic Review*, 32, 705-726;
- [192] Soper, J.B., G. Norman, M.L. Greenhut and B.L. Benson (1991), "Basing Point Pricing and Product Concentration," *Economic Journal*, 101, 539-556;
- [193] Spar, D.L. (1994), "*The Cooperative Edge: The Internal Politics of International Cartels*," Ithaca: Cornell University Press;
- [194] Staiger, R.W. and F.A. Wolak (1992), "Collusive Pricing with Capacity Constraints in the Presence of Demand Uncertainty," *RAND Journal of Economics*, 23(2), 203-220;
- [195] Stigler, G.J. (1949), "A Theory of Delivered Price Systems," *American Economic Review*, 39, 477-489;
- [196] Stigler, G.J. (1950), "Monopoly and Oligopoly by Merger," *American Economic Review*, 40(2), 23-34;
- [197] Stigler, G.J. (1964), "A Theory of Oligopoly," *Journal of Political Economy*, 72(1), 44-61;
- [198] Stigler, G.J. (1983), "*The Organization of Industry*," Chicago: University of Chicago Press;

- [199] Stocking, G.W. and M.W. Watkins (1948), “*Cartels or Competition?: The Economics of International Controls by Business and Government*,” New York: The Twentieth Century Fund;
- [200] Thisse, J-F and X. Vives (1988), “On the Strategic Choice of Spatial Price Policy,” *American Economic Review*, 78, 122-137;
- [201] Thisse, J-F and X. Vives (1992), “Basing-Point Pricing: Competition versus Collusion,” *Journal of Industrial Economics*, 40, pp. 249-260.
- [202] Thoron, S. (1998), “Formation of a Coalition-Proof Stable Cartel,” *Canadian Journal of Economics*, 31(1), 63-76;
- [203] Tirole, J. (1988), “*The Theory of Industrial Organization*,” MIT Press, Cambridge, Massachusetts;
- [204] Van Cayseele, P. and D. Furth (1996a), “Bertrand-Edgeworth Duopoly with Buyouts or First Refusal Contracts,” *Games and Economic Behavior*, 16, 153-180;
- [205] Van Cayseele, P. and D. Furth (1996b), “von Stackelberg’s equilibria for Bertrand-Edgeworth duopoly with buyouts,” *Journal of Economic Studies*, 23, 96-109;
- [206] Van Cayseele, P. and D. Furth (2001), “Two is Not Too Many for Monopoly,” *Journal of Economics*, 74(3), 231-258;
- [207] Vasconcelos, H. (2005), “Tacit Collusion, Cost Asymmetries, and Mergers,” *RAND Journal of Economics*, 36, 39-62;
- [208] Wang, Z. (2005), “Edgeworth Price Cycle and Oligopoly Coordination: Trial Evidence from Australia,” *Northeastern University Working Paper*;
- [209] Whinston, M.D. (2006), “*Lectures on Antitrust Economics*,” MIT Press, Cambridge, Massachusetts;
- [210] Yi, S.S. (1997), “Stable Coalition Structures with Externalities,” *Games and Economic Behavior*, 20, 201-237.

Samenvatting (Summary in Dutch)

Eén van de belangrijkste kenmerken van een kapitalistisch economisch systeem is vrije marktwerking. Vrije marktconcurrentie leidt in de regel tot hoge maatschappelijke welvaart, maar ondermijnt tegelijkertijd de winstgevendheid van ondernemingen. Bedrijven zullen bijvoorbeeld in het gevecht om de klant hun prijzen laag moeten houden. Concurrentiedruk is bevorderlijk voor de consument en de maatschappelijke welvaart, maar leidt tegelijkertijd tot relatief geringe winstmarges. Dit is de reden dat in een vrije markteconomie bedrijven een natuurlijke neiging hebben om de onderlinge concurrentie te beperken of zelfs uit te schakelen.

Eén van de manieren om de concurrentie te beperken is door middel van een kartelafsprake. Bedrijven kunnen bijvoorbeeld een (minimum) prijs afspreken, waaronder geen product wordt aangeboden. Ook kunnen zij afzetgebieden onderling verdelen en een kunstmatige schaarste creëren door productiequota af te spreken. Dergelijke afspraken leiden tot hogere winsten, maar komen voor rekening van de consument en de maatschappelijke welvaart. Dit is de reden dat in veel vrije markteconomieën kartelafspraken verboden zijn.

