



EUROPEAN UNIVERSITY INSTITUTE
Department of Economics

Three Essays on Collusion and Mergers

Helder Vasconcelos

Thesis submitted for assessment with a view to obtaining
the degree of Doctor of the European University Institute

Florence
October 2002

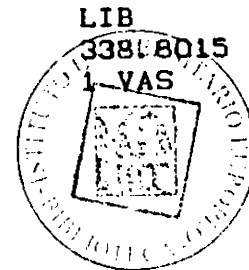
B/C →



EUROPEAN UNIVERSITY INSTITUTE
Department of Economics

Three Essays on Collusion and Mergers

Helder Vasconcelos



The Thesis Committee consists of:

Prof. Pierpaolo Battigalli, Università Bocconi

- Luis Cabral, Leonard Stern School of Business, NYU
- Massimo Motta, EUI, Supervisor
- Karl Schlag, EUI
- John Sutton, London School of Economics

To Anabela

Acknowledgments

In the process of writing this dissertation, I have accumulated many debts. Time has come to express my gratitude to those people and institutions that have contributed, in a variety of ways, for the preparation of this work.

My first debt is to my supervisor, Massimo Motta. I could not have been more fortunate with my supervisor. Massimo's supervision was outstanding: very knowledgeable and demanding. His permanent availability to discuss my research work and very fast feedback on preliminary drafts of my research papers should also be highlighted. Working under his supervision proved to be extremely enriching and stimulating.

My second debt is to Pierpaolo Battigalli. His co-supervision was also excellent. He very often read my work with patience and provided me very detailed comments on the more technical parts of the chapters, which substantially improved them. With him I learned how important it is to use clear and precise notation in economic modelling.

My special thanks are also due to Karl Schlag. Even though he was not formally my thesis advisor, I benefited enormously from several discussions in his office about my thesis two last chapters. I am grateful for a substantial amount of work and very useful suggestions on those chapters.

Many other people have helped to improve the quality of this thesis, by commenting on specific chapters of this dissertation. I am particularly indebted to Pedro

Pita Barros, Francis Bloch, Natalia Fabra, Chiara Fumagalli, Joseph E. Harrington, Jr., Andrea Ichino, Soren Johansen, Patrick Rey, Margaret Slade and Larry Samuelson. Thanks are also due to Robert Porter for providing me with the data set on the Joint Executive Committee, which I use in the empirical part of chapter two.

On the institutional side, a very warm acknowledgment goes to the Economics of Industry Group (STICERD) at the London School of Economics and Political Science, for the hospitality provided while I spent four months as a visiting research student. There I started the last chapter of this thesis under Professor John Sutton's supervision. I am indebted for his encouragement and invaluable suggestions. Moreover, I would like to thank Professor Sutton for having welcomed me as one of *his* students from the very first minute. Financial support from the Portuguese Ministry of Foreign Affairs, the Portuguese Ministry of Science and Technology and the European University Institute is gratefully acknowledged.

Seminar participants at the European University Institute, Research Institute for Industrial Economics (IUI, Stockholm), Universitat Pompeu Fabra, the University of Warwick, University of Lausanne, ISEG (Lisbon), Universidade do Minho and Universidade Nova de Lisboa as well as conference participants at the 1999 C.I.E. Summer School (University of Copenhagen), Second CEPR Conference in Applied Industrial Organization, the 2000 and 2001 annual conferences of the *European Economic Association* (EEA), 27th Conference of the *European Association*

for Research in Industrial Economics (EARIE) and ASSET 2001 Euroconference have commented on various stages of the work presented here.

I am also grateful to Jacqueline Bourgonje, Jessica Spataro and Marcia Gastaldo for their very friendly support and invaluable administrative help over the last four years.

Finally, a very special mention to those who have provided me the most support: my parents, my beloved sister Sandra and my wife Anabela. My parents and sister encouraged me to follow my own decision to perform this project even if this meant having to live apart from each other for four years time. I have no words to express my gratitude to them. They were simply exceptional, as always. My greatest debt is, however, to my wife Anabela. She sacrificed her professional career to be in Florence with me and provided me with her love and continuous support throughout the whole project. This thesis is dedicated to her.

Contents

Introductory overview	1
1 Entry Effects on Cartel Stability and the Joint Executive Committee	7
1.1 Introduction	7
1.2 The basic model	10
1.3 Considering entry	16
1.3.1 The new model structure	16
1.3.2 Case 1. Cartel Breakdown	19
1.3.3 Case 2. Accommodation	25
1.3.4 Stability discussion	30
1.4 The Joint Executive Committee	33
1.5 The data	34
1.6 The econometric model	37
1.7 Conclusion	42
2 Tacit Collusion, Cost Asymmetries and Mergers	44
2.1 Introduction	44
2.2 The model	48
2.3 The analysis of perfect efficient collusion	52
2.3.1 Collusive profits	53
2.3.2 Deviation profits	54

2.3.3 Distribution of assets and scope for collusion	55
2.4 Perfect non-efficient collusion.....	76
2.5 Conclusion.....	80
2.A The Cournot equilibrium.....	83
3 Towards a Characterization of the Upper Bound to Concentration in Endogenous Sunk Cost Industries	84
3.1 Introduction.....	84
3.2 The basic model.....	88
3.2.1 The game	89
3.2.2 Equilibrium analysis	92
3.3 Conclusion.....	129
3.A Stability requirements	131
3.B Upper bound to the number entrants anticipating a duopoly coalition structure.....	134
Bibliography	135

Introductory overview

This thesis aims at contributing to the analysis of the issues of collusion and mergers from an industrial organization perspective. The thesis is composed of three main chapters. In what follows each of those chapters will be discussed in greater detail. In particular, I will briefly describe the model considered in each chapter, its relation to the literature, and the contribution each chapter makes to the existing literature.

Chapter one deals with the issue of entry and collusion both theoretically and empirically. In particular, it re-considers Green and Porter's (1984) model of collusion with imperfect monitoring (in the price version, as in Tirole's (1988) textbook). In this model with homogeneous goods, a firm can either charge the monopoly price to share the market with other firms or secretly undercut its rivals to get the whole market. Since market demand is affected by stochastic factors, firm's demand could be zero even when no firm is deviating. Given that the only available information to each firm is its own sales, it is difficult for them to infer whether a low demand is caused by a secret price cut or it is just due to a "bad" demand shock. In equilibrium, the firms need to start a price-war (i.e., charge the competitive price) when the demand is zero to maintain the incentive to charge the monopoly price in the collusive phases. As a result, a price-war occurs with positive probability as an equilibrium phenomenon.

New in this chapter is the existence of potential entrants in the cartel. The chapter examines the structure and success of a cartel, of the type described above, when there exists a pool of potential entrants. To do so, a no-entry condition has to be taken into account

in addition to the standard incentive compatibility constraint to maintain the collusive price level in the collusive phases. A potential entrant is assumed to enter the market (and the cartel) in any period (which is an irreversible decision) if its benefit from a continuation equilibrium exceeds the initial entry fixed sunk cost. Hence, how incumbent firms react to an entrant is important to determine how severe the no-entry constraint is. Two different incumbents' response to entry are considered: (i) reversion to finite-punishment phase, and (ii) accommodation of entrants. Naturally, it is shown that the no-entry constraint is most difficult to be satisfied when incumbent firms accommodate the entrant, which was actually the case for the entry episodes for the nineteenth-century US railroad cartel whose pricing decisions are discussed in the applied part of the chapter.

The chapter shows that the optimal length of a price war increases as the number of firms in the cartel agreement increases if and only if the entry cost is so high that the no-entry constraint is not binding. The intuition that underlies this result is simple. When there are more firms in the agreement, each firm has a smaller share of the market. Then, first, the gain from secret price cutting becomes larger as a deviating firm can obtain a larger share of the market, and, second, each firm's continuation payoff in the collusive equilibrium decreases, which reduces the impact of any finite-period punishment. Combining these two effects, one concludes that the duration of a "price-war" has to *increase* to maintain the incentive for the firms to stick to the collusive price.

Since the demand shock is assumed to be independent and identically distributed (i.i.d.), the probability to trigger a price-war is the same for every period in a collusive phase. Given this, it is also shown that the percentage of periods spent in a price-war in-

creases with the length of the punishment phase, and hence with the number of active firms in the cartel (if entry costs are sufficiently large). Using Porter's weekly data set on the Joint Executive Committee (JEC) from 1880 to the 16th week of 1886, the chapter does find empirical support for the hypotheses that entry increases the probability of observing a breakdown of the cartel, and that it increases the length of the punishment period. Interestingly, Porter (1985) found support for the latter but not for the former.

Chapter two contributes to the analysis of the nature of the difficulty for collusion when firms differ in "size", depending on the underlying reason which explains the differences in firms' sizes. In particular, the chapter analyses the conditions under which an industry-wide collusive outcome can be supported when an infinitely repeated game is played between asymmetric quantity setting firms which produce a homogeneous product. Following Perry and Porter (1985), asymmetries are dealt with by assuming that firms can have a different share of a specific asset (say, capital) which affects marginal costs. In this context, a firm is considered "large" if it owns a large fraction of the capital stock, and "small" if it owns only a restricted proportion of the capital available in the industry. The model assumes that firms use optimal punishment strategies with a *stick and carrot* structure in the style of the ones which have been characterized in general by Abreu (1986, 1988). In particular, Abreu's work is extended to consider a class of "proportional penal codes" (that is, along the punishment path firms produce in proportion to their assets).

The chapter's main results can be summarized as follows. First, it is shown that firms' incentives to disrupt the collusive agreement crucially depend on the distribution of assets amongst firms involved in the agreement. Second, from the analysis of the im-

impact of changes in the distribution of asset holdings (due to mergers, transfers or split-offs) on the sustainability of tacit collusion, it is found that if a merger (or other asset transfer) induces a more even distribution of assets, this tends to foster collusion. These results embody interesting implications for practical application of competition policy. In particular, the conclusions of this chapter shed some light on the analysis of the complex problem of assessing how a merger would affect collusion possibilities, an issue widely debated in today's competition policy (under the name of *joint dominance*, which refers to the possibility that firms reach a collusive outcome after a merger). The analysis clearly suggests that asymmetries in cost functions should be taken into consideration when predictions are made regarding the facility of collusion after an asset transfer takes place. In addition, our results confirm that, as initially stressed by Compte Jenny and Rey (1997) and more recently by Kühn and Motta (1999), a systematic analysis of market shares and concentration indexes does not always provide a reliable guide to evaluate potential effects on the level of competition in the market induced by an horizontal merger.

Chapter three aims at providing empirically testable implications regarding the relationship between market size and concentration in *endogenous sunk cost industries* (Sutton (1991, 1998)), that is, industries where firms are involved in research and development (henceforth, R&D) activities with the aim of enhancing the perceived quality of their products. Sutton (1991, 1998) has shown that in industries of this type very fragmented outcomes cannot arise as equilibrium outcomes in large markets. He shows that, under very general conditions, a lower bound to concentration exists and is bounded away from zero, no matter how large the market is. However, as pointed out by Bresnahan (1992) and

Scherer (2000), an important question left open by Sutton's analysis is whether an upper bound to the degree of concentration can be characterized in this type of industries. Chapter three addresses this question by extending the linear-demand model with horizontal product differentiation proposed by Sutton (1998). The chapter incorporates a post-entry additional stage into Sutton's framework, where firms may endogenously form coalitions. By forming coalitions (merging), firms cooperate and eliminate duplication efforts in R&D activities to enhance product quality. Hence, apart from reducing competition in the market, a merger allows firms to realize a cost advantage over the unmerged rivals.

A novel feature of this chapter is that it employs a coalitional stability concept which assumes that firms are endowed with foresight, in the sense that when making merger decisions, they look ahead and anticipate the ultimate outcome of their actions. This chapter is, therefore, related to a relatively new strand of the literature on farsighted stability (see, for instance, Chwe (1994), Xue (1998) and Diamantoudi and Xue (2001)).

The analysis leads to the following two main conclusions. First, independently of the size of the market, arbitrarily concentrated outcomes can arise in equilibrium. Therefore, this chapter shows that in *endogenous sunk cost industries*, an upper bound to concentration exists and is independent of the size of the market. Interestingly, in some equilibria in which a merger to monopoly is the unique equilibrium outcome, it turns out that firms belonging to the monopoly 'grand' coalition earn strictly positive profits even under the threat of entry. Second, it is also shown that if products are sufficiently good substitutes (or, if investment in R&D is sufficiently effective), duopoly coalition structures can only arise in equilibrium

if composed of sufficiently size asymmetric coalitions. The results, therefore, complement those of Sutton (1991, 1998).

Finally, as a practical remark, it should be stressed that all chapters of this thesis are self-contained and, therefore, can be read independently from each other as an independent article. This has the advantage that readers who are interested in only a part of the work presented here can gain easy access to their point of interest.

Chapter 1

Entry Effects on Cartel Stability and the Joint Executive Committee

1.1 Introduction

The objective of this chapter is twofold. From a theoretical perspective, the chapter contributes to the analysis of entry effects on cartel stability under demand uncertainty. From a more applied perspective, we empirically test some theoretical predictions about how entry can affect the pattern of collusive behavior of a group of firms organized in a cartel agreement to coordinate prices, making use of data on the US railroad cartel of the turn-of-the century.

We develop an extended version of the model proposed by Green and Porter (1984). In their seminal article, Green and Porter analyze infinitely repeated oligopoly games where market demand is subject to exogenous shocks and the firm's (past) actions are not observable, but they do not consider the possibility of entry. We thus reexamine their model to understand how the stability of the collusive price structure can be influenced by an increase in the number of firms in the agreement or by the existence of a pool of potential competitors.

The framework we use in order to explicitly model the entry process is related to Harrington (1989). We study two types of collusive equilibria in a repeated Bertrand game (with random demand and unobservable prices) between a set of active firms and a set of

potential competitors that can enter the market by paying a one-time, fixed, sunk cost (entry cost). In a collusive equilibrium, the initially active firms have no incentive to cut prices in non-reversionary periods (because this would trigger a “price war”) and the potential entrants have no incentive to enter, because the present value of the expected profits from entering the cartel is not sufficient to cover the entry cost. In the first type of equilibrium it is expected that entry would trigger a price war. In the second type of equilibrium it is expected that entry would be accommodated with a more inclusive agreement. It should be noted, however, that a major difference exists between the model developed in this chapter and Harrington’s (1989) framework. While Harrington’s results are obtained in a context of almost perfect information,¹ our analysis restricts the information available to firms, in the sense that at each period of time, apart from past entry decisions a firm knows only its own past prices and output levels. We find that entry does reduce the scope of collusion under both types of equilibria. In addition, in contrast with what the previous literature has pointed out with respect to the experience of the US railroad cartel, it is shown that, from an *ex-ante* point of view, the existence of a pool of competitors is a more important constraint on the maintenance of a stable agreement when a potential entrant expects to be accommodated by incumbents if entry occurs. We then look for empirical evidence on the role played by entry using the data set from the experience of the Joint Executive Committee (henceforth JEC), a pre-Sherman Act (legal) cartel.

The Green and Porter (1984) model has been subject to previous empirical tests using the JEC data set. The model suggests that in industries working in a context of imper-

¹ In the model he presents there are simultaneous moves at each stage of the dynamic game.

fect observability, price patterns include shifts between collusive regimes and competitive regimes along the collusive equilibrium path. Porter (1983b) and Ellison (1994), among others, demonstrated the existence of such regime shifts while examining this railroad cartel.² In another paper, which is probably the closest to the empirical application developed in this study, Porter (1985), still using the JEC data set, analyzes empirically the determinants of both frequency and duration of competitive reversions of finite length. The main goal of the econometric work developed by Porter was to determine whether the JEC was less successful in maintaining cooperative prices because of entry of new firms. To this end, he ran regressions for both the full sample and two subsets of the overall sample in which there were structural changes due to entry occurrence, to test for the impact of entry on the likelihood of a price war beginning. As noted by Porter himself, using the incidence of competitive episodes reported by the press at the time as the dependent variable, the results obtained are quite discouraging. Using the same data set but considering a different specification of the econometric model and creating new variables to explore what causes price wars to occur, our findings confirm the predictions of the reexamined version of the Green and Porter (1984) model presented in this chapter. In particular, unlike Porter, we find that a larger number of firms in the industry increases the percentage of periods spent in a price war. Two different forces justify this result. First, the higher number of firms in the agreement is, the higher the (one-shot) incentives to disrupt the collusive agreement are. Second, as is shown theoretically and contrary to what Porter (1985) claims,³ the opti-

² However, while Porter (1983b) obtained the result that firms' price cost mark-ups were consistent with a Cournot behavior, Ellison (1994), allowing for serial correlation in the demand between periods, came to the conclusion that cartel members were setting prices collusively between price wars.

³ See Porter (1985, p. 419).

mal punishment length is always increasing with the number of firms in the basic version of the model and increases as well with the number of active firms when entry is explicitly modeled, as long as the entry (sunk) costs are sufficiently high.

The chapter is organized as follows. Section 1.2 includes a brief description of the structural features of the basic theoretical model and some preliminary predictions about entry effects on cartel stability. The structure presented is based entirely upon the model developed in Green and Porter (1984). Departures from their assumptions are noted below. In Section 1.3 we develop the central analytical argument, extending the analysis of the preceding section to consider the possibility of entry of new firms. Section 1.4 briefly reviews the operations of the JEC. A description of the data used in the empirical application and the estimation results are provided in sections 1.5 and 1.6, respectively. Section 1.7 contains some concluding comments.

1.2 The basic model

Following Tirole (1988), we will develop a model in the spirit of Green and Porter (1984) in which a cartel is sustained by oligopolistic firms acting noncooperatively in a context of demand uncertainty.⁴ In this section we briefly remind the reader of the main features of this well-known model for the case where entry is not allowed. In the following Section 1.3, we extend the model to analyze entry.

⁴ We depart from Tirole's approach by considering that there are n firms in the agreement rather than two and also by allowing for a wider range of possible prices along the collusive path.

There exist n firms producing a homogeneous good and facing the same unit cost c . Firms choose prices in every period. Demand fluctuates randomly and its realizations are assumed to be independent and identically distributed (i.i.d.) over time. In each period there are two possible states of nature. With probability α the demand is zero ("low-demand state") and there is a positive demand with probability $(1 - \alpha)$ (the "high-demand state"). In the latter case, demand is split into equal parts corresponding to those firms charging the lowest price.

Firms do not observe their rivals' prices. Thus, from the point of view of each single firm in a cartel, a low demand for its product may be due to either secret price cutting by some competitors or a bad market demand shock.

A strategy, that is, a contingent plan of action, for a firm i in the repeated game is an infinite sequence $S_i = (S_i^0, S_i^1, \dots, S_i^t, \dots)$, where $S_i^0 \in R_+$ is a determinate initial price level, and $S_i^t : (R_+^2)^t \rightarrow R_+$ is a function that maps the prices charged and quantities faced by firm i in periods $1, 2, \dots, t-1$ into a price p_i^t , for firm i in period t . In this game, we look for a Nash equilibrium with the strategy for the i -th player defined in the following way:

$$S_i^0 = p^*, \quad (1.1)$$

$$S_i^t((q_i^0, p_i^0), \dots, (q_i^{t-1}, p_i^{t-1})) = \begin{cases} p^* & \text{if } q_i^{t-1} > 0, p_i^{t-1} = p^*, \text{ or} \\ & \forall \tau \in [t-T, t-1], p_i^\tau = p_c, q_i^\tau > 0. \\ p_c & \text{otherwise} \end{cases} \quad (1.2)$$

where $t = 1, 2, \dots$, $((q_i^0, p_i^0), \dots, (q_i^{t-1}, p_i^{t-1}))$ is the partial history with length t observed by firm i , $p^* \in (p_c, p_m]$ is the collusive price, while p_c and p_m represent the competitive and the monopoly price, respectively.

In words, the game lasts forever and initially each firm charges the collusive price. As soon as one of the n firms in the cartel observes a zero demand (and, therefore, earns zero profits), a punishment phase of T periods is triggered in which every firm adopts a Bertrand behavior.⁵ At the end of the reversionary episode all firms return to collusive behavior and share the collusive profit (Π^*) until a zero demand is again observed by some (or all) firms.⁶ Note that the length of the optimal punishment period can be neither zero nor infinite.

Let V_n^+ represent the present discounted value of a firm's profit from date t on, assuming that date t belongs to a collusive phase. Analogously, let V_n^- denote the present discounted value of a firm's profit from date t on, assuming that date t is the beginning of a punishment period.⁷ In this context we have:

$$V_n^+ = (1 - \alpha) \left(\frac{\Pi^*}{n} + \delta V_n^+ \right) + \alpha (\delta V_n^-) \quad (1.3)$$

and,

$$V_n^- = \delta^T V_n^+. \quad (1.4)$$

Equation (1.3) says that with probability $(1 - \alpha)$ the demand state is high, each firm earns its share of the collusive profit and the game remains in the collusive phase, and thus each firm has the valuation V_n^+ . However, with probability α , the "low-demand state" is achieved and in the next period a punishment phase starts. Equation (1.4) gives the present

⁵ As was pointed out by Fudenberg and Tirole (1991), "no player can gain by deviating in the punishment phase, since play there is a fixed number of repetitions of a static equilibrium." (p. 186).

⁶ In the light of Abreu *et al.* (1986), we know that it is possible that a global optimum might not be achieved with Nash reversion. However, since the competitive price Nash equilibrium achieves the minmax payoff profile, there is no room for a stronger punishment than Nash reversion in this setting.

⁷ Because of stationarity, neither V_n^+ and V_n^- depend on time. In addition, the subscripts denote the number of firms in the agreement.

discounted value of the expected stream of profits at the beginning of the punishment phase phase.

By solving the system of equations (1.3) and (1.4) one obtains:

$$V_n^+ = \frac{1}{1 - (1 - \alpha)\delta - \alpha\delta^{1+T}} (1 - \alpha) \frac{\Pi^*}{n}, \quad (1.5)$$

$$V_n^- = \frac{\delta^T}{1 - (1 - \alpha)\delta - \alpha\delta^{1+T}} (1 - \alpha) \frac{\Pi^*}{n}. \quad (1.6)$$

As we have seen, cartel members' strategies (1.1) and (1.2) prescribe a mechanism to punish deviations from the price structure agreed upon. Whether this punishment mechanism is self-enforcing will depend on the trade-off between potential short-run gains from deviation and the present value of expected future losses. This trade-off is captured by the analysis of the incentive compatibility constraint:

$$V_n^+ \geq (1 - \alpha) (\Pi^* + \delta V_n^-) + \alpha (\delta V_n^-), \quad (1.7)$$

where $\delta \in (0, 1)$ is the common discount factor. The right-hand side of the inequality shows that if the demand is high and the firm decides to undercut, it earns all the one-shot collusive profit, but by deviating it will trigger a punishment reversal in the following T periods. If the demand is zero, it will earn zero profits in the current period and a punishment phase starts in the next period.

Using equation (1.3) we can rewrite the incentive compatibility constraint in the following way:

$$\delta (V_n^+ - V_n^-) \geq \Pi^* - \frac{\Pi^*}{n}. \quad (1.8)$$

This relation says that a prospective deviant will decide to respect the agreement if the present discounted value of the long-run net gain from collusion is greater than the short-run gain from deviation.

The next proposition derives the optimal length of the punishment period T^* as a function of the parameters in the model and shows that it is increasing in the number of firms belonging to the agreement. The intuition which underlies this result is simple. On the one hand, when more firms belong to the collusive agreement, the gain from secret price cutting becomes larger as the deviating firm can obtain more share. On the other hand, each firm's continuation payoff in the collusive equilibrium decreases, which reduces the impact of any finite-period punishment. Combining these two effects, one concludes that the optimal length of the punishment has to increase in order to maintain the incentive for the firms to stick to the collusive price.

Proposition 1 *If the number of firms is sufficiently low, i.e., $n < 1/(1 - \delta(1 - \alpha))$, the optimal punishment length T^* is determined by*

$$T^* = \frac{1}{\ln \delta} \ln \frac{n(1 - \delta(1 - \alpha)) - 1}{\delta(\alpha n - 1)},$$

where T^* is an increasing function of the number of firms in the agreement.

Proof. Substituting (1.5) and (1.6) into (1.8), some algebra shows that, in order for the incentive compatibility constraint to hold, one must have that:

$$n - 1 \leq \delta^{T+1}(\alpha n - 1) + \delta n(1 - \alpha), \quad (1.9)$$

where we assume $\alpha < 1/n$ in order for the r.h.s. of eq. (1.9) to increase with T . As Tirole (1988) and others have mentioned, the optimal punishment length T^* should be chosen to maximize the discounted joint profits and, therefore, should be the minimal T for which eq. (1.8) (or, equivalently, (1.9)) holds. Rewriting (1.9), one obtains:

$$\delta^T \leq \frac{n - 1 - \delta n (1 - \alpha)}{\delta (\alpha n - 1)}, \quad (1.10)$$

which in turn implies that

$$T \geq \frac{1}{\ln \delta} \ln \frac{n (1 - \delta (1 - \alpha)) - 1}{\delta (\alpha n - 1)} \equiv T^*, \quad (1.11)$$

where, since $\alpha < 1/n$, one must have that $n < 1 / (1 - \delta (1 - \alpha))$ in order for T^* to exist.

Now, taking the derivative of T^* with respect to the number of firms, one obtains:

$$\frac{\partial T^*}{\partial n} = - \frac{1}{\ln \delta} \frac{(1 - \alpha) (1 - \delta)}{(\alpha n - 1) (n (1 - \delta (1 - \alpha)) - 1)}, \quad (1.12)$$

which is always positive given that, as already mentioned, $\alpha < 1/n$ and $n < 1 / (1 - \delta (1 - \alpha))$.

This completes the proof. ■

First, it should be stressed that an optimal punishment length only exists if the number of firms in the agreement is sufficiently low. Otherwise, the cartel is never (internally) stable. Second, and most importantly, notice that this result is in sharp contrast with what Porter (1985) indicates as a possibility, i.e., that as the number of firms increases, "the cartel may want to employ shorter, and so more forgiving, punishments when reversions occur, in order to partially offset the increased fraction of time spent at competitive prices." (p. 419). This result will play a central role in the explanation of the main hypothesis tested in the empirical section of this chapter, namely that cartel stability is negatively correlated with the number of firms in the cartel agreement.

1.3 Considering entry

We will now describe a new version of the Green and Porter (1984) model in which the entry process is modeled explicitly. We will study a particular subset of the set of Perfect Bayesian Equilibria. In particular, we will assume that once entry occurs, incumbent firms can either engage in aggressive behavior for a while (Case 1) or accommodate the new entrants (Case 2).⁸ Throughout the analysis of each of these cases, we will first give the necessary and sufficient conditions for a non-trivial degree of collusion to be sustainable and characterize the optimal length of the punishment period which is required by the equilibrium conditions. We then explore the stability issue in further detail, by analyzing the determinants of the proportion of periods spent by the cartel in price wars.

1.3.1 The new model structure

Assume a countably infinite number of active and inactive firms represented by the set Z . $A^t \subset Z$, where the inclusion is *strict*, denotes the set of active firms in period t and $|A^t| = N^t$ (where $|A^t|$ is the number of elements of A^t).⁹ Hence, $Z - A^t$ represents the set of potential entrants in each period t . Further, let both active and inactive firms have the same marginal costs of production.

⁸ We are not considering here the situation where incumbent firms coordinate their behavior to force an entrant to exit from the industry. During the operation of the JEC, it was common knowledge that new competitors faced a "no-exit constraint". As a consequence, "it would not be rational for a railroad cartel to engage in predatory pricing practices in response to entry." (Porter 1985, p. 420). Thus, we have decided not to cover predatory pricing here, since it is clear that the use of this kind of threat against potential entrants will not have been credibly carried out by the railroad firms belonging to the nineteenth-century railroad cartel, whose pricing strategies we will discuss in the applied part of this chapter.

⁹ Firm i is considered active at time t ($i \in A^t$) if it faces a positive demand in that period.

We also assume that, for a potential entrant, entry and price decisions are not simultaneous. At each period t , the N^t active firms simultaneously announce the price to charge (p_i^t denotes the i -th firm's price in period t). At the same time, potential entrants decide about entry. A one-time entry (sunk) cost K (where $K \geq 0$) has to be incurred if entry takes place. It allows the firm to begin production one period later. Hence, prices are a post-entry decision for a potential competitor and past entry decisions are assumed to be perfectly observed by all active firms.¹⁰

If a firm i is initially active ($i \in A^0$), its overall strategy - S_i - can be represented as an infinite sequence of action functions (one for each period) $S_i = (S_i^0, S_i^1, \dots, S_i^t, \dots)$, where $S_i^0 \in R_+$ represents the initial price charged by firm i and $S_i^t : (R_+^2 \times 2^Z)^t \rightarrow R_+$. The domain of an action function S_i^t is the Cartesian product between the set of feasible prices p_i^τ , the set of possible outputs¹¹ q_i^τ and the set of active firms in period τ , where $\tau \in \{0, \dots, t-1\}$. The range of S_i^t is represented by the set of possible prices that firm i can charge in period t . Hence, period t action function of an arbitrary active firm i tells it which price to set in the t -th period as a function of the feasible histories observed over the periods $\{0, \dots, t-1\}$.

More formally:

$$S_i^t((q_i^0, p_i^0, A^0), \dots, (q_i^{t-1}, p_i^{t-1}, A^{t-1})) = p_i^t,$$

where $((q_i^0, p_i^0, A^0), \dots, (q_i^{t-1}, p_i^{t-1}, A^{t-1}))$ is the partial history with length t observed by firm i , which is denoted by $h_i^t \in H_i^t$. Notice that h_i^t can be partitioned into the public partial

¹⁰ Entry is generally a time-consuming process; therefore, following Harrington (1989), we assume that incumbent firms are able to change their price decisions in response to entry.

¹¹ q_i^t is the demand faced by firm i in period t .

history (A^0, \dots, A^{t-1}) and the private partial history $((q_i^0, p_i^0), \dots, (q_i^{t-1}, p_i^{t-1}))$ observed by firm i .

If firm i is instead initially inactive ($i \in (Z - A^0)$), then its overall strategy is $E_i = (E_i^0, E_i^1, \dots, E_i^t, \dots)$. In the first period of the game ($t = 0$), $E_i^0 \in \{\text{Out}, \text{In}\} \times \{\infty\}$, where Out means “Do Not Enter the Market” and In means “Enter the Market”. Moreover, with respect to the price decision, the set of feasible prices is a singleton.¹² For each period $t \in \{1, 2, \dots\}$, there exists an action function E_i^t , which, given the history observed up to period $t - 1$, tells the firm whether or not to enter at the beginning of period $(t + 1)$. We consider entry as an irreversible decision. Thus, if entry occurs, firm i 's strategy specifies the prices to be charged for the remainder of the horizon. The observed (partial) history is composed by the own price and entry decisions and by the demand faced by the firm up to the previous period. In formal terms, $E_i^t : \tilde{H}_i^t \rightarrow \{\text{Out}, \text{In}\} \times \{R_+ \cup \{\infty\}\}$, where $\tilde{H}_i^t \subseteq (\{\text{Out}, \text{In}\} \times R_+^2)$. Moreover,

$$\forall \tilde{h}_i^t \in \tilde{H}_i^t, \text{ if } \tilde{h}_i^t = (\dots, (x_i^\tau, p_i^\tau, q_i^\tau), \dots), \text{ where } x_i^\tau = \text{Out}, \text{ then } q_i^{\tau+1} = 0.$$

It should be also noted that:

$$E_i^t((\text{Out}, \infty, 0), \dots, (\text{Out}, \infty, 0)) \in \{\text{Out}, \text{In}\} \times \{\infty\},$$

while

$$E_i^t((\text{Out}, \infty, 0), \dots, (\text{In}, \infty, 0)) \in \{\text{In}\} \times R_+.$$

¹² The only feasible price is $p_i^0 = \infty$, thus the firm will not face a positive demand in the first period of the game.

In the discussion that follows, we will identify, for each of the considered cases of post-entry reaction, the conditions which should be satisfied in order for the incumbent firms to sustain a non-trivial degree of market power without giving rise to either internal or external defection. Following Harrington (1991), we consider that internal defection takes place when some active firm secretly undercuts the price while external defection occurs if some potential competitor decides to enter the market.¹³

1.3.2 Case 1. Cartel Breakdown

In this case we consider a situation in which incumbent firms respond to entry by reverting to competitive pricing for a finite length of time (T periods). Since we want to understand how incumbent firms effectively sustain collusion, avoiding not only internal deviations for profit, but also entry of new competitors, an active firm's strategy is designed in the following way:

$$S_i^0 = p^*, \quad (1.13)$$

$$S_i^t(h_i^t) = \begin{cases} p^* & \text{if } q_i^{t-1} > 0, p_i^{t-1} = p^* \text{ and } A^t = A^{t-1}, \text{ or} \\ & \forall \tau \in [t-T, t-1], p_i^\tau = p_c, q_i^\tau > 0 \text{ and } A^\tau = A^{\tau-1}. \\ p_c & \text{otherwise} \end{cases} \quad (1.14)$$

where $t = 1, 2, \dots, i \in A^0, N^0 = |A^0| = n$.

Hence, the incumbent firms charge the collusive price (p^*) at the beginning of the game and continue to set it as long as no firm faces a zero demand and no entry occurs. However, if some firm (active or inactive) has deviated from the proposed path or if a

¹³ Note that the basic framework of Section 1.2 can be viewed as a special case of the extended model we present in this section, where the entry sunk cost is prohibitive ($K \rightarrow \infty$) and, therefore, external stability is not an issue.

low-demand state has occurred, then the collusive price is reestablished after a (temporary) punishment phase.

Our aim is to find a trigger strategy equilibrium such that strategies (1.13), (1.14) are optimal and there is no incentive for the entry of new firms into the industry. To this end, consider the following strategy for a potential entrant:

$$E_i^0 = \text{Out} \quad (1.15)$$

$$E_i^t \left(\tilde{h}_i^t \right) = \begin{cases} \text{Out} & \text{if } i \in Z - A^t \\ S_i^t & \text{if } i \in A^t \end{cases}, \quad (1.16)$$

where $t = 1, 2, \dots, i \in Z - A^0$.

This strategy calls for a potential entrant not to enter the industry. However, if the firm decides to enter, then the strategy also prescribes the price conduct which it should follow after its entry.¹⁴ Remember that the incumbents will react to entry by setting a price p_c during a finite punishment period. Hence, once inside the market, the best response of the entrant is to set p_c during the punishment period which is triggered by its entry, since the Bertrand solution constitutes a Nash equilibrium for the one-shot price game which is played in each single-period, given the number of active firms in the industry.