Bestaande economische theorieën laten zien dat een kartelafsprake leidt tot een maximale vergroting van de winst wanneer alle bedrijven in een markt participeren. Met andere woorden, het optimale kartel is een compleet kartel, omdat zo een afspraak alle marktconcurrentie uitschakelt. Echter, in de praktijk zijn de meeste veroordeelde kartels incompleet in de zin dat de kartelleden een gezamenlijk marktaandeel hadden van minder dan 100%. Er is sprake van een incompleet kartel wanneer minstens één bedrijf in de markt niet meedoet.

Deze dissertatie handelt over incomplete kartels en is georganiseerd rond vier hoofdonderzoeksvragen: (i) Onder welke condities verschilt de optimale kartelgrootte van de maximale kartelgrootte? (ii) Gegeven dat het kartel incompleet is, welke (soort) bedrijven doen mee in het kartel en welke (soort) bedrijven blijven onafhankelijk con-

curreren? (iii) Wat is de relatie tussen marktstructuur en optimale kartelgrootte? (iv) Hoe kan de economische wetenschap worden ingezet om (incomplete) kartels op te sporen? Het proefschrift bestaat uit twee delen. De eerste drie vragen worden behandeld in het eerste deel (Hoofdstuk 2 en 3) en de laatste vraag komt aan de orde in het tweede deel (Hoofdstuk 4 en 5).

In Hoofdstuk 2 wordt een literatuuroverzicht gepresenteerd, waarbij zowel empirische als theoretische studies over incomplete kartels worden bediscussieerd. Er bestaat een beperkt aantal gegevens over incomplete kartels in praktijk, maar deze zijn wel min of meer eenduidig. Op basis van empirische studies naar kartels kunnen de volgende conclusies worden getrokken: (i) kartels zijn vaak incompleet (ii) incomplete kartels domineren vaak de industrie waarin zij opereren (iii) het marktaandeel van een incompleet kartel daalt over de tijd, en (iv) deelnemers aan een incompleet kartel zijn vaak de grotere spelers in de markt.

De theoretische literatuur over incomplete kartels kan worden ingedeeld aan de hand van drie hoofdvragen. Ten eerste, wanneer is een incompleet kartel winstgevend? Ten tweede, wanneer is een incompleet kartel intern stabiel? Ten derde, hoe komt een incompleet kartel tot stand? De bestaande theorie laat zien dat wanneer het kartel haar beslissingen neemt voordat niet-kartelleden dat doen, alle kartels winstgevend zijn ongeacht het marktaandeel. Echter, wanneer kartelleden en niet-kartelleden hun beslissingen tegelijkertijd nemen, dan is een incompleet kartel nooit winstgevend wanneer goederen homogeen zijn, terwijl alle incomplete kartels winstgevend zijn in markten met gedifferentieerde goederen. Wanneer bedrijven concurreren in hoeveelheden in plaats van in prijs, dan zijn slechts de incomplete coalities met voldoende marktaandeel winstgevend.

De tweede vraag gaat over de interne stabiliteit van incomplete kartels. Met andere woorden, onder welke condities willen de kartelleden zich vrijwillig aan de kartelafpraak houden? Het blijkt dat wanneer het kartel en niet-kartelleden hun beslissingen tegelijkertijd nemen en goederen homogeen zijn, een incompleet kartel intern stabiel is wanneer een compleet kartel dat ook is. Echter, in andere theoretische modellen is het mogelijk dat een incompleet kartel intern stabiel is, terwijl een kartel met honderd procent marktaandeel dat niet is. De noodzaak voor interne stabiliteit van kartelafspraken vormt daarmee een belangrijk argument voor het bestaan van incomplete kartels. Bedrijven vormen een winstgevend incompleet kartel, omdat een volledig kartel niet kan worden gehandhaafd.

Een derde belangrijke vraag die aan de orde komt is onder welke omstandigheden bedrijven een prikkel hebben om een incompleet kartel te vormen. Dit is een kernonderwerp in de literatuur en het wordt in het proefschrift samengevat onder de noemer “participation puzzle”. Deze “kartelformatie puzzel” ontstaat omdat ieder bedrijf profiteert van kartelvorming, maar buitenstaanders profiteren meer van het kartel dan de kartelleden zelf. De verklaring hiervoor is dat kartelleden hun productieniveau moeten verlagen om een prijsstijging te bewerkstelligen, terwijl niet-kartelleden vaak zowel hun prijs als productieniveau verhogen. Ieder bedrijf wil daarom het liefst dat al zijn concurrenten een (incomplete) kartel vormen, zodat het, zonder zelf iets illegaals te ondernemen, zijn winsten meer ziet stijgen dan de winsten van zijn concurrenten. De “participation puzzle” maakt het moeilijk te verklaren hoe incomplete kartels ontstaan.