A free entry trigger strategy equilibrium is defined as a triple (p^*, \tilde{T}_1, n^*) such that the strategies in (1.13)-(1.16) form a Perfect Bayesian Equilibrium (PBE). The following two conditions are necessary and sufficient for (1.13)-(1.16) to form a PBE:¹⁵

$$\delta (V_{n^*}^+ - V_{n^*}^-) \geq \Pi^* - \frac{\Pi^*}{n^*} \quad (1.17)$$

¹⁴ It should follow the strategy of an initially active firm (strategy (1.14)) for the remainder of the horizon.

¹⁵ Notice that subgame perfection cannot be used as the equilibrium concept. This is a dynamic game with unobservable actions in which the only proper subgame is the whole game itself. Moreover, active firms are able to observe actions which are off the equilibrium path (remember that they know the set of active firms in the past periods).

and,

$$\delta V_{n^*+1}^- - K \leq 0. \quad (1.18)$$

Notice that not only the collusive price, but also the length of the punishment are chosen (optimally) by the cartel to maximize the expected discounted joint profit subject to the constraints (1.17) and (1.18). As can be easily verified from equation (1.5), V_n^+ is a decreasing function of T .

As in Section 2, expression (1.17) represents the incentive compatibility constraint. In Proposition 1, it has been shown that this constraint holds as long as the punishment length is large enough. In particular, if $T \geq T^*$, where T^* is given by eq. (1.11), then each of the n^* active firms finds it optimal to go along with the collusive path and to charge the collusive price p^* since the discounted loss from cheating is greater than the one-shot gain from deviation. On the other hand, condition (1.18) makes further entry into the industry unprofitable. Thus, when this latter condition holds, existing profits might be positive because profits with further entry¹⁶ are expected to be negative, which means that "Out" is, in fact, a best response for an inactive firm.¹⁷

In what follows, the optimal punishment length in a cartel breakdown scenario is derived and characterized.

Proposition 2 *Let T^* be as in Proposition 1. In a scenario in which incumbent firms respond to entry by reverting to a price war during a finite length of time, and assuming*

¹⁶ Defined net of the (sunk) costs of entry.

¹⁷ The condition in (1.18) can be interpreted as a violation of the participation constraint corresponding to the pool of potential competitors.

$0 < K < (\Pi^*/(n^* + 1))(1/(1 - \delta))$, the optimal punishment length \tilde{T}_1 is given by $\tilde{T}_1 = \max\{T^*, T_{pc1}\}$, where

$$T_{pc1} = \frac{1}{\ln \delta} \ln \frac{K(n^* + 1 - \delta(1 - \alpha)(n^* + 1))}{\delta((1 - \alpha)\Pi^* + K\alpha(n^* + 1))}.$$

Proof. From Proposition 1, we know that the incentive compatibility constraint (1.17) holds if and only if $T \geq T^*$, where T^* is defined by eq. (1.11). Now, using eq. (1.6), it is straightforward to show that:¹⁸

$$V_{n^*+1}^- = \frac{\delta^{T'}}{1 - (1 - \alpha)\delta - \alpha\delta^{1+T'}} (1 - \alpha) \frac{\Pi^*}{n^* + 1}. \quad (1.19)$$

Hence, substituting (1.19) into eq. (1.18), some algebra shows that incumbent firms will be able to coordinate their price strategies without giving rise to entry of new competitors if and only if:

$$\delta^{T'+1} \leq \frac{K(n^* + 1)(1 - (1 - \alpha)\delta)}{(1 - \alpha)\Pi^* + K\alpha(n^* + 1)}. \quad (1.20)$$

The l.h.s. of condition (1.20) is easily seen to be decreasing in T' . Notice also that when $K \geq \frac{\Pi^*}{n^*+1} \left(\frac{1}{1-\delta}\right)$,¹⁹ the r.h.s. of condition (1.20) is greater than or equal to one and, therefore, condition (1.20) is trivially satisfied. Let us, therefore, consider in what follows the case in which $K < \frac{\Pi^*}{n^*+1} \left(\frac{1}{1-\delta}\right)$. If this is the case, some algebra shows that condition (1.18) holds if and only if

$$T' \geq \frac{1}{\ln \delta} \ln \frac{K(n^* + 1 - \delta(1 - \alpha)(n^* + 1))}{\delta((1 - \alpha)\Pi^* + K\alpha(n^* + 1))} \equiv T_{pc1}. \quad (1.21)$$

¹⁸ We here obviously assume that the length of a price war depends on the number of active firms in the agreement. Hence, given that $V_{n^*+1}^-$ is computed for a case in which there are $n^* + 1$ firms in the agreement, notice that we are now considering a general value T' (different from the one used in $V_{n^*}^-, T$).

¹⁹ When this is the case, entry (sunk) costs are extremely high (say, prohibitive) and, therefore, external stability is not an issue.

Hence, the minimum length of the punishment period for which both the incentive and the participation constraints (eqs. (1.17) and (1.18)) hold is given by $\tilde{T}_1 = \max \{T^*, T_{pc1}\}$. ■

Two important remarks are in order at this point. First, using (1.21) it can be easily shown that

$$\frac{\partial T_{pc1}}{\partial n^*} = \frac{1}{\ln \delta (n^* + 1)} \frac{\Pi^* (1 - \alpha)}{((1 - \alpha) \Pi^* + K \alpha (n^* + 1))} < 0, \quad (1.22)$$

and

$$\frac{\partial T_{pc1}}{\partial K} = \frac{1}{\ln \delta (\Pi^* (1 - \alpha) + K \alpha (n^* + 1))} < 0. \quad (1.23)$$

It should be stressed at this point that the *initial* number of firms n^* is a parameter of the model affecting the viability of collusive agreements. In equilibrium, the number of firms does not change, and so it makes sense to do comparative statics with respect to n^* . Let us now turn to the interpretation of the previous results. From (1.22), it can be concluded that the greater the number of firms in the agreement or the higher the height of the entry barriers is, the lower will be the length of the punishment period which must exist in order to avoid entry and allow incumbents to sustain a non-trivial degree of cooperation. Second, eq. (1.23) calls attention to the fact that for sufficiently high values of the entry (sunk) cost, condition (1.18) is non-binding. In particular, and as already mentioned in the proof of Proposition 2, if $K \geq \Pi / (n + 1) (1 / (1 - \delta))$,²⁰ then (1.18) is trivially satisfied. In the next Lemma, however, tighter predictions are given regarding the ranking between the two relevant thresholds for the punishment length, T^* and T_{pc1} .

²⁰ That is, the value of the entry cost is higher than or equal to the present value of individual shares in the collusive profits in case no punishment is triggered in the future periods.

Lemma 1 *Let \tilde{T}_1 , T^* and T_{pc1} be as in Proposition 2. $\tilde{T}_1 = T^*$ if and only if $K \geq K_1$, where*

$$K_1 = \frac{\Pi^* (1 - n^* (1 - \delta (1 - \alpha)))}{n^* + 1 (1 - \delta)}.$$

Proof. Making use of (1.11) and (1.21), some algebra shows that $T^* \geq T_{pc1}$ if and only if

$$\frac{(1 - n^* (1 - \delta (1 - \alpha))) ((1 - \alpha) \Pi^* + K \alpha (n^* + 1))}{(1 - \alpha n^*) K (n^* + 1) (1 - \delta + \delta \alpha)} \leq 1, \quad (1.24)$$

which in turn implies that

$$K \geq \frac{\Pi^* (1 - n^* (1 - \delta (1 - \alpha)))}{n^* + 1 (1 - \delta)} \equiv K_1. \quad (1.25)$$

Hence, when $K \geq K_1$, $\tilde{T}_1 \equiv \max \{T^*, T_{pc1}\} = T^*$. ■

In words, Lemma 1 shows that for sufficiently high values of the entry (sunk) cost K , the incentive compatibility constraint is the one that is binding and, therefore, the optimal length of the punishment is given by the minimum value of T for which this incentive constraint is satisfied, $\tilde{T}_1 = T^*$. As the next Corollary shows, this result is important to understand what is, in this case of the extended version of the model, the impact of entry on the duration of the optimal length of the punishment phase.

Corollary 1 *Let K_1 be as in Lemma 1. The optimal length of the punishment in a cartel breakdown scenario, \tilde{T}_1 , increases with the number of firms if and only if $K \geq K_1$.*

Proof. If $K_1 \leq K < (\Pi^*/(n^* + 1)) (1/(1 - \delta))$, then applying Proposition 2 and Lemma 1, one concludes that $\tilde{T}_1 = T^*$. Now, from (1.12) we know that $\frac{\partial T^*}{\partial n^*} > 0$. If, instead, $K \geq (\Pi^*/(n^* + 1)) (1/(1 - \delta))$, then entry costs are prohibitive, the incentive

compatibility constraint is the one which is binding and, as already mentioned, from (1.12) we know that $\frac{\partial T^*}{\partial n^*} > 0$. Lastly, if $0 < K < K_1$, then from Lemma 1, one has that $\tilde{T}_1 = T_{pc1}$ and from (1.22) we know that $\frac{\partial T_{pc1}}{\partial n^*} < 0$. ■

So, even when potential entrants are threatened with a subsequent (temporary) cartel breakdown if entry takes place, strictly positive entry sunk costs must exist in order to discourage entry into the industry. In addition, if sunk entry costs are sufficiently high, then the optimal punishment length chosen by cartel members is increasing in the number of firms belonging to the collusive agreement.

1.3.3 Case 2. Accommodation

Rather than following the policy of reacting to entry by starting a finite punishment period, cartel members might decide to accommodate the entrant by achieving a new collusive outcome.²¹ If a potential entrant anticipates that, entering at period t , the N^t active firms will adopt post-entry accommodation behavior, then the necessary and sufficient conditions for a triple (p^*, \tilde{T}_2, n^*) to constitute a free entry trigger strategy equilibrium are now given by:

$$\delta (V_{n^*}^+ - V_{n^*}^-) \geq \Pi^* - \frac{\Pi^*}{n^*} \quad (1.26)$$

and,

$$\delta V_{n^*+1}^+ - K \leq 0. \quad (1.27)$$

²¹ Wenders (1971) calls this policy a "price maintenance strategy".

Whenever condition (1.26) holds, an individual active firm predicts that it will make more profit by being loyal to the cartel agreement than by being disloyal; therefore, the agreement is unlikely to breakdown because of an internal deviation.²²

Condition (1.27) specifies that, even though incumbents will allow new competitors to join the collusive process, potential entrants anticipate a negative post-entry profit, which means that "Out" is an optimal choice for them.²³

As in the previous section, we can now derive the optimal punishment length in a context where entering firms are accommodated by incumbents in a more inclusive agreement.

Proposition 3 *Let T^* be as in Proposition 1. In a scenario in which incumbent firms respond to entry by accommodating the new entrants, and assuming $(\Pi^*/(n^* + 1))(\delta(1 - \alpha)/(1 - \delta(1 - \alpha))) < K < (\Pi^*/(n^* + 1))(\delta/(1 - \delta))$, the optimal punishment length \tilde{T}_2 is given by $\tilde{T}_2 = \max\{T^*, T_{pc2}\}$, where*

$$T_{pc2} = \frac{1}{\ln \delta} \ln \frac{K(n^* + 1)(1 - \delta(1 - \alpha)) - \delta\Pi^*(1 - \alpha)}{K\delta\alpha(n^* + 1)}.$$

Proof. From Proposition 1, we know that the incentive compatibility constraint (1.26) holds if and only if $T \geq T^*$, where T^* is defined by eq. (1.11). Now, from (1.5) one can

²² It should be noted that although incumbent firms follow a policy of post-entry accommodation, we are still assuming that internal defection is followed by a finite period of cartel breakdown. Hence, conditions (1.17) and (1.26) coincide.

²³ Again, both p^* and \tilde{T}_2 are chosen (by the cartel) to maximize the discounted profits, but now subject to constraints (1.26) and (1.27).

easily conclude that

$$V_{n^*+1}^+ = \frac{1}{1 - (1 - \alpha)\delta - \alpha\delta^{1+T''}} (1 - \alpha) \frac{\Pi^*}{n^* + 1}. \quad (1.28)$$

Substituting now (1.28) into eq. (1.27), some algebra shows that condition (1.27) is satisfied if and only if:

$$\delta^{1+T''} \leq \frac{K(n^* + 1)(1 - (1 - \alpha)\delta) - \delta(1 - \alpha)\Pi^*}{K(n^* + 1)\alpha}, \quad (1.29)$$

where we have to assume two restrictions regarding values which the entry sunk cost can take. First, we assume that $K > \frac{\Pi^*}{n^*+1} \frac{\delta(1-\alpha)}{1-\delta(1-\alpha)}$ in order for the r.h.s. of eq. (1.29) to be positive. Second, we suppose $K < \frac{\Pi^*}{n^*+1} \left(\frac{\delta}{1-\delta}\right)$ to rule out the case in which the r.h.s. of eq. (1.29) is greater than or equal to one and, therefore, the previous condition is trivially satisfied. Now, since the l.h.s. of eq. (1.29) is easily seen to decrease in T'' , some algebra shows that condition (1.29) (and, hence, (1.27)) holds if and only if:

$$T'' \geq \frac{1}{\ln \delta} \ln \frac{K(n^* + 1)(1 - \delta(1 - \alpha)) - \delta\Pi^*(1 - \alpha)}{K\delta\alpha(n^* + 1)} \equiv T_{pc2}. \quad (1.30)$$

Hence, the minimum length of the punishment period for which both the incentive and the participation constraints (eqs. (1.26) and (1.27)) hold is given by $\tilde{T}_2 = \max\{T^*, T_{pc2}\}$. ■

Notice that now, contrary to what has been found for the cartel breakdown scenario, in order for T_{pc2} to exist, the entry (sunk) cost has to be sufficiently high (and not only positive).

As in Case 1, the analysis will now focus upon the sensitivity of the critical level for the punishment length for which the participation constraint (1.27) holds, T_{pc2} , to changes in the number of firms in the agreement or in the level of the entry cost. Making use of

(1.30), it is straightforward to show that:²⁴

$$\frac{\partial T_{pc2}}{\partial n^*} = \frac{\Pi^* (1 - \alpha)}{(n^* + 1) (K (n^* + 1) (1 - \delta (1 - \alpha)) - \delta \Pi^* (1 - \alpha))} \frac{\delta}{\ln \delta} < 0, \quad (1.31)$$

and

$$\frac{\partial T_{pc2}}{\partial K} = \frac{\Pi^* (1 - \alpha)}{K (K (n^* + 1) (1 - \delta (1 - \alpha)) - \delta \Pi^* (1 - \alpha))} \frac{\delta}{\ln \delta} < 0. \quad (1.32)$$

Hence, as in Case 1, the higher the values of n^* and K , the lower the minimal length of the punishment period which is compatible with an anticipated unprofitable entry. As before, one can now show that, since T_{pc2} is decreasing in K , for sufficiently high levels of the entry sunk cost, condition (1.27) is a non-binding constraint. The next Lemma formalizes this result.

Lemma 2 *Let \tilde{T}_2 , T^* and T_{pc2} be as in Proposition 3. $\tilde{T}_2 = T^*$ if and only if $K \geq K_2$,*

where

$$K_2 = \frac{\Pi^*}{n^* + 1} \frac{\delta (1 - \alpha n^*)}{1 - \delta}.$$

Proof. From (1.11) and (1.30), some algebra shows that $T^* \geq T_{pc2}$ if and only if:

$$\frac{K \alpha (n^* + 1) (n^* - 1 - \delta n^* (1 - \alpha))}{(\alpha n^* - 1) (K (1 - (1 - \alpha) \delta) (n^* + 1) - \delta \Pi^* (1 - \alpha))} < 1, \quad (1.33)$$

which in turn implies that²⁵

$$K \geq \frac{\Pi^*}{n^* + 1} \frac{\delta (1 - \alpha n^*)}{1 - \delta} \equiv K_2. \quad (1.34)$$

Hence, if $K \geq K_2$, $\tilde{T}_2 \equiv \max \{T^*, T_{pc2}\} = T^*$. ■

²⁴ Remember that, as was mentioned in the proof of Proposition 3, we assume that $K > \frac{\Pi^*}{n^* + 1} \left(\frac{\delta(1-\alpha)}{1-\delta(1-\alpha)} \right)$. Otherwise, there is no (finite) optimal punishment length compatible with cartel external stability.

²⁵ Remember that $\alpha < 1/n^*$.

Corollary 2 *Let K_2 be as in Lemma 2. The optimal length of the punishment in an accommodation of the new entrants scenario, \tilde{T}_2 , increases with the number of firms if and only if $K \geq K_2$.*

Proof. If $K \geq K_2$, then applying Lemma 2, one has that $\tilde{T}_2 = T^*$. Then, from (1.12) we know that $\frac{\partial T^*}{\partial n^*} > 0$. If, instead, $\frac{\Pi^*}{n^*+1} \frac{\delta(1-\alpha)}{1-\delta(1-\alpha)} < K < K_2$, then by Proposition 3 and Lemma 2, one has that $\tilde{T}_2 = T_{pc2}$ and from (1.31) we know that $\frac{\partial T_{pc2}}{\partial n^*} < 0$. Lastly, if $K \leq \frac{\Pi^*}{n^*+1} \frac{\delta(1-\alpha)}{1-\delta(1-\alpha)}$, as already mentioned in the proof of Proposition 3, the cartel is not externally stable for any (finite) value of the punishment length. ■

Before closing this section, let us compare the two critical levels for the entry cost derived in Lemma 1 and 2. Using (1.25) and (1.34), simple algebra shows that:

$$K_2 - K_1 = \frac{\Pi^*}{n^* + 1} (n^* - 1) > 0. \quad (1.35)$$

This result stresses the fact that, for a given market structure (n^*), in order for the participation constraint to be a non-binding constraint, the height of the entry barriers should be higher in the second case (accommodation of the new entrants) than in the first one (cartel breakdown). In other words, the existence of a pool of potential competitors is a more important constraint on the maintenance of a stable agreement when cartel members decide to accommodate the entrants. The intuition behind this result is just that the anticipation of a tougher price competition in Case 1 makes entry less attractive. This finding is particularly relevant to us since, even though it is *ex-ante* optimal for cartel members to threaten entrants with a breaking up of the cartel if entry occurs,²⁶ accommodating the new railroad

²⁶ Notice, however, that we are not considering the possibility of "renegotiation" of equilibria. This is certainly an important further development of this study, since, as was stressed by Fudenberg and Tirole

firms by allocating them market shares was considered by some authors as the incumbents' best reply to entry during the operation of the JEC.²⁷

1.3.4 Stability discussion

The purpose of this section is to pin down what are, in the extended version of the model, the theoretical predictions regarding the primary question addressed in the empirical part of this chapter - understanding how entry of new firms affects the firms' collusive pricing behavior.

Let W_t be the indicator function that takes value 1 if a price war occurs at period t . In addition, let $\Pr(W_t = 1)$ denote the stationary probability of a price war occurrence in period t . In the next Proposition, the percentage of periods spent in a price war is derived and shown to be increasing in the punishment length.

Proposition 4 *The percentage of periods spent in a price war is given by*

$$\Pr(W_t = 1) = \frac{\alpha T}{1 + \alpha T},$$

where $\Pr(W_t = 1)$ increases with the length of the punishment phase T .

Proof. Within this framework, along the equilibrium path and whatever the assumed incumbents post-entry reaction, there is no cheating and/or entry of new firms. All price wars are induced by low demand shocks. The percentage of periods spent in a price war is

(1991), if "players have the opportunity to negotiate anew at the beginning of each period, then equilibria that enforce *good* outcomes by the threat that deviations will trigger a *punishment equilibrium* may be suspect, as a player might deviate and then propose abandoning the punishment equilibrium for another equilibrium in which all players are better off" (p. 175).

²⁷ As will be explained in further detail in Section 1.4.

just the stationary probability of being in a price war $\Pr(W_t = 1)$. If there is a price war at period $t - 1$, then the probability that $t - 1$ was not the last period of the price war is given by $(T - 1)/T$. $\Pr(W_t = 1 | W_{t-1} = 1) = (T - 1)/T$. If, instead, there is no price war at period $t - 1$, then the probability of a price war occurrence in the following period t is $\Pr(W_t = 1 | W_{t-1} = 0) = \alpha$. Therefore, the unconditional probability of being in a price war in period t is defined by

$$\Pr(W_t = 1) = \Pr(W_{t-1} = 1) \left(\frac{T-1}{T} \right) + \Pr(W_{t-1} = 0) \alpha. \quad (1.36)$$

Now, since $\Pr(W_t = 1) = \Pr(W_{t-1} = 1)$ as it is the stationary probability, one can rewrite eq. (1.36) as follows:

$$\Pr(W_t = 1) = \Pr(W_t = 1) \left(\frac{T-1}{T} \right) + [1 - \Pr(W_t = 1)] \alpha. \quad (1.37)$$

Solving eq. (1.37) with respect to $\Pr(W_t = 1)$, one obtains that:

$$\Pr(W_t = 1) = \frac{\alpha T}{1 + \alpha T}. \quad (1.38)$$

Carrying out now a simple exercise of comparative statics using (1.38), it is straightforward to show that:

$$\frac{\partial \Pr(W_t = 1)}{\partial T} = \frac{\alpha}{(1 + \alpha T)^2} > 0. \quad (1.39)$$

This completes the proof. ■

The results of the previous sections suggest that in the case the number of cartel members n^* varies, there is a conflict between having collusion immune to internal defection, on the one hand, and to the entry of new competitors, on the other. In particular, eqs. (1.12), (1.22) and (1.31) show that, whatever the incumbent's post-entry reaction is, an increase in n^* induces, on the one hand, an increase in the minimum length of the punishment T^*

for which the incentive compatibility constraint holds and a decrease in the critical threshold for which the participation constraint is satisfied (T_{pc1} , T_{pc2}), on the other. Whether, in the end, this change in the number of firms leads to an increase or to a decrease of the optimal length of the punishment phase (\tilde{T}_1, \tilde{T}_2) depends, as already shown, on the height of the actual entry sunk costs. In particular, Corollary 1 and 2 have shown that the punishment length increases with the number of firms if the actual level of the sunk entry cost is sufficiently high. If this is the case, then by Proposition 4 we also have that this increase in the optimal punishment length induced by an increase in the number of cartel members, will in turn lead to an increase in the percentage of periods the cartel spends in a price-war. As will be shown in the empirical part of the chapter, this seems to have occurred during the operation of the JEC cartel.

As a final remark, it should be stressed that as we are interested here in finding out the necessary and sufficient conditions for which the cartel is (internally and externally) stable, entry is, in this framework, treated as a disequilibrium phenomenon. Entry is never expected in equilibrium. When it occurs, the enlarged set of active firms coordinate on a *new* equilibrium taking for granted that entry will not occur any more and an anticompetitive conduct by the cartel members becomes more difficult to attain.²⁸ This is important for the interpretation of the empirical results: since entry is actually observed during the period under study, this means that the data cannot be fully explained by the equilibrium model. It has to be explained with unexpected changes in exogenous variables such as a decrease in the sunk costs of entering the railway business (or to the arrival of a lone potential

²⁸ That is, as entry occurs the cartel becomes (internally) less stable (see eq. (1.12)).

entrant with a low cost of entry).²⁹ Unfortunately, we do not have data on entry costs associated with the working of this specific cartel. But still, we can empirically test whether the new (accommodated) entrants contributed to an increase in the instability of the cartel agreement after entry has taken place. Specifically, an increase in cartel instability after entry can be justified, according to the theoretical model at hand, by a consequent adoption of longer (optimal) punishment periods once entry has occurred.

1.4 The Joint Executive Committee

In this section we will briefly review the history of the Joint Executive Committee railroad cartel in the period between 1880 and 1886.³⁰

The JEC was a public and legal agreement³¹ formed in April 1879, involving the railroads in the market. The aim of this cartel was to control the transport of grain, flour and provisions from Chicago to the East Coast. The colluding firms agreed upon a transport price structure and each member of the cartel was allocated a market share. Our attention can be focused on the movements of grain without loss of generality, since the prices for transporting flour and provisions were very closely related to the grain rate.³²

²⁹ Notice, however, that the equilibrium in the original model could be slightly modified such that entry could be interpreted as an equilibrium phenomenon. Formally, assume that the entry sunk cost is K most of the time, but it gets small enough to allow entry with small probability. If the probability of entry is very small, then it does not affect the continuation payoff of the incumbent firms (hence, their incentive constraints) very much. If the equilibrium of the original model is slightly modified so that the incentive constraints of the incumbent firms hold strictly (i.e. incumbent firms prefer the collusive price strictly to secret price cutting), then it will be still an equilibrium which allows entry to occur with positive but small probability.

³⁰ A more detailed analysis of the history of the JEC cartel can be found in MacAvoy (1965).

³¹ Note that this cartel took place before the Sherman Act (1890).

³² With respect to this, MacAvoy (1965) pointed out that "it was not possible to isolate grain rates from the rest so that grain agreements could have been broken while agreements on provision and merchandise rates were not." (p. 71).

Two sources of arguments have led us to consider that it is reasonable to accept that the Green and Porter (1984) model fits this specific case. First, prices set by cartel members were not perfectly observable by its rivals.³³ Moreover, variability of aggregate demand was not only related to prices charged by cartel participants but also to some “unpredictable stochastic forces”.³⁴

Entry occurred twice during the sample period we cover in this article. However, Porter (1983b) states that “bankrupt railroads were relieved by the courts of most of their fixed costs and instructed to cut prices to increase business”(p. 303). This fact led some authors to defend the position that because of the existence of this “no-exit constraint”, the incumbent firms’ best response to entry was to accommodate entrants, allocating them market shares. Our claim here is that even though entry can lead the incumbent firms to an accommodation strategic behavior, the greater the number of firms in the agreement, the higher will be the probability of future price wars. In other words, we are interested in testing how entry may affect cartel stability.

1.5 The data

In this section we describe the data used in the econometric model. We deal here with weekly time-series data which was gathered and disseminated to member firms by the JEC.

The data available corresponds to the period which starts at the first week of 1880 and finishes on week 16 in 1886. Within the 328 sample points, five different periods related

³³ In this direction, Hajivassiliou (1997, p. 6) remarks that “special shipping rates were sometimes secretly arranged with selected costumers”.

³⁴ Porter (1985, p. 420).

Variable	Description
GR	Grain rate; in dollars per 100 lbs.
TQG	Total quantity of grain shipped (in tons).
Lakes	Dummy variable; =1 if Great Lakes were open to navigation.
PW	Cheating dummy variable; =1 when cheating was reported to have occurred.
N	Number of firms in the cartel.

1. List of Variables

to changes in the cartel composition can be distinguished. During the first 27 weeks there were 3 firms in the cartel: the New York Central, the Penn and the Baltimore and Ohio. The Grand Trunk Railway entered the cartel in week 28 of 1880. In week 11 of 1883 the New York Central added another line to its network. A fifth firm - the Chicago and Atlantic - entered the cartel in week 26 of 1883. Finally, in week 12 of 1886 this last entrant decided to leave the cartel due to a dispute with a non-JEC railroad (the Erie). A list of variables which will be used in the econometric models is presented in Table 1.

The price variable, GR, is the weekly reported price of grain (in dollars per lbs). This is an index provided by the JEC after pooling all the member firms. The quantity variable, TQG, is the aggregate tonnage of grain which was shipped by the JEC members in each of the weeks included in the sample period. The Lakes variable is a dummy variable which takes the value one whenever the Great Lakes were open to navigation and steamers could, therefore, compete with the railroads.³⁵ PW is the "cheating" dummy, and equals one when cheating was reported to have occurred in the Railway Review and is used as a proxy of cartel breakdowns. The original dataset developed by Porter (1983b) is here expanded to include also the number of firms in the agreement in each period of time, variable *N*. This

³⁵ Briggs (1996) considers lake steamers as another type of entrant which competed with the cartel for traffic. We do not follow his position. Instead, we consider lake traffic as alternative transportation services through the Great Lakes to ports along the Lake Erie, while the JEC cartel was operating.

Variable	Mean	Standard Deviation	Minimum Value	Maximum Value
GR	.2465	.0665	.125	.4
TQG	25384.4	11632.77	4810	76407
Lakes	.5732	.4954	0	1
PW	.3811	.4864	0	1
N	4.3506	.6273	3	5

2. Summary Statistics

new variable will allow us to control for structural changes in the cartel composition due to entry of new firms or departures from the JEC.

Table 2 presents some summary statistics of the variables.

Remember that one of the main arguments of the theoretical model presented in sections 2 and 3 is that a (tacit) agreement may be more difficult to reach and sustain when the number of firms in the cartel is larger. Using the information provided by variable PW, we constructed some descriptive statistics about the proportion of weeks during which the member firms colluded, controlling for the number of firms in the agreement.

Table 3 presents these results and can be seen as a first empirical approach to the main problem we want to address in this chapter.³⁶

³⁶ Using the same data set, Briggs (1996) presents a similar table. However, he does not take into account the fact that from week 12 of 1886 on there were only four firms in the cartel, because of the departure of the Chicago and Atlantic.

Number of Firms	Number of Weeks	Collusive Periods (%)
3	27	100 %
4	159	72,33 %
5	142	42,96 %

3. Proportion of Collusive Periods and Cartel Composition

Note that when only the initial three firms composed the cartel, collusion was sustained during all the first twenty-seven weeks of the sample. However, when four and five firms were in the cartel, collusion was successfully sustained for only 72% and 43% of the correspondent periods, respectively. Therefore, these results, although based on very simple descriptive statistics, reveal that it is reasonable to admit that the entry of additional firms to the cartel affected its profitability and stability.

1.6 The econometric model

In this section we will present a binary choice econometric model and discuss its results.

Binary choice models have been widely used in empirical applications to cross-section data. Nevertheless, there are also some time-series applications in which these tools can play a very important role. In this study we cover one of the latter cases since we are examining the economic decisions that cartel members had to make, in every single period of time, with regard to whether to cooperate or to begin a price war.

The use of binary choice models relies on a strong assumption of independence across observations. In this particular model, we think that it is reasonable to assume that this is the case. Remember that the Green and Porter (1984) model is developed in a context of uncertainty about the level of demand and unobservability of rival's actions. It defends that

a price war is triggered whenever a firm observes an unusual drop in its demand. However, the probability of a breakdown of the cartel agreement depends on a “signal-extraction problem”³⁷ which is faced by each of the cartel participants. In addition, the model assumes that the realizations of demand are independently and identically distributed (i.i.d.) over time. In those circumstances, each member firm cannot verify if the absence of demand was caused by a realization of a low demand state or by a deviation of someone else in the agreement.

It is, therefore, clear that an equilibrium outcome in which the collusive price is charged forever is not sustainable. Price wars must occur in order for the collusive behavior to be sustained and, indeed, during the history of the JEC, attempts to achieve a price coordination met with phases of aggressive price competition.

The aim of the econometric model presented in this section is to identify the causes of the price wars which occurred in the American railroad cartel during the nineteenth century. Namely, we test the importance of entry effects on the collusive price scheme designed by the JEC.

Let us start by assuming that the probability of a price-war occurrence in period t ($\Pr(W_t = 1)$) can be explained according to the following logit model:

$$\Pr(W_t = 1) = \frac{1}{1 + e^{-Z_t}}, \quad (1.40)$$

where

$$Z_t = \beta_1 + \beta_2 N_t + \beta_3 \text{Lakes}_{t-1} + \beta_4 TQ G_{t-1} * (1 - W_{t-1}) + \varepsilon_t,$$

where $t = 2, \dots, 328$ and ε is an unobserved disturbance.

³⁷ Tirole (1988, p. 263).

Equation (1.40) represents the (cumulative) logistic distribution function. It shows that $\Pr(W_t = 1)$ is not only non-linear in the independent variables, but in the β 's as well. Therefore, we cannot use the OLS procedure to estimate the parameters. However, using (1.40) we have that,

$$\frac{\Pr(W_t = 1)}{1 - \Pr(W_t = 1)} = e^{Z_t}. \quad (1.41)$$

Now, $\Pr(W_t = 1) / (1 - \Pr(W_t = 1))$ is simply the *odds ratio* in favor of a cartel breakdown. If we take the natural log of (1.41), we obtain:

$$L_t = \ln \left(\frac{\Pr(W_t = 1)}{1 - \Pr(W_t = 1)} \right) = Z_t. \quad (1.42)$$

For estimation purposes, we write (1.42) as follows:

$$L_t = \beta_1 + \beta_2 N_t + \beta_3 \text{Lakes}_{t-1} + \beta_4 TQG_{t-1} * (1 - W_{t-1}) + \varepsilon_t.$$

Notice that L_t , the log of the odds ratio, is not only linear in the independent variables, but linear in the parameters also. The Maximum-likelihood estimation results of this model are reported in Table 4³⁸ and the interpretation of the values presented in the odds ratio column is as follows: β_i measures how the log-odds in favor of a cartel breakdown changes as the independent variable associated with β_i changes by one unit, controlling for all other predictors in the model.

We will start with a brief comment about the goodness of fit of the model. First of all, it should be stressed that models with discrete dependent variable are never constructed with the aim of maximizing goodness of fit, but we would like to say a few words about it. Moreover, note that analyzing the Likelihood Ratio Index³⁹ may be misleading, since

³⁸ Standard errors are reported in parentheses with $p < 0.1 = *$ and $p < 0.01 = **$.

³⁹ $LRI = 1 - \frac{\ln L}{\ln L_0}$, where $\ln L_0$ is the log-likelihood computed with only one constant.

Independent Variable	Odds Ratio
N_t	4.298*** (1.594)
$Lakes_{t-1}$	0.460* (0.195)
$TQG_{t-1} * (1 - W_{t-1})$	0.805*** (0.019)
N. Obs.	327
Prob > Chi ²	0.000
Log Likelihood	-91.457
Pseudo R ²	0.58

4. Estimation Results

on the one hand values between 0 and 1 have no natural interpretation and, on the other, an $LRI = 1$ may be indicative of a flaw in the model. For all these reasons, we should focus our attention on the χ^2 test on the significance of the parameters. Following this line of reasoning, Table 4 shows that the (null) hypothesis that all the independent variables' coefficients are zero is strongly rejected. Hence, the model at hand can be considered to be statistically significant.

Turning to the interpretation of the findings, it should be noticed first that the variable $Lakes_{t-1}$ is the regressor with less explanatory power. Its estimated coefficient indicates that price wars were less likely to occur when Great Lakes were open to navigation.