De bevindingen in Hoofdstuk 2 vormen het uitgangspunt voor de analyse in Hoofdstuk 3. In Hoofdstuk 3 wordt een dynamisch spel geanalyseerd, waarin de participatiekeuze endogeen is. Een belangrijke uitbreiding op bestaande theorieën is dat bedrijven niet-identiek zijn; zij verschillen in beschikbare productiecapaciteit. De totale winst van het kartel evenals de winst per kartellid hangt positief af van de hoeveelheid productiecapaciteit die onder de controle van het kartel is. In feite hangt de winstgevendheid van een kartel daarom af van het aantal bedrijven dat deelneemt in de coalitie. In deze setting is de optimale kartelgrootte gelijk aan de maximale kartelgrootte zolang kartelvorming geen kosten met zich meebrengt. Anders gezegd, wanneer bedrijven een kartel kunnen vormen zonder extra kosten, dan kan de totale kartelwinst altijd zo worden verdeeld dat ieder bedrijf wil participeren.

De aanname dat kartelvorming “gratis” is, is echter verre van realistisch. Bedrijven moeten bijvoorbeeld onderhandelen over de inhoud van het kartelcontract. Ook zullen kartelleden in de regel investeringen moeten doen om het kartel intern stabiel te houden. Bovendien zijn de meeste kartelafspraken in praktijk illegaal. Dit betekent dat bedrijven een risico nemen wanneer zij deelnemen in een kartel. Er bestaat een kans dat het kartel ontdekt wordt en dit betekent doorgaans substantiële boetes voor kartelparticipanten. De kosten van kartelvorming worden gemodelleerd door middel van een kostenfunctie die positief afhangt van het aantal deelnemers. De positieve relatie wordt aangenomen omdat onderhandelingen, *ceteris paribus*, moeilijker zullen zijn wanneer meer bedrijven rond de tafel zitten. Ook zal de kans op ontdekking groter zijn, omdat de kans groter is dat één van de kartelleden het kartel verklikt in ruil voor een sterke verlaging van de boete of volledige boete-immuniteit. Tot slot, zal de impact van het kartel op de markt afhangen van het aantal deelnemers waardoor grotere kartels, *ceteris paribus*, meer zichtbaar zijn, wat de kans op ontdekking weer vergroot.

Rekening houdend met de kosten van kartelvorming worden de volgende voorname resultaten gevonden. Allereerst is een incompleet kartel meer winstgevend dan een volledig kartel wanneer de kleinste bedrijven in de markt ‘voldoende klein’ zijn. De reden hiervoor is dat als een heel klein bedrijf toetreedt tot het kartel de winst nauwelijks zal toenemen, terwijl de kosten van het kartel significant stijgen. Ten tweede hangt de prikkel om toe te treden tot het kartel samen met de grootte van het bedrijf. Meer specifiek, de prikkel voor kartelvorming hangt positief af van de grootte van een bedrijf. Zeer kleine bedrijven dragen dus niet alleen weinig bij aan de winstgevendheid van het kartel; zij hebben ook het minst de behoefte om deel te nemen. Meer algemeen zijn incomplete kartels te verwachten in markten met een of meer relatief kleine ondernemingen. Ten derde, een kartel bestaande uit de grootste ondernemingen is een uitkomst van het model, maar deze uitkomst is in de regel niet uniek. Het is dus zeer wel mogelijk dat kartels voorkomen waarbij de grootste buitenstaanders groter zijn dan de kleinste deelnemers. Tot slot wordt in het hoofdstuk onderzocht hoe verschuivingen in productiecapaciteit kartelvorming beïnvloeden. Zo een verschuiving vindt bijvoorbeeld plaats wanneer twee of meer bedrijven fuseren; een aantal bedrijven vormen dan samen een groter bedrijf. Dit kan een impact hebben op kartelvorming, omdat twee kleine bedrijven die geen prikkel hebben om in een kartel deel te nemen na de fusie die prikkel wel hebben. Uit de analyse blijkt dat de grootste impact komt van fusies tussen middelgrootte ondernemingen. Dit levert een tegenstelling op met

de bestaande literatuur, waarin wordt gesteld dat fusies tussen de grootste en/of de kleinste ondernemingen de grootste coördinerende effecten hebben.