We have also included the variable $TQG_{t-1} * (1 - W_{t-1})$ to capture the effect of a negative shock in the previous period aggregate quantity, provided that period $t - 1$ was a collusive period. It turns out that, as predicted by the theoretical model under consideration, price wars were more likely to occur the smaller the aggregate quantity sold in the last (collusive) period. This result goes in the direction of the predictions of the theoretical model we are considering here. Remember that, along the equilibrium path, a price war

can only be triggered by a realization of a "low-demand state" in the previous (collusive) period, which is exactly the causal effect we intend to measure with this last variable.

Lastly, and most importantly, in the case of variable N_t , as predicted by Stigler (1964), our results confirm that the greater the number of firms in the cartel, the more difficult it is to support a collusive outcome, that is, the likelihood of a price war occurrence increases as the number of cartel members increases. Two different forces justify this result. On the one hand, as the number of firms in the cartel increases, the higher the one-shot incentives to cheat on the collusive agreement are. Second, it appears that the equilibrium length of the punishment was made longer when there were more firms in the agreement. This evidence seems, therefore, to suggest that the actual entry sunk cost during the operation of the JEC was not high enough to discourage new competitors from entering the market. However, this actual sunk cost of entry seems to be at the same time sufficiently high so that after entry has taken place, the enlarged set of active firms coordinated on a new equilibrium from that point onwards, adopting a longer (optimal) punishment length and taking for granted that entry would not occur any more (see Corollary 2). As a result, the proportion of periods spent in price-wars once more firms joined the cartel was higher, as shown by the estimation results.

To summarize, the results support the conjectures of the extended version of the Green and Porter (1984) model presented above. In particular, it reveals that the greater the number of firms belonging to the cartel agreement, the less likely it was that firms would succeed in co-ordinating their pricing behavior; thus entry played an important role as a determinant of the likelihood of a cartel breakdown.

1.7 Conclusion

In this chapter we have defined which conditions should be met in order for cartel members to maintain a non-trivial degree of market power in a context of imperfect observability of rivals' behavior and when there existed the threat of entry of new firms. To this end, we have developed an extended version of the Green and Porter (1984) model in which entry is modeled explicitly. In particular, the model considers the cases where incumbent firms can respond to entry either by adopting a perfectly competitive behavior for a while or by accommodating the new firms. The model suggests that entry barriers are necessary for an oligopoly with some degree of collusion to prevail. It also shows that even when strictly positive entry barriers exist, the existence of a pool of competitors is a more important constraint on the maintenance of a stable agreement when a potential entrant expects that incumbents will allow him to join the (tacit) agreement.

A natural extension of the model we develop in this chapter would examine cases in which the firms have the opportunity to negotiate anew at the beginning of each period of the infinite horizon game. For the purposes of this study, however, models of this sort are classified as the subject of future research.

In the empirical part of this chapter, we focus on time-series data on the Joint Executive Committee to test key theoretical predictions about collusive behavior using models of discrete dependent variable. Specifically, assuming a context of demand uncertainty, we analyze the determinants of the probability with which a price war occurs.

We have found that the empirical results are consistent with our reexamined theoretical model's prediction that the number of firms in the industry plays an important role as a determinant of cartel instability.

Chapter 2

Tacit Collusion, Cost Asymmetries and Mergers

2.1 Introduction

It has long been recognized that "...the more cost functions differ from firm to firm, the more trouble the firms will have maintaining a common price policy, and the less likely joint maximization of profits will be"(Scherer (1980), p. 205). Unfortunately, however, most of the studies which have discussed the factors that facilitate or hinder tacit collusion have only examined the not very realistic case in which firms are perfectly symmetric in terms of their costs. The present study investigates how asymmetry in cost functions across firms affects the scope for collusion and provides conditions under which a collusive outcome involving all firms in the market can be supported.

We employ a model in which cost asymmetric firms repeatedly set quantities and use *optimal penal codes* to enforce collusion. Asymmetry will be dealt with by assuming that firms have a different share of a specific asset (say, capital) which affects marginal costs. Therefore, a firm is considered "large" if it owns a large fraction of the capital stock, and "small" if it owns only a restricted proportion of the capital available in the industry.

We start by characterizing firms' incentives to deviate from the collusive phase as well as their incentives to deviate from the punishment scheme. We show that these incentives turn out to crucially depend on the asset holdings of the firms in the industry. Specifi-

cally, joint profit maximization implies that output is shifted away from small (inefficient) firms towards large (efficient) firms. This implies that the smallest firm in the industry is the one that has the highest potential to steal the business of its rivals and, hence, has the highest incentives to disrupt the collusive agreement. This study thus provides a theoretical rational to the finding of Mason, Phillips and Nowell (1992) that, in experimental duopoly games, "low-cost agents are unable to induce high cost agents to collude" (pp. 665-666). In addition, it is also shown that the incentives to deviate are exactly reversed when the equilibrium calls for punishments. Following Abreu (1986, 1988), we assume that if a deviation occurs, all firms expand output for one period so as to drive price below cost and return to the most collusive sustainable output in the following periods, provided that every player went along with the first phase of the punishment. Since the largest firm is the one that proportionally loses most in the first period of the punishment, it will have the highest incentive to deviate from the punishment strategy.

We then identify a minimal threshold for the discount factor in order for collusion to be sustained and study the impact of changes in firms' asset holdings on this minimal threshold. In spite of the simplifying assumptions (namely, the particular demand and cost functions used), the results offer some interesting merger policy implications. Specifically, they allow us to discuss the issue of *joint or oligopolistic dominance* which has been gaining increasing importance in European merger control.⁴⁰ The analysis suggests that the evaluation of whether the structural change implied by a merger creates more favorable

⁴⁰ The issue of joint dominance was first used by the European Commission in the *Nestlé/Perrier* case. However, only in more recent cases, such as the *Kali und Salz* case and the *Airtours/First Choice* case, it has become clear that this concept can be used to block mergers within the European merger control. For a detailed analysis on this see Motta (2000).

conditions for tacit collusion to arise between the remaining firms, depends on which firms the merger involves. In particular, it is shown that two different effects can be induced by a merger: (i) if firms were already colluding before the merger takes place, then the merger will only have effects on the scope for collusion if it affects the size of the largest firm in the industry. A merger increasing the size of the largest firm gives rise to a more asymmetric distribution of assets and this offsets the increased risk of anticompetitive behavior due to higher concentration; (ii) If, instead, firms were not colluding before the merger, then a merger might make collusion enforceable afterwards. This will occur when the merger involves very small (and, hence, inefficient) firms, which, as already mentioned, turn out to have very high incentives to disrupt the collusive agreement.

To the best of our knowledge, the only paper that discusses the impact of cost heterogeneity on the stability of tacit collusion is Rothschild (1999). In a repeated game setting, Rothschild shows that the stability of tacit collusion depends crucially, and in quite a complex way, on the relative efficiencies of the deviant and nondeviant firms. There exist, however, two major differences between Rothschild's framework and the setting used in this chapter. First, while Rothschild assumes that firms adopt standard 'grim' trigger strategies, in this model firms' strategies incorporate optimal punishments with a *stick and carrot* structure in the style of Abreu (1986, 1988) to sustain a mutually desirable collusive outcome.⁴¹ Specifically, Abreu's work is extended to consider a class of "proportional penal codes".⁴² Second, in this study costs are not exogenous but depend on assets, that is, on

⁴¹ As Vives (2000) observes, "in general, the threat of Nash reversion does not provide the most severe credible punishment to deviants to a collusive agreement. This fact is important because the more severe the punishment is, the more 'cooperative' outcomes can be sustained." (p. 311).

⁴² In this particular class of penal codes, firms outputs along the punishment path are proportional to their

each firm's share in the industry capital. This fact allows for the discussion of the impact of transfer of asset holdings amongst firms on their incentives to collude.

In two recent papers, which are probably the closest to our study, Compte, Jenny and Rey (1997) and, more recently, Kühn and Motta (1999), working respectively with a Bertrand supergame with asymmetric capacity constraints and with a differentiated goods framework where firms produce different numbers of products, discuss the joint dominance issue based on asset transfers. Both studies reach - despite different mechanisms at work - the same conclusion that a more symmetric industrial structure enhances collusion. However, while in Compte, Jenny and Rey (1997) firms are endowed with different capacities and it turns out that the largest firm (the one with the highest capacity) is the one that has the highest incentive to disrupt the collusive agreement, in Kühn and Motta (1999) firms' assets are product varieties, and the firm which tends to have the largest incentive to deviate is the one with the most limited range of products (the smallest one in the industry). For that reason, Kühn and Motta (1999) conclude that "the specific incentive structure for collusion for small and large firms may vary greatly depending on the type of asset one is concerned about" (p. 2). In the present study, as already mentioned, firms own some share of an industry tangible asset (capital) which affects marginal costs. In line with the two previous works, the outcome which emerges is that a more asymmetric distribution of firms' asset holdings tends to hurt tacit collusion. Nevertheless, it also appears that the smallest firms constitute the main obstacle for the stability of the collusive agreements. Hence, although the asset under consideration implicitly captures the importance of firms' capacity,

share in the industry capital.

as in Compte, Jenny and Rey (1997), the results obtained regarding the mapping between firm's asset holdings and their incentives to collude are much closer to those obtained by Kühn and Motta (1999). In addition, new insights are drawn for practical application of competition policy.

The rest of the chapter is organized as follows. The model is laid out in the next section. In section 3, the case of *perfect efficient collusion* is considered. In this section, firms are assumed to maximize joint profits during the collusive phases. Section 4, discusses the case of *perfect non-efficient collusion*, i.e., firms are assumed to coordinate on the jointly production of the monopoly (aggregate) quantity, but deviate away from the joint profit maximization rule to try and enhance collusion sustainability. This section, therefore, discusses the trade-off that exists between efficiency (joint-profit maximization) and sustainability in an industry-wide cartel of the type presented in this chapter. Finally, section 5 offers some concluding comments.

2.2 The model

Consider n firms which produce in the same market for infinitely many periods. Suppose they make output decisions simultaneously at the beginning of each period. Let $q_{i,t}$ be the quantity chosen by firm i , $i \in \{1, \dots, n\}$, in period t , $t = 1, 2, \dots$.

We assume that the industry inverse demand is piecewise linear:

$$p(Q) = \max \{0, a - bQ\}, \quad (2.43)$$

where $Q \in [0, n\frac{a}{b}]$ is the industry output, p is the price of the output and $a, b > 0$ are demand parameters.

Following Perry and Porter (1985), we assume that what distinguishes firms is the amount of capital they own. Total supply of capital is assumed to be fixed to the industry. For the sake of simplicity, the total quantity of capital is normalized to be one.⁴³ Let k_i be the fraction of the industry capital stock owned by firm i , $i \in \{1, \dots, n\}$. Notice that the assumption of a fixed supply of the industry capital is a key feature of the model which will affect our discussion of the effects of changes on firms' size (as measured by k_i) and in the number of firms on the scope for collusion.

The cost function of a firm that owns a fraction k_i of the capital stock and produces q_i units of output is given by:

$$C_i(q_i, k_i) = cq_i + \frac{q_i^2}{2k_i}, \quad (2.44)$$

where $0 < c < a$,⁴⁴ $0 < k_i < 1$ and $\sum_{i=1}^n k_i = 1$. Without loss of generality, we assume that $k_1 \geq \dots \geq k_n$,⁴⁵ $q_i \in [0, \frac{a}{b}]$ and fixed costs are taken to be zero.

The resulting marginal cost function is linearly increasing:

$$C'_i(q_i, k_i) = c + \frac{q_i}{k_i}. \quad (2.45)$$

Notice that the marginal cost function rotates about the intercept as the proportion of capital owned by firm i (k_i) increases or decreases. Hence, this way of characterizing efficiency differences amongst firms implicitly captures the importance of firms' capacity.⁴⁶

⁴³ As pointed out by Perry and Porter (1985), "this supresses de novo entry into the industry" (p. 220).

⁴⁴ To exclude the trivial case in which production is not viable.

⁴⁵ Firms are ranked by decreasing efficiency.

⁴⁶ If a firm is endowed with a small share of the industry capital, it will face a rapidly rising marginal cost

Assume that in the absence of collusion, firms behave like Cournot competitors. A basic insight from the supergame literature is that nonstationary equilibria of quantity setting oligopoly repeated games are much larger sets than just the Cournot (Nash) equilibrium repeated in every round. Reductions in output below the Cournot levels can benefit all the players, but they also create incentives for some firms to undermine the tacitly collusive agreement by (secretly) expanding their individual output. Hence, in order for tacit collusion to be possible, firms have to use their ability to punish each other's deviations from any supposed equilibrium path, by using a *credible penal code*. A penal code is a rule which specifies what players should do in the event that a firm deviates either from the collusive path or from the behavior specified in the penal code. If a penal code is *credible*, then, in each period of the game, given that all the other firms have decided to follow the behavior prescribed by the penal code, each individual player maximizes the present value of its profits stream by also obeying the penal code. Subgame perfection is here used as the equilibrium concept. Unfortunately, as has been shown by different versions of the Folk Theorem,⁴⁷ there exists a large set of subgame perfect equilibrium strategies, if players are sufficiently patient. However, in order to carry out comparative statics exercises, which turn out to be particularly relevant to analyze the effects of mergers (and other capital transfers) on the scope for collusion, a plausible equilibrium will, in what follows, be selected and characterized within the set of subgame perfect equilibria.

A standard example of a *credible penal code* is the one proposed by Friedman (1971), which consists in a Cournot-Nash reversal forever. This penal code is easily seen to be

curve.

⁴⁷ See, for example, Fudenberg and Maskin (1986).

credible since no player can gain by deviating in the punishment phase, because play there is just an infinite number of repetitions of a static (Nash) equilibrium. Notice, however, that such an infinite unforgiving punishment might seem rather extreme. This fact justifies the importance of two papers by Abreu (1986, 1988). In order to derive the highest level of profits which can be sustained by a fixed number of firms as a subgame perfect equilibrium, Abreu examined a class of more sophisticated punishments than reversion to the one-shot Nash equilibrium. He pointed out that, without loss of generality, attention can be restricted to what he defined as *simple penal code*. A simple penal code has a very simple structure. If firms conform with the strategies of the prescribed equilibrium, then they will earn the value of the best continuation equilibrium. If, instead, a single deviation occurs,⁴⁸ all firms (including the deviant one) revert to a punishment which gives the deviant its worst possible continuation equilibrium.⁴⁹ Assume that period t payoffs are received at the end of period t . For each simple penal code, there exists a vector (v_1, \dots, v_n) , where v_i represents the present value of profits that firm $i \in \{1, \dots, n\}$ receives after its deviation has occurred, discounted to the beginning of the first period after deviation.

Let q^c represent a collusive output vector. Denote the profit corresponding to firm i under collusion as $\pi_i^c(k_i)$ and the (one period) gain from deviation as $\pi_i^d(k_i)$. The collusive output is said to be sustainable if, for some simple penal code and for all i , the potential short-run gains from cheating are no greater than the present value of expected future losses

⁴⁸ Simultaneous deviations are ignored since in seeking a Nash or subgame-perfect equilibria, we ask only if a player can gain by deviating *assuming his opponents play as originally specified*.

⁴⁹ In other words, the most effective way to prevent a player from deviating is to threaten to respond to a deviation from a proposed strategy by playing the subgame perfect equilibrium of the infinitely repeated game which yields the lowest payoff of all such equilibria for the deviator.

which are due to the subsequent punishment. This trade-off is captured by the analysis of the incentive compatibility constraint:

$$\frac{\pi_i^c(k_i)}{1-\delta} \geq \pi_i^d(k_i) + v_i, \quad (2.46)$$

which can also be written as follows:

$$\left(\frac{\delta}{1-\delta}\right) \pi_i^c(k_i) - v_i \geq \pi_i^d(k_i) - \pi_i^c(k_i), \quad (2.47)$$

where $\delta \in (0, 1)$ is the common discount factor. If the condition in (2.46) (or, equivalently, (2.47)) holds for all i , then the collusive solution is self-enforcing for every single firm in the coalition.

In the analysis which follows, we consider the case of perfect efficient collusion. The analysis suggests that the extent to which an industry can sustain a stable collusive agreement depends crucially upon the asset distribution amongst coalition members. In particular, it is shown that small firms represent the main obstacle for industry-wide collusion.

2.3 The analysis of perfect efficient collusion

In this section, the case in which firms are assumed to adopt a joint profit maximization behavior on the collusive path is analyzed, concentrating particularly on understanding the effect of the structure of asset distribution amongst member firms on the scope for collusion. To do so, we start by computing the collusive and the optimal deviation profits for a generic firm i , owning a fraction k_i of the industry capital stock.

2.3.1 Collusive profits

In the specific case of full collusion, the coalition operates as a monopolist with n plants, so the marginal cost of production (mc) must be equalized amongst firms:⁵⁰

$$mc = c + \frac{q_1}{k_1} = c + \frac{q_2}{k_2} = \dots = c + \frac{q_n}{k_n}. \quad (2.48)$$

Using relation (2.48),⁵¹ we can easily derive the 'aggregate' marginal cost of production ($mc(Q)$) by horizontally summing the individual marginal cost functions of the member firms:

$$mc(Q) = c + Q. \quad (2.49)$$

Hence, it is straightforward to show that the collusive aggregate quantity, individual output and market price are, respectively:

$$Q^c = \frac{a - c}{2b + 1}, \quad (2.50)$$

$$q_i^c = \frac{a - c}{2b + 1} k_i, \quad (2.51)$$

and

$$p^c = \frac{a + b(a + c)}{2b + 1}. \quad (2.52)$$

From the above expressions for the equilibrium price and individual quantity, it follows that the profit earned by firm i in a collusive period equals:

$$\pi_i^c(k_i) = \frac{1}{2} \frac{(a - c)^2}{2b + 1} k_i > 0. \quad (2.53)$$

⁵⁰ As opposed to models in which firms have different but constant marginal costs, the cartel problem at hand is not a trivial one.

⁵¹ Notice that $\forall i \in \{1, \dots, n\}, q_i = (mc - c) k_i$.

At this stage, it is worth mentioning that the allocation rule adopted by the cartel in order to share the joint profit maximizing output (as expressed by eq. (2.48)) reflects the firms' different sizes (as shown by eq. (2.51)). Since, as output increases, marginal cost rises more rapidly for a small firm than for a large firm, joint profit maximization implies that the smaller (and, hence, the more inefficient) a member firm is, the lower its share in the aggregate output is. Banning side payments, this implies a correspondingly smaller share in the joint profit (see eq.(2.53)).

2.3.2 Deviation profits

If firm i considers deviating from the collusive agreement, it assumes that all the opponents will keep their quantities constant at the collusive level in the current period. Hence, it takes as given the combined rival's (collusive) output and chooses its deviation quantity (q_i^d) by maximizing the following profit function:

$$\pi_i^d(q_1, \dots, q_n; k_i) = \left(a - bq_i - b \sum_{\substack{j=1 \\ j \neq i}}^n q_j^c \right) q_i - cq_i - \frac{q_i^2}{2k_i}. \quad (2.54)$$

Making use of equation (2.51), one finds that $\sum_{\substack{j=1 \\ j \neq i}}^n q_j^c = \left(\frac{a-c}{2b+1} \right) (1 - k_i)$. Therefore, (2.54) can be rewritten as follows:

$$\pi_i^d(q_i; k_i) = \left(a - bq_i - b \left(\frac{a-c}{2b+1} \right) (1 - k_i) \right) q_i - cq_i - \frac{q_i^2}{2k_i}. \quad (2.55)$$

By maximizing (2.55) with respect to q_i , it turns out that firm i 's optimal deviation quantity equals,

$$q_i^d = \frac{(a-c)(1+b+bk_i)}{(2b+1)(2bk_i+1)} k_i. \quad (2.56)$$

Using eq. (2.56) to substitute for q_i in expression (2.55), one obtains:

$$\pi_i^d(k_i) = \frac{1}{2} \frac{(a-c)^2 (1+b+bk_i)^2}{(2bk_i+1)(2b+1)^2} k_i > 0. \quad (2.57)$$

Using now expressions (2.53) and (2.57) in order to carry out a simple exercise of comparative statics, it can be shown that $\frac{\partial \pi_i^c(k_i)}{\partial k_i} = \frac{1}{2} \frac{(a-c)^2}{2b+1} > 0$ and also that $\frac{\partial \pi_i^d(k_i)}{\partial k_i} = \frac{1}{2} \frac{(a-c)^2 (1+b+bk_i)(bk_i(4bk_i+3)+b+1)}{(2b+1)^2 (2bk_i+1)^2} > 0$.⁵² Hence, the more efficient a member firm is, the higher its share on the collusive profit, on the one hand, and the higher are its deviation profits, on the other.⁵³

2.3.3 Distribution of assets and scope for collusion

Having defined the profits of a representative firm i both at the joint profit maximum and in a deviation scenario, respectively, we now turn to the study of the conditions which must be satisfied in order for a stable fully collusive outcome to exist. This section will show how the incentives to deviate from the fully collusive agreement depend on the asset holdings of the firm, focusing on pure strategy (subgame) perfect equilibria. In particular, we propose a specific class of penal codes more severe than Cournot-Nash reversion, in the style of the ones which have been characterized in general by Abreu (1986, 1988).

In his paper, Abreu (1988) shows that repeated games with discounting may be completely analyzed in terms of simple strategy profiles. A simple strategy profile is "a rule specifying an initial path (i.e., an infinite stream of one-period action profiles), and *pun-*

⁵² Remember that the total quantity of capital was assumed to be exogenously given and normalized to be one. Therefore, when performing exercises of comparative statics with respect to k_i , we are simply comparing the effect of interchanging a more efficient firm with a less efficient one.

⁵³ As we will see below, this does not imply that the larger the firm, the higher its incentive to deviate from the collusive agreement.

ishments (also paths, and hence infinite streams) for any deviations from the initial path or from a previously prescribed punishment" (Abreu (1988), p. 383). More formally, as already mentioned, $q_{i,t}$ denotes the quantity chosen by firm i , $i \in \{1, \dots, n\}$, in period t , $t = 1, 2, \dots$. Let $q(t) \equiv (q_{1,t}, \dots, q_{n,t})$. An action profile $\{q(t)\}_{t=1}^{\infty}$ is referred to as a path or punishment and is denoted by $P \in \Omega$, where Ω represents the set of paths.

Definition 1 (cf. Abreu(1988)) Let $P^i \in \Omega$, $i = 0, 1, \dots, n$. A simple strategy profile $SSP(P^0, P^1, \dots, P^n)$ specifies: (i) play P^0 until some player deviates unilaterally from P^0 ; (ii) for any $j \in \{1, \dots, n\}$, play P^j if the j -th player deviates unilaterally from P^i , $i = 0, 1, \dots, n$, where P^i is an ongoing previously specified path; continue with P^i if no deviations occur or if two or more players deviate simultaneously.

A simple strategy profile is therefore history-independent in the sense that it specifies the same punishment P^i for any deviation (from the initial path P^0 or from a previously prescribed punishment) by player i . When there is an unilateral deviation, the subsequent sequence of action profiles depends on the identity of the deviant and not on the history that preceded its deviation.⁵⁴ Now, a *simple penal code* is defined by an n -vector of punishments (P^1, \dots, P^n) , where P^i is inflicted if player i deviates. Notice that the elements of simple strategy profiles that define a simple penal code differ only with respect to the initial path they prescribe.

⁵⁴ As was highlighted by Abreu (1988), "in no sense is there any need to 'make the punishment fit the crime'." (p. 385).

Necessary and sufficient conditions for perfect efficient collusion

The following Lemma shows that under the assumptions of our model, a simple penal code exists which is an optimal penal code. A penal code is said to be optimal if it yields to the deviant player the lowest possible continuation payoff in any (subgame) perfect equilibrium of the model at hand.

Lemma 1 *An optimal simple penal code exists.*

Proof. To prove that a simple penal code exists, all we need to show is that the model fits Assumptions 2 to 4 in Abreu (1988). Notice that:

1. $(q_1, \dots, q_n) \in [0, \frac{a}{b}]^n$, which is a compact topological space;
2. The one period profit function of a generic firm i , $\pi_i(q_1, \dots, q_n; k_i)$, is continuous.
3. Since the cost functions $C_i(q_i, k_i)$ are strictly convex in the first argument $\left(\frac{\partial^2 C_i(q_i, k_i)}{\partial q_i^2} > 0\right)$ and the (inverse) demand function is piecewise linear, each firm's profit function is strictly quasi-concave in its own output. Moreover, the profit functions are continuous and their domain is compact. Hence, the one-shot (stage) game has a pure-strategy Nash equilibrium (a Cournot equilibrium exists) and,⁵⁵ therefore, the set of perfect equilibrium strategy profiles of the supergame with discounting is nonempty.

Hence, by Proposition 2 in Abreu (1988), an optimal simple penal code exists. ■

⁵⁵ See Appendix A.

Lemma 1 establishes the existence of an n -vector of punishments. The i -th vector is an infinite stream of action profiles specifying what each player should do in the event of a (single) deviation by firm i from the agreed upon initial path, or from a previously prescribed punishment. In the case that player i 's specific punishment is imposed, this player will earn its lowest possible perfect equilibrium payoff. Notice, however, that this result does not provide us with the specific intertemporal structure of the optimal punishment paths. In what follows, we show that although asymmetry amongst firms' cost functions is assumed, the optimal punishments inflicted to deviant players may have very simple structures, as the following definition suggests.

Definition 2 $\sigma(q^1, q^2)$ denotes a simple "proportional" two-phase penal code, where:

- $q^j = (q_1^j, \dots, q_n^j) \in [0, \frac{a}{b}]^n$, for $j = 1, 2$;
- $Q^j = \sum_{i=1}^n q_i^j$, for $j = 1, 2$;
- $q_i^j = k_i Q^j$ for $i \in \{1, \dots, n\}$, $j = 1, 2$;
- $q_{i,t} = \begin{cases} q_i^1, & \text{if } t = 1 \\ q_i^2, & \text{if } t = 2, 3, \dots \end{cases}$.

Two remarks are in order at this point. First, notice that the proposed class of penal codes (proportional to capital shares) uses the *same* punishment path for *every* deviating firm. Second, this particular class of punishments has a two-phase structure, implying that punishments are stationary after the first period.

Definition 3 $v_i(q^1, q^2) \equiv \delta (\pi_i(q^1) + (\frac{\delta}{1-\delta}) \pi_i(q^2))$, where $\pi_i(q^j)$, for $i \in \{1, \dots, n\}$, $j = 1, 2$, is the profit earned by firm i when q^j is the vector of quantities produced by all firms.

Abreu (1986) applies the systematic framework presented in Abreu (1988) to oligopolistic quantity-setting supergames. Our setting respects Assumptions (A2)-(A5) in Abreu (1986). However, instead of identical firms producing a homogeneous good at constant marginal cost, we consider cost asymmetric firms whose efficiency differences are characterized by the cost function (2.44). The next Lemma aims at showing that, under the assumptions of our model and if players are sufficiently patient, a “proportional” two-phase penal code exists, yielding every player a payoff of zero. Since zero is the minmax payoff for every firm in the component game, the proposed penal code turns out to be *globally optimal*.⁵⁶

Lemma 2 *There exists a lower bound $\underline{\delta} < 1$ such that for every $\delta \geq \underline{\delta}$ there exists a pair $(q^1, q^2) \in [0, \frac{a}{b}]^{2n}$ such that $\sigma(q^1, q^2)$ is an optimal simple “proportional” two-phase penal code yielding $v_i(q^1, q^2) = 0$ for all $i \in \{1, \dots, n\}$ if and only if $\delta \geq \underline{\delta}$.*

Proof. For any $(q^1, q^2) \in [0, \frac{a}{b}]^{2n}$, $v_i(q^1, q^2) = 0$ if and only if:

$$-\pi_i(q^1) = \delta (\pi_i(q^2) - \pi_i(q^1)). \quad (2.58)$$

On the other hand, $\sigma(q^1, q^2)$ is a perfect equilibrium if and only if no member firm has incentives to deviate from any phase of the punishment, that is, if and only if, for all i ,

$$\pi_i^*(q_{-i}^1) - \pi_i(q^1) \leq \delta (\pi_i(q^2) - \pi_i(q^1)), \quad (2.59)$$

$$\pi_i^*(q_{-i}^2) - \pi_i(q^2) \leq \delta (\pi_i(q^2) - \pi_i(q^1)), \quad (2.60)$$

⁵⁶ Zero is the lowest possible payoff that firms are willing to accept in order to go along with the punishment, since firms can guarantee themselves a zero payoff by producing nothing forever.

where $q_{-i}^j = (q_1^j, \dots, q_{i-1}^j, q_{i+1}^j, \dots, q_n^j)$, $j = 1, 2$, and $\pi_i^* : [0, \frac{a}{b}]^{n-1} \rightarrow \mathbb{R}$ denotes firm i 's best response profit, that is, $\pi_i^*(q_{-i}^j) = \max \{ \pi_i(x, q_{-i}^j) \mid x \in [0, \frac{a}{b}] \}$. Since $\sigma(q^1, q^2)$ satisfies (2.58), eqs. (2.59) and (2.60) can be rewritten as follows, respectively:

$$\pi_i^*(q_{-i}^1) = 0, \quad (2.61)$$

and,

$$\pi_i^*(q_{-i}^2) \leq \pi_i(q^2) - \pi_i(q^1). \quad (2.62)$$

Let $D = \{ \delta \in (0, 1) \mid (\delta, q^1, q^2) \text{ satisfies Eqs. (2.58), (2.61) and (2.62) for some } (q^1, q^2) \in [0, \frac{a}{b}]^{2n} \}$. Continuity of $\pi_i(\cdot)$ and $\pi_i^*(\cdot)$ implies that D is closed.⁵⁷ Since demand is piecewise linear, $p(0) > C'_i(0)$ (that is, $a > c$) and $C''_i(q_i) > 0$ (assumptions (A2)-(A4) in Abreu (1986) hold in our framework), there exists $q^0 \in [0, \frac{a}{b}]^n$ and, therefore, a $Q^0 = \sum_{i=1}^n q_i^0$ such that $p(Q^0(1 - k_i)) \leq C'_i(0) = c$.⁵⁸ Hence, by (A2) in Abreu (1986) and since $C''_i(q_i) > 0$, one has that $\pi_i(q^0) < 0$ and $\pi_i^*(q_{-i}^0) = 0$. Observe that there exists a proportional fully collusive output vector $q^c \in [0, \frac{a}{b}]^n$, where q_i^c is given by equation (2.51). Let $\delta' = \frac{-\pi_i(q^0)}{\pi_i(q^c) - \pi_i(q^0)} < 1$. Then, (δ', q^0, q^c) satisfies Eqs. (2.58), (2.61) and (2.62), and D is nonempty.⁵⁹ Let $\underline{\delta} = \min D$. Then, for $\delta < \underline{\delta}$ there exists no (δ, q^1, q^2) such that $\sigma(q^1, q^2)$ is a Perfect Equilibrium and $v_i(q^1, q^2) = 0$. So, the "only if" part of the proof is complete.

⁵⁷ Notice that the condition $\delta > 0$ in the definition of D is not binding/relevant. When $\delta = 0$, condition (2.58) is not satisfied. As a consequence, continuity of $\pi_i(\cdot)$ and $\pi_i^*(\cdot)$ implies that $\min D$ exists (and is positive).

⁵⁸ It is worth mentioning at this point that since in this setting $C'_i(0) = c$, Q^0 can be chosen to fit all possible deviating firm i .

⁵⁹ Notice that (2.60) is satisfied, since one can chose Q^0 to be high enough.

Let $(q^{1*}, q^{2*}) \in [0, \frac{a}{b}]^{2n}$ satisfy Eqs. (2.58), (2.61) and (2.62) for $\delta = \underline{\delta}$. Now consider $\hat{\delta} \geq \underline{\delta}$ and let $\hat{q}^1 \in [0, \frac{a}{b}]^n$ satisfy Eq. (2.58) for $q^2 = q^{2*}$ and $\delta = \hat{\delta}$. By (A2)-(A4) in Abreu (1986), \hat{q}^1 exists, $-\pi_i(\hat{q}^1) \geq -\pi_i(q^{1*})$, and $\hat{Q}^1 \geq Q^{1*} \geq Q^0$, so that $(\hat{\delta}, \hat{q}^1, q^{2*})$ satisfies Eqs. (2.58), (2.61) and (2.62). Thus, $\sigma(\hat{q}^1, q^{2*})$ is a perfect equilibrium and yields $v_i(\hat{q}^1, q^{2*}) = 0$ for all $i \in \{1, \dots, n\}$. ■

The intuition behind this result is as follows. A potential deviant has to trade off the short-run gains from deviation with the *future* discounted loss due to the restarting of a punishment phase.⁶⁰ Hence, the discount factor should be high enough for firms not to deviate from the punishment strategy. Besides, if δ is higher, firms can increase the severity of the punishment by appropriately choosing a higher aggregate output in the first phase of the punishment scheme (Q^1 in our notation). Therefore, there is some minimum discount factor, denoted $\underline{\delta}$, such that it is credible to impose a punishment which yields each firm a zero payoff.

Lemma 2 simplifies the characterization of the relationship between asset distribution and firms' incentives to collude, considered in the next Proposition.

Proposition 1 *Let $\underline{\delta}$ be as in Lemma 2. Perfect efficient collusion is sustainable in a subgame perfect equilibrium if and only if $\delta \geq \max\{\underline{\delta}, \tilde{\delta}_n\}$, where $\tilde{\delta}_n = \frac{b^2(1-k_n)^2}{(1+b+bk_n)^2}$.*

⁶⁰ Notice, in this direction, that, as pointed out by Abreu (1988), "the early stages of an optimal punishment must be more unpleasant than the remainder." (p. 385) Therefore, the punishment is made credible by the threat of being restarted should any player deviate from the punishment strategy.

Proof. Notice that the incentive compatibility constraint (eq. (2.46)) can be rewritten in the following way:

$$\pi_i^c(k_i) - (1 - \delta) \pi_i^d(k_i) \geq (1 - \delta) v_i. \quad (2.63)$$

By Lemma 2, for every $\delta \geq \underline{\delta}$, an optimal “proportional” penal code exists yielding $v_i = 0, \forall i \in \{1, \dots, n\}$. Hence, if $\delta \geq \underline{\delta}$, the r.h.s. of condition (1.8) equals zero.