In Hoofdstuk 4 wordt de aandacht verlegd naar karteldetectie. In dit hoofdstuk wordt bediscussieerd hoe economische theorieën kunnen helpen bij het opsporen van kartels. Omdat kartelafspraken meestal illegaal zijn, zullen kartelleden trachten de afspraak verborgen te houden. Men zal bijvoorbeeld vergaderen op geheime locaties en weinig sporen nalaten die eventueel in de rechtszaal als bewijs kunnen dienen. Het voornaamste voordeel van economische detectiemethoden is dat wordt uitgegaan van het resultaat van de kartelafpraak. Wellicht zullen managers het niet heel moeilijk vinden om stiekem een hogere prijs overeen te komen, maar het implementeren van deze afspraak zal leiden tot een prijsstijging en dit laatste is in principe zichtbaar.

De standaardbenadering in de economische detectieliteratuur is om voor een bepaalde markt zowel concurrentie als kartelgedrag te modelleren. Op basis van beschikbare data wordt vervolgens getest door welk van de twee modellen het geobserveerde marktgedrag het beste kan worden verklaard. Meer algemeen is een groep bedrijven verdacht wanneer hun marktgedrag significant afwijkt van het verwachte competitieve gedrag. Om een idee te vormen over hoe competitief marktgedrag er in een bepaalde markt uit zal zien kan worden gekeken naar marktgedrag op vergelijkbare markten. Ook kan het marktgedrag in vorige perioden als ijkpunt dienen. Als het vermoeden bestaat dat sprake is van een incompleet kartel kan bovendien worden gekeken naar het gedrag van bedrijven onderling. Dit komt omdat kartelleden in de regel significant ander marktgedrag vertonen dan bedrijven die geen onderdeel uitmaken van het kartel. Een belangrijke conclusie in dit hoofdstuk is dat economische detectiemethoden vooral effectief zullen zijn wanneer zij rekening houden met de bijzonderheden van een bepaalde industrie.

In Hoofdstuk 5 wordt een nieuwe detectiemethode ontwikkeld die kan worden toegepast op industrieën, waarin gebruik wordt gemaakt van het zogenaamde “basing-point systeem”. Bij deze prijsmethode wordt een klant een prijs in rekening gebracht inclusief transportkosten die worden berekend op basis van de afstand tussen een klantlocatie en een andere locatie; het basing-point. Dit systeem wordt veelvuldig toegepast in markten voor homogene bulkgoederen. Voorbeelden zijn de markt voor cement, staal en suiker. Er bestaan voorbeelden van incomplete kartels die dit systeem hebben misbruikt om extra winsten mee te genereren. Het “basing-point systeem” is een handig systeem voor kartels, omdat bedrijven slechts een basing-point hoeven af te spreken van waaraf alle prijzen worden berekend. In competitie zullen, gegeven een bepaalde klantlocatie, bedrijven hun dichtstbijzijnde concurrent als basing-point gebruiken. Samen kunnen bedrijven echter een punt ver weg van de klantlocaties afspreken en aldus extra transportkosten in rekening brengen die in werkelijkheid niet worden gemaakt. Omdat de klant een offerte ontvangt met daarop de totaalprijs is het moeilijk te controleren welk basing-point is toegepast. In het hoofdstuk wordt aangetoond dat dit soort kartelafspraken moeilijk te ontdekken zijn met bestaande detectiemethodes.

De hier ontwikkelde detectiemethode maakt vervolgens gebruik van het inzicht dat bedrijven in concurrentie en onder een kartelregime een verschillend basing-point hanteren. De toegepaste basing-points zijn op het eerste gezicht onbekend, maar kunnen worden achterhaald met behulp van transactiedata. Gegeven een bepaalde prijsformule kan op basis van enkele transacties de meest waarschijnlijke basing-point

locatie worden achterhaald. De techniek die wordt gehanteerd lijkt sterk op die van het “global positioning system” (GPS). Simpel gezegd kun je om iedere consumentenlocatie een cirkel trekken die de afstand van het gebruikte basing-point weergeeft en de positie waar al deze cirkels snijden is het meest waarschijnlijke basing-point. Op basis van de beschreven theorie kan een conclusie worden getrokken over de aanwezigheid van een kartel. Basing-points die relatief dicht bij elkaar liggen en ver weg van klantlocaties en bedrijven zijn verdacht, terwijl basing-points meer verspreid en in de buurt van bedrijfslocaties duiden op concurrentie. Een software is ontwikkeld om deze methode toe te passen en te analyseren. De benodigde gegevens zijn consumentenlocaties, bedrijfslocaties, prijzen en hoeveelheden. Het programma gebruikt deze gegevens om het meest waarschijnlijke basing-point te bepalen. Als uitkomst geeft het programma een waarde tussen 0 en 1. Een waarde dicht bij 0 komt overeen met een competitieve situatie, terwijl waarden dicht bij 1 een indicatie zijn voor kartelpraktijken.