Now, making use of eqs. (2.53) and (2.57), one concludes that $\pi_i^c(k_i) - (1 - \delta) \pi_i^d(k_i) \geq 0$, if and only if, for all $i \in \{1, \dots, n\}$:

$$\delta \geq \frac{b^2 (1 - k_i)^2}{(1 + b + bk_i)^2} \equiv \tilde{\delta}_i. \quad (2.64)$$

Notice also that, from (2.64), it can be easily shown that,

$$\frac{\partial \tilde{\delta}_i}{\partial k_i} = -2b^2 \frac{(1 - k_i)(2b + 1)}{(1 + b + bk_i)^3} < 0.$$

Hence, the main problem is to prevent firms with rapidly rising marginal cost curves from deviating. Since in our setting k_n denotes the share in the capital corresponding to the smallest firm in the industry, the condition which should be taken into account in order to evaluate the stability of an industry-wide cartel whose members maximize joint profits at the collusive path is

$$\delta \geq \frac{b^2 (1 - k_n)^2}{(1 + b + bk_n)^2} \equiv \tilde{\delta}_n. \quad (2.65)$$

Hence, if $\delta < \tilde{\delta}_n$, then perfect efficient collusion cannot be sustainable, since (2.65) is a necessary condition under maximal punishments. Take now the case in which $\tilde{\delta}_n < \underline{\delta}$. When this is the case, firms cannot enforce perfect *efficient* collusion when $\delta < \underline{\delta}$. This is so because firms are assumed to use the most severe punishment strategies. These punishment strategies are optimal in the sense that given a certain fixed collusive allocation (namely,

the one given by eq. (2.51)), firms can implement this same allocation with the lowest critical discount factor, $\underline{\delta}$. Thus, a perfect efficient collusion is sustainable if and only if $\delta \geq \max \{ \underline{\delta}, \tilde{\delta}_n \}$. ■

Proposition 1 captures the fact that, as already shown, joint profit maximization implies that output is shifted away from small (inefficient) firms towards large (efficient) firms. As a result, the smallest firm is the one which is allotted a share in the collusive aggregate output that is too low with respect to its optimal deviation output. However, as was highlighted by Martin (1988), a small inefficient firm "may well judge that over the long run its bargaining power within the cartel will be tied to its market share. If this is the case, accepting a lower market share to maximize joint profit will amount to cutting its own throat within the cartel." (p. 137). Notice, in this direction, that, from eqs. (2.51) and (2.56), it is straightforward to show that:

$$\frac{q_i^d}{q_i^c} = \frac{1 + b + bk_i}{2bk_i + 1}. \quad (2.66)$$

For a collusive agreement to be stable, this ratio should not be too high for any member firm i . Working through some algebra, one can show that $\frac{\partial}{\partial k_i} \left(\frac{q_i^d}{q_i^c} \right) = -b \frac{2b+1}{(2bk_i+1)^2} < 0$. Therefore, the smaller the firm is, the higher its potential to profitably capture demand from its opponents by deviating.⁶¹ It should also be stressed at this point that the fact that small firms are the less keen to accept the collusive agreement relies on our assumption of absence of monetary transfers, which we think is the most realistic in most circumstances.

⁶¹ At this point it is worth contrasting this result with the one obtained by Rothschild (1999). In his paper, Rothschild does not give sharp predictions as to the relationship between firms' cost conditions and their ability to sustain a collusive agreement to restrict output. His Proposition 4 shows that the most inefficient firms *might be* the ones that determine the stability a fully collusive agreement. However, in his result the propensity of these firms to deviate depends crucially upon the Cournot outputs of the nondeviant firms, and this in turn depends on the relative efficiencies of the deviant and nondeviant firms.

The previous proposition also highlights the fact that in order for firms to credibly participate in this efficient collusive scheme, they should be willing to comply with the collusive path, on the one hand, and with the punishment strategy, on the other. It was shown that on the collusive equilibrium path the incentive constraint which is binding is that of the smallest firm. In the next proposition it is shown that if the smallest firm is not too small (inefficient), then we can also identify the firm for which the incentive constraint is binding on the punishment path.

Proposition 2 *If the smallest firm in the industry is not too small, that is, if $k_n \in [k^*, k_1]$, where*

$$k^* = \frac{b^3 (k_1^2 + 6k_1 + 1) + 2bk_1 (1 + 4b)}{(1 - k_1)^2 b^3} + \frac{(2b + 1) \left((2b + 1) - (1 + bk_1 + b) \sqrt{((2b + 1)(1 + 2k_1b))} \right)}{(1 - k_1)^2 b^3},$$

then $\underline{\delta} \geq \tilde{\delta}_n$, where $\underline{\delta} = \frac{(1+2bk_1)(2b+1)}{(1+b+bk_1)^2}$.

Proof. From (2.58), one concludes that:

$$\delta = \frac{-\pi_i(q^1)}{\pi_i(q^2) - \pi_i(q^1)}. \quad (2.67)$$

Since, in this setting, $\pi_i(q^1) = (a - c - bQ^1)k_iQ^1 - \frac{(k_iQ^1)^2}{2k_i}$ and $\pi_i(q^2) = (a - c - bQ^2)k_iQ^2 - \frac{(k_iQ^2)^2}{2k_i}$, one can reevaluate (2.67), obtaining

$$\delta = Q^1 \frac{(2b + 1)Q^1 - 2(a - c)}{(Q^1 - Q^2)((Q^1 + Q^2)(2b + 1) - 2(a - c))}, \quad (2.68)$$

where we assume that $Q^1 > 2 \left(\frac{a-c}{2b+1} \right)$ in order for δ to be positive. Now, from (2.68), one can easily show that:

$$\frac{\partial \delta}{\partial Q^2} = \frac{2Q^1 (Q^1 (2b+1) - 2(a-c)) ((2b+1) Q^2 - (a-c))}{(Q^1 - Q^2)^2 ((Q^2 + Q^1) (2b+1) - 2(a-c))^2}, \quad (2.69)$$

which is always *non-negative* for the following two reasons. First, we have just assumed that $Q^1 > 2 \left(\frac{a-c}{2b+1} \right)$ in order for δ to be positive. Second, for all $i = \{1, \dots, n\}$, the individual profit function is strictly concave in its own output and the joint profit maximum is achieved when the aggregate output equals $\frac{a-c}{2b+1}$ (see eq. (2.50)). Hence, $Q^2 \geq \frac{a-c}{2b+1}$. Since we are looking for a minimum value for the discount factor δ , let us, therefore, set $Q^2 = \frac{a-c}{2b+1}$. Reevaluating (2.68) for this specific value of Q^2 , one gets

$$\delta^* = \frac{Q^1 (2b+1) ((2b+1) Q^1 - 2(a-c))}{((2b+1) Q^1 - (a-c))^2}. \quad (2.70)$$

Notice now that, from (2.70), it can be easily shown that

$$\frac{\partial \delta^*}{\partial Q^1} = \frac{2(2b+1)(a-c)^2}{((2b+1) Q^1 - (a-c))^3} > 0. \quad (2.71)$$

Hence, in order to minimize δ^* , we want the lowest possible value of Q^1 for which conditions (2.61) and (2.62) hold. Notice that, as long as $Q^1 \leq \frac{a}{b}$, $\pi_i^*(q_{-i}^1) = \frac{1}{2}(a-c-bQ^1(1-k_i))^2 \frac{k_i}{1+2bk_i}$.⁶² Hence, in order for condition (2.61) to hold, we must have that $Q^1 \geq \frac{a-c}{b(1-k_1)}$.⁶³ Let us, therefore, set $Q^1 = \frac{a-c}{b(1-k_1)}$. Now, using the analytical expressions of $\pi_i(q^1)$ and $\pi_i(q^2)$ specified above and given that $\pi_i^*(q_{-i}^2) = \frac{1}{2}(a-c-bQ^2(1-k_i))^2 \frac{k_i}{1+2bk_i}$, some algebra shows that when we set $Q^1 = \frac{a-c}{b(1-k_1)}$ and

⁶² In the case that $Q^1 > \frac{a}{b}$, $\pi_i^*(q_{-i}^1) = 0$ and, therefore, condition (2.61) is trivially satisfied.

⁶³ Notice that $\forall k_1 \in (0, 1)$, $\frac{a-c}{b(1-k_1)} > 2 \left(\frac{a-c}{2b+1} \right)$ and, therefore, δ^* is nonnegative.

$Q^2 = \left(\frac{a-c}{2b+1}\right)$, condition (2.62) is satisfied if, for all i ,

$$b^4(1-k_1)^2(1+k_i^2) - (2b+1)^2(1+2b(k_1+k_i)) - 2b^4k_i(1+k_1^2+6k_1) - 4b^2k_ik_1(1+4b) \leq 0. \quad (2.72)$$

It can be easily shown that the derivative of the l.h.s. of the previous condition with respect to k_i is always negative. Therefore, the condition which is binding is that of the smallest firm in the industry (whose capital share is given by k_n). In addition, the l.h.s. of (2.72) is a polynomial of second degree in k_i and one of its roots is greater than one (and, therefore, should be discarded). Hence, after some manipulation, one concludes that in order for condition (2.62) to hold, one must have that $k_n \in [k^*, k_1]$, where

$$k^* = \frac{b^3(k_1^2 + 6k_1 + 1) + 2bk_1(1+4b)}{(1-k_1)^2 b^3} + \frac{(2b+1) \left((2b+1) - (1+bk_1+b) \sqrt{((2b+1)(1+2k_1b))} \right)}{(1-k_1)^2 b^3}. \quad (2.73)$$

If this is the case, i.e., if $k_n \in [k^*, k_1]$, making use of (2.70) and setting $Q^1 = \frac{a-c}{b(1-k_1)}$, we obtain:

$$\underline{\delta} = \frac{(1+2bk_1)(2b+1)}{(b+1+bk_1)^2}. \quad (2.74)$$

In addition, making use of eqs. (2.65) and (2.74), one can easily show that if $k_n \in [k^*, k_1]$, then $\underline{\delta} \geq \tilde{\delta}_n$.

This completes the proof of Proposition 2. ■

Notice that from (2.74), one concludes that $\underline{\delta}$ only depends on the capital share owned by the largest firm in the industry. The intuition here rests on the fact that in the first period of the punishment path the aggregate output produced has to be large enough such that a very sharp price cut occurs leading all firms to earn negative profits in this period. Besides,

the largest firm is the one which is proportionally most affected by this price cut since it is the one with the highest market share in the agreement. As a result, a lower bound on the discount factor is clearly necessary. The discount factor has to be sufficiently high so that the largest firm (and, therefore, all the other firms) can recoup the one-period losses on the most attractive (second) phase of the punishment.

The insights of the two previous propositions can be summarized as follows. On the one hand, during the collusive phases the cartel maximizes its joint profit, which implies that the smaller the firm the lower its share in the collusive output. This implies that the incentive constraint that matters is that of the smallest firm. Hence, there exists a minimal discount factor $\tilde{\delta}_n$, only depending on the share in the capital of the smallest firm, above which all firms in the industry find it optimal to keep the cartel agreement. On the other hand, if the smallest firm in the market is not too small, there exists a lower bound on the discount factor $\underline{\delta}$, which only depends on the size of the largest firm, above which an optimal "proportional" two-phase penal code exists. This penal code yields for every firm a continuation payoff of zero after a deviation has occurred. Studying the ranking between the identified thresholds for the discount factor, one concludes that the necessary and sufficient condition which must be met in order perfect efficient collusion to be sustainable is the following:

$$\delta \geq \frac{(1 + 2bk_1)(2b + 1)}{(b + 1 + bk_1)^2} \equiv \underline{\delta}. \quad (2.75)$$

In the next subsection, we study the implications of changes in the distribution of asset holdings (due to mergers, transfers or split-offs) on the sustainability of tacit collusion. By doing so, we draw some merger policy implications.

The impact of mergers on collusion

In this setting, a specific asset (namely, capital) is introduced and assumed to affect firms' marginal costs (see eq. (2.45)). Hence, any merger gives rise to endogenous efficiency gains since it brings the individual capital of the merging firms under a single larger (and, hence, more efficient) resulting firm. A more delicate problem, however, is to understand the impact of merger-induced changes in firm's capital allocations on the sustainability of tacit collusion. This is the issue we address in the present section.

The common wisdom is that mergers tend to create structural conditions which facilitate collusion. The argument typically used is that the lower the number of market participants, the easier it will be for them to coordinate their actions (e.g. the easier it is to allocate market shares) or to monitor departures from agreed-upon output levels.⁶⁴ In what follows, however, it is shown that when cost-asymmetric firms co-exist in the market, two distinct effects can be induced by a merger: (i) If firms were already colluding before the merger, then the merger either has no effect on the scope for collusion or it hinders collusion; (ii) If, instead, before the merger collusion is not feasible, then a merger might make collusion enforceable afterwards.

In our setting, the effect of a merger is not restricted to a decrease in the number of firms. A merger also gives rise to a different distribution of assets amongst the remaining firms (a different post-merger capital allocation). Therefore, a natural question at this point is which capital reallocations can be induced by a merger. Figure 1 shows that two cases should be considered when analyzing this question.

⁶⁴ This is a well-established argument which extends at least as far back as Stigler (1964).

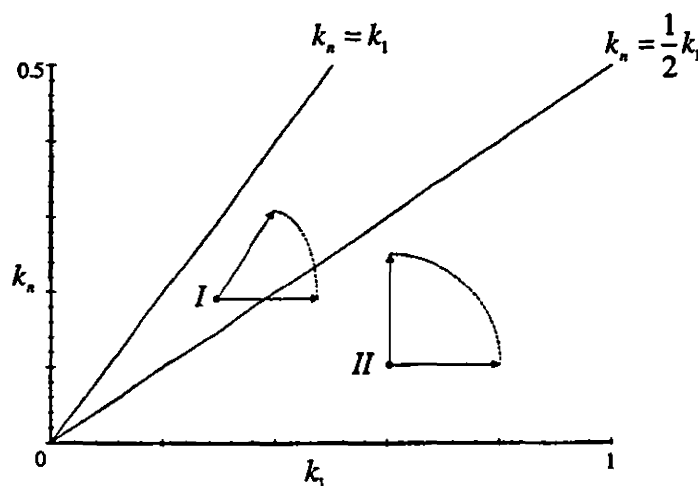


Figure 1: Merger Effects on Capital Allocation.

First, consider the situations in which before the merger $k_n > \frac{1}{2}k_1$ (e.g. allocation I in Figure 1). When this is the case, any merger will lead to an increase in the size of the largest firm. Notice, in this direction, that even if the merging firms are the two smallest firms in the industry, then the size of the resulting merged firm ($k_n + k_{n-1}$) will certainly be greater than k_1 . Take now the case in which the largest firm merges with any other firm but the smallest one. This will lead to an increase in the size of the largest competitor, but the size of the smallest firm will remain unaffected. We, thus, move along the horizontal arrow starting from point I in the Figure. Hence, the capital reallocation induced by any merger can be represented by an arrow starting from point I which lies between the two arrows that form the acute angle presented in the diagram.

Next, take any pre-merger capital allocation such that $k_n < \frac{1}{2}k_1$ (e.g. allocation II in Figure 1). We can still have mergers affecting the size of the largest firm but not that of the smallest firm (we move along the horizontal arrow starting from point II in the picture).

However, now we can also have a merger in which the size of the smallest firm increases but the size of the largest one remains unaffected. Take, for instance, the situation in which before the merger there are two equal-sized smallest firms. If they decide to merge, then the size of the new smallest firm will be $k'_n = \min \{2k_n, k_{n-1}\}$, but the size of the largest firm in the market remains unchanged since $2k_n \leq k_1$. As a result, if allocation II in the diagram is the pre-merger capital allocation, the capital reallocation induced by any merger can be represented by an arrow which starts from II and lies within the right angle formed by the arrows shown in Figure 1.

The previous discussion identified the possible capital reallocations which a merger can induce. Since in this setting both the initial number of firms is given and there is a fixed supply of the industry capital, we can now also identify natural bounds within which capital shares k_1 and k_n can vary. Figure 2 shows a symmetry line along which $k_n = k_1 = \frac{1}{n}$. Since, by definition $k_n \leq k_1$, any feasible capital allocation must lie within the region below this symmetry line.

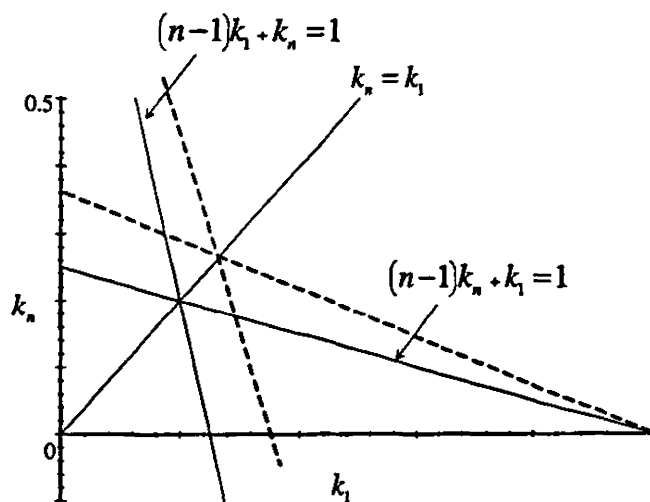


Figure 2: Merger Effects on the Feasibility Region.

The two other solid straight lines represent two extreme cases regarding the industry configuration. In the first case, there exists a large firm owning a share k_1 of the industry capital and the residual capital share $(1 - k_1)$ is equally shared by the remaining (small) firms. In the second case, the industry is composed of a single small firm and $n - 1$ symmetric large firms. In order for the capital constraint to hold before the merger, the pre-merger capital allocation (k_n, k_1) must lie within the triangle formed by these two solid straight lines. If two firms decide to merge, then the number of independent firms in the market is reduced by one. Therefore, the two extreme cases for the industry configuration are now represented by the dashed lines along which $k_1 + (n-2)k_n = 1$ and $(n-2)k_1 + k_n = 1$, respectively. As a result, the feasibility region in which the capital constraint holds after the merger is now represented by the triangle formed by the dashed lines.

We know that in this framework small firms have the highest incentives to deviate from the collusive path. This explains why in Proposition 2 a minimal level of efficiency k^*

for the smallest firm is required in order for perfect efficient collusion to be enforceable, that is, in order for the smallest firm not to have incentives to disrupt the collusive agreement. Some algebra shows that $\forall b > 0 \forall k_1 \in [k_n, 1), \frac{\partial k^*}{\partial k_1} < 0$. This is illustrated in Figure 3 by a solid curve along which $k_n = k^*$.⁶⁵ Notice that collusion can be enforced for every pair (k_n, k_1) above this solid locus. In addition, $\lim_{k_1 \rightarrow 1} k^* = -\frac{1}{2b}$. Hence, as shown in Figure 3, k^* assumes negative values for high enough values of the capital share owned by the largest firm in the market. This is so because, according to the punishment scheme firms adopt in this setting, all the firms in the market must earn a negative profit in the first phase of the punishment. This implies that a lower bound exists for the aggregate output produced in this first phase of the punishment. More precisely, condition (2.61) implies that $Q^1 \geq \frac{a-c}{b(1-k_1)}$. Since this lower bound obviously increases with the size of the largest firm, an increase in k_1 leads to an increase in the first period losses of every firm in the market, that is, induces an increase of the severity of the punishment. As a result, if the largest firm is sufficiently large, then the severity of the punishment becomes so high that any small firm, no matter how small it is, will have no incentives to disrupt the collusive agreement.

⁶⁵ As is shown in Proposition 2, the minimal level of efficiency k^* does not depend on the number of firms in the industry, n (see eq. (2.73)). Therefore, the solid curve $k_n = k^*$ is unaffected by a merger.

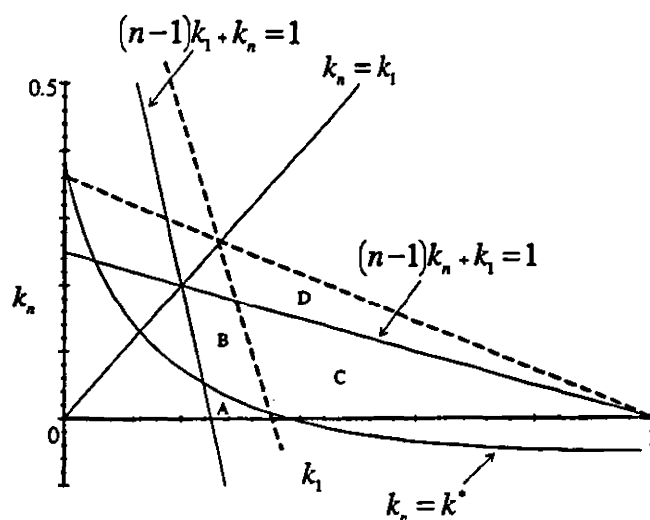


Figure 3: Merger Effects on the Scope for Collusion (Panel a)

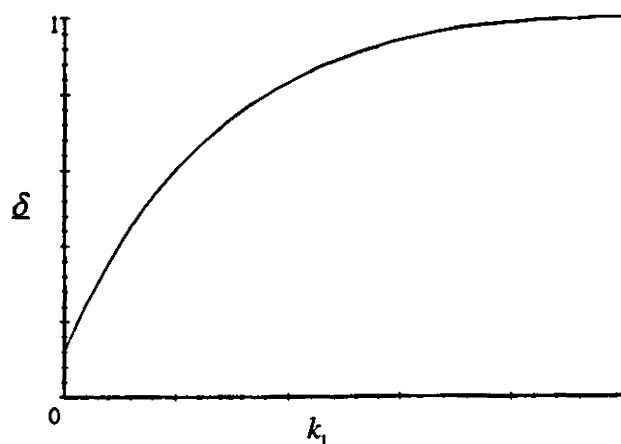


Figure 4: Merger Effects on the Scope for Collusion (Panel b)

In order to be able to discuss the impact of mergers on the scope for collusion, we also need to know how the minimal discount factor reacts to changes in firms' capital allocations. From (2.74), one concludes that the sustainability of tacit collusion only depends

on the share in the capital of the largest firm. In addition, some simple algebra shows that:

$$\frac{\partial \delta}{\partial k_1} = \frac{2b^2 (2b + 1) (1 - k_1)}{(1 + b + bk_1)^3} > 0, \quad (2.76)$$

and

$$\frac{\partial^2 \delta}{\partial k_1^2} = 2b^2 (2b + 1) \frac{2b(k_1 - 2) - 1}{(b + 1 + bk_1)^4} < 0. \quad (2.77)$$

Hence, an increase in the size of the largest firm will induce an increase in the minimal threshold on the discount factor above which perfect efficient collusion is sustainable. This is illustrated in Figure 4. The intuition which underlies this result is the following. The larger the largest firm is, the higher will be its share on the one period losses due to the first phase of the punishment strategy. Therefore, the more weight has to be attached to the future stream of payoffs in order for this firm to comply with the punishment strategy.

Let us now turn to the discussion of merger effects on the scope for collusion. We know that before the merger the initial capital allocation belongs to the triangle formed by areas A , B and C in Figure 3, whereas after the merger, the final capital allocation (k_n, k_1) has to lie somewhere on the triangle formed by areas C and D in the same diagram. Hence, two different scenarios should be taken into account. First, we consider the case in which firms were already colluding before the merger takes place. Then, we analyze the situation in which industry-wide collusion was not feasible before the merger.

If before the merger firms were already colluding, this means that the initial capital allocation (k_n, k_1) lies somewhere in regions B or C of Figure 3. When this is the case, any merger leading to an increase in the size of the largest firm will hurt collusion. All other mergers will have no impact on the scope for collusion. In particular, if before the merger $k_n > \frac{1}{2}k_1$, then any merger will affect the size of the largest competitor and, hence,

hurt collusion. This result stresses the fact that a merger induces a more asymmetric post-merger industry configuration when it increases the size of the largest firm.⁶⁶ Thus, this analysis stresses the fact that if collusion already exists in the industry, there may be no room for an improvement in firms' ability to collude after a merger has taken place.

Notice, however, that if before the merger the number of firms is sufficiently high,⁶⁷ then the pre-merger capital allocation might lie in region *A* of Figure 3. If this is the case, then it may be rational for more efficient firms to merge with very small and inefficient firms if by doing so collusion turns out to be enforceable after the merger takes place (that is, if after the merger the smallest firm size $k'_n \geq k^*$). Therefore, according to the present model, only in this case does a merger have anti-competitive effects as the common wisdom would suggest.

As a final remark, notice that (2.76) also shows that a split-off reducing the size of the largest firm in the industry tends to improve the scope for collusion and, hence, should give rise to antitrust concerns since it contributes to a more symmetric distribution of capital shares amongst the existing firms.

Our results therefore reveal that the conventional wisdom that mergers tend to enhance collusion, whereas split-offs have pro-competitive effects, may actually give mis-

⁶⁶ In our framework, merger parties are taken to be exogenous. However, if the analysis was extended to endogenize mergers (split-offs) decisions, one would probably conclude that in the case where collusion was already taking place before the merger, any merger affecting the size of the largest firm would fail to occur in equilibrium. This would be justified by the fact that not only the merger would hurt collusion possibilities, but it would also have no effect on merging firm's profits. Indeed, the sum of the profits of the merging parties equals the (aggregate) profit of the resulting firm (see eq. (2.53)). Notice, however, that in such a situation a split-off of the largest competitor would decrease asymmetries in the distribution of asset holdings and, therefore, would possibly be part of the equilibrium since it would enhance collusion possibilities.

⁶⁷ More precisely, if $n > \frac{2b^3 + (2b+1)^2 - (\sqrt{(2b+1)})^3(b+1)}{(b+1)(b^2 + 3b + 1 - (\sqrt{(2b+1)})^3)}$.

leading predictions about the facility of collusion after an asset transfer takes place, if we disregard the fact that asymmetries in cost functions tend to make coordination amongst oligopolists less likely. More importantly, this analysis clearly suggests that, as was emphasized by Compte, Jenny and Rey (1997), a systematic analysis of market shares and concentration indexes, such as the Herfindahl-Hirshman Index (HHI), does not always provide a reliable guide to assess potential effects on the level of competition in the market induced by a horizontal merger.⁶⁸ Antitrust authorities, when assessing whether a merger between two firms is likely to enhance oligopolistic coordination in the market, should give special attention to firms' cost conditions and to the degree of post-merger symmetry among the firms in the industry.⁶⁹ Even though any merger gives rise to an increase in the size of the merged parties and also reduces the number of competitors operating in the (relevant) market, this fact is not enough to conclude that the scope for collusion increases with the merger. It might well happen that asymmetries outweigh the collusion-enhancing effects of a proposed merger.

2.4 Perfect non-efficient collusion

The previous analysis suggests that in an industry consistent with the assumptions of our model, the smallest and largest firms in the agreement play a crucial role on the stability of perfect efficient collusion. Being allotted a very small share in the collusive output (and profits), small firms may have no incentive to credibly participate in the collusive

⁶⁸ As was highlighted by Fisher (1987), a "serious analysis of market power and oligopoly cannot be subsumed in a few spuriously precise measurements." (p. 39).

⁶⁹ It is important to note that the HHI index tends to *penalize* asymmetry.

agreement. On the other hand, when the equilibrium calls for punishments, the largest firm is the most penalized in the punishment first phase. It is, therefore, natural to wonder whether stability of the collusive agreement could be enhanced if large firms accepted to transfer part of their output share to the smallest ones, therefore allowing small firms to produce a disproportionate (higher) share of the collusive aggregate output. This section deals with precisely this issue by studying a situation in which firms accept a distortion of the efficient allocation rule (2.48) in order to try and enlarge the set of discount factors for which perfect industry-wide collusion is sustainable. Proposition 3 shows that there is a trade-off between efficiency and the stability of perfect collusion.

Proposition 3 *Let $\underline{\delta}$ be as in Proposition 2. There exists a $\delta^{ne} < \underline{\delta}$ such that there exists a subgame perfect equilibrium in which perfect (non-efficient) collusion is enforced for every $\delta \geq \delta^{ne}$.*

Proof. Consider an initial situation in which $k_n = k^*$ and $\delta = \underline{\delta}$, where k^* and $\underline{\delta}$ are defined by equations (2.73) and (2.74), respectively. Hence, we depart from a situation in which the two incentive constraints that matter along an optimal maximal punishment - that of the smallest firm to stick to the collusive output and that of the largest firm to stick to the first phase of the punishment - are just binding. Assume also, without loss of generality, that the number of largest and smallest firms in the market is the same. Let \tilde{k}_i denote the output share of firm i , whereas k_i continues to denote the share in the industry capital owned by firm i , $i \in \{1, \dots, n\}$. Let us now consider an ε -transfer of *output share* from the largest to the smallest firm(s). When this is the case, $\tilde{k}_1 = k_1 - \varepsilon$, $\tilde{k}_n = k_n + \varepsilon$, where

$\varepsilon < k_{n-1} - k_n$, whereas $\tilde{k}_j = k_j$, for all $j \notin \{1, n\}$. Firms are assumed to keep on using a simple proportional two-phase penal code $\sigma(\tilde{q}^1, \tilde{q}^2)$, where $\tilde{Q}^1 = \frac{a-c}{b(1-k_1)}$, $\tilde{Q}^2 = \frac{a-c}{2b+1}$ and $\tilde{q}_i^j = \tilde{k}_i \tilde{Q}^j$, for $i \in \{1, \dots, n\}$, $j = 1, 2$. In order for $\sigma(\tilde{q}^1, \tilde{q}^2)$ to be a perfect equilibrium, one must have that, for all i ,

$$\pi_i^*(\tilde{q}_{-i}^1) - \pi_i(\tilde{q}^1) - \underline{\delta}(\pi_i(\tilde{q}^2) - \pi_i(\tilde{q}^1)) \leq 0, \quad (2.78)$$

and

$$\pi_i^*(\tilde{q}_{-i}^2) - \pi_i(\tilde{q}^2) - \underline{\delta}(\pi_i(\tilde{q}^2) - \pi_i(\tilde{q}^1)) \leq 0, \quad (2.79)$$

where $\tilde{q}_{-i}^j = (\tilde{q}_1^j, \dots, \tilde{q}_{i-1}^j, \tilde{q}_{i+1}^j, \dots, \tilde{q}_n^j)$, $j = 1, 2$, and $\pi_i^* : [0, \frac{a}{b}]^{n-1} \rightarrow \mathbb{R}$ denotes firm i 's best response profit. Since we are assuming an initial situation in which $k_n = k^*$ and $\delta = \underline{\delta}$, then the previous analysis shows that when $\varepsilon = 0$ (initially), condition (2.78) is binding for the largest firm(s) in the market (firm 1), whereas condition (2.79) is binding for the smallest firm in the agreement (firm n). To show that there exists a subgame perfect equilibrium in which firms enforce perfect collusion within a larger set of discount factors by accepting a distortion of the efficient allocation rule (2.48), it suffices to show that the binding incentive constraints (2.78) and (2.79) are both relaxed when there is a transfer of output share from the largest to the smallest firms in the market. Let us start with condition (2.78). This condition is binding for firm 1 when $\varepsilon = 0$. In addition, in this setting, $\pi_1^*(\tilde{q}_{-1}^1) = 0$. Hence, one can rewrite condition (2.78) for the largest firm, obtaining

$$(1 - \underline{\delta})(-\pi_1(\tilde{q}^1)) - \underline{\delta}\pi_1(\tilde{q}^2) \leq 0, \quad (2.80)$$

where $\pi_1(\tilde{q}^j) = (a - c - b\tilde{Q}^j)(k_1 - \varepsilon)\tilde{Q}^j - \frac{((k_1 - \varepsilon)\tilde{Q}^j)^2}{2k_1}$, for $j = 1, 2$. Let $L_1 = (1 - \underline{\delta})(-\pi_1(\tilde{q}^1)) - \underline{\delta}\pi_1(\tilde{q}^2)$. Then, after some algebra and making use of the fact

that $\widetilde{Q}^1 = \frac{a-c}{b(1-k_1)}$, $\widetilde{Q}^2 = \frac{a-c}{2b+1}$ and $\underline{\delta} = \frac{(1+2bk_1)(2b+1)}{(b+1+bk_1)^2}$, one concludes that

$$\frac{\partial L_1}{\partial \varepsilon} = -(a-c)^2 \frac{k_1 - 2\varepsilon}{k_1(2b+1)(b+1+bk_1)} < 0.$$

Hence, the incentive constraint of the largest firm to stick to the first phase of the punishment is relaxed when there is a (marginal) transfer of output share from the largest to the smallest firm(s).

Let us now analyze the effect of the above-mentioned transfer of output share on the incentive constraint of the smallest firm to stick to the collusive output. Condition (2.79) for the smallest firm n is given by

$$\pi_n^*(\widetilde{q}_{-n}^2) - \pi_n(\widetilde{q}^2) - \underline{\delta}(\pi_n(\widetilde{q}^2) - \pi_n(\widetilde{q}^1)) \leq 0, \quad (2.81)$$

where we have that $\pi_n^*(\widetilde{q}_{-n}^2) = \frac{1}{2}(a-c-b(1-(k_n+\varepsilon))\widetilde{Q}^2)^2 \frac{k_n}{1+2bk_n}$ and $\pi_n(\widetilde{q}^j) = (a-c-b\widetilde{Q}^j)(k_n+\varepsilon)\widetilde{Q}^j - \frac{((k_n+\varepsilon)\widetilde{Q}^j)^2}{2k_n}$, for $j = 1, 2$. Now, let $L_n = \pi_n^*(\widetilde{q}_{-n}^2) - \pi_n(\widetilde{q}^2) - \underline{\delta}(\pi_n(\widetilde{q}^2) - \pi_n(\widetilde{q}^1))$. Some algebra shows that $\frac{\partial L_n}{\partial \varepsilon} = (a-c)^2(1+b(1+k_n+\varepsilon))k_n \frac{b}{(2b+1)^2(1+2bk_n)} - (1+\underline{\delta})\left((a-c)^2 \frac{bk_n-\varepsilon}{(2b+1)^2k_n} - \underline{\delta}(a-c)^2 \frac{k_n(1+bk_1)+\varepsilon}{b^2(1-k_1)^2k_n}\right)$.

Now, notice that

$$(a-c)^2(1+b(1+k_n+\varepsilon))k_n \frac{b}{(2b+1)^2(1+2bk_n)} - (a-c)^2 \frac{bk_n-\varepsilon}{(2b+1)^2k_n}$$

can be rewritten as

$$(a-c)^2(1+bk_n) \frac{\varepsilon(1+bk_n) - bk_n(1-k_n)}{(2b+1)^2(1+2bk_n)k_n}.$$

Hence, $\varepsilon