The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus University Rotterdam, University of Amsterdam and VU University Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. The following books recently appeared in the Tinbergen Institute Research Series:

419. L. RATNOVSKI, *A Random Walk Down the Lombard Street: Essays on Banking.*
420. R.P. NICOLAI, *Maintenance models for systems subject to measurable deterioration.*
421. R.K. ANDADARI, *Local clusters in global value chains, a case study of wood furniture clusters in Central Java (Indonesia).*
422. V.KARTSEVA, *Designing Controls for Network Organizations: A Value-Based Approach.*
423. J. ARTS, *Essays on New Product Adoption and Diffusion.*
424. A. BABUS, *Essays on Networks: Theory and Applications.*
425. M. VAN DER VOORT, *Modelling Credit Derivatives.*
426. G. GARITA, *Financial Market Liberalization and Economic Growth.*
427. E.BEKKERS, *Essays on Firm Heterogeneity and Quality in International Trade.*
428. H.LEAHU, *Measure-Valued Differentiation for Finite Products of Measures: Theory and Applications.*
429. G. BALTUSSEN, *New Insights into Behavioral Finance.*
430. W. VERMEULEN, *Essays on Housing Supply, Land Use Regulation and Regional Labour Markets.*
431. I.S. BUHAI, *Essays on Labour Markets: Worker-Firm Dynamics, Occupational Segregation and Workplace Conditions.*
432. C. ZHOU, *On Extreme Value Statistics.*
433. M. VAN DER WEL, *Riskfree Rate Dynamics: Information, Trading, and State Space Modeling.*
434. S.M.W. PHILIPPEN, *Come Close and Co-Create: Proximities in pharmaceutical innovation networks.*
435. A.V.P.B. MONTEIRO, *The Dynamics of Corporate Credit Risk: An Intensity-based Econometric Analysis.*
436. S.T. TRAUTMANN, *Uncertainty in Individual and Social Decisions: Theory and Experiments.*
437. R. LORD, *Efficient pricing algorithms for exotic derivatives.*
438. R.P. WOLTHOFF, *Essays on Simultaneous Search Equilibrium.*
439. Y.-Y. TSENG, *Valuation of travel time reliability in passenger transport.*

440. M.C. NON, *Essays on Consumer Search and Interlocking Directorates.*
441. M. DE HAAN, *Family Background and Children's Schooling Outcomes.*
442. T. ZAVADIL, *Dynamic Econometric Analysis of Insurance Markets with Imperfect Information.*
443. I.A. MAZZA, *Essays on endogenous economic policy.*
444. R. HAIJEMA, *Solving large structured Markov Decision Problems for perishable-inventory management and traffic control.*
445. A.S.K. WONG, *Derivatives in Dynamic Markets.*
446. R. SEGERS, *Advances in Monitoring the Economy.*
447. F.M. VIEIDER, *Social Influences on Individual Decision Making Processes.*
448. L. PAN, *Poverty, Risk and Insurance: Evidence from Ethiopia and Yemen.*
449. B. TIEBEN, *The concept of equilibrium in different economic traditions: A Historical Investigation.*
450. P. HEEMEIJER, *Expectation Formation in Dynamic Market Experiments.*
451. A.S. BOOIJ, *Essays on the Measurement Sensitivity of Risk Aversion and Causal Effects in Education.*
452. M.I. LÓPEZ YURDA, *Four Essays on Applied Microeconometrics.*
453. S. MEENTS, *The Influence of Sellers and the Intermediary on Buyers' Trust in C2C Electronic Marketplaces.*
454. S. VUJIĆ, *Econometric Studies to the Economic and Social Factors of Crime.*
455. F. HEUKELOM, *Kahneman and Tversky and the Making of Behavioral Economics.*
456. G. BUDAI-BALKE, *Operations Research Models for Scheduling Railway Infrastructure Maintenance.*
457. T.R. DANIËLS, *Rationalised Panics: The Consequences of Strategic Uncertainty during Financial Crises.*
458. A. VAN DIJK, *Essays on Finite Mixture Models.*
459. C.P.B.J. VAN KLAVEREN, *The Intra-household Allocation of Time.*
460. O.E. JONKEREN, *Adaptation to Climate Change in Inland Waterway Transport.*
461. S.C. GO, *Marine Insurance in the Netherlands 1600-1870, A Comparative Institutional Approach.*
462. J. NIEMCZYK, *Consequences and Detection of Invalid Exogeneity Conditions.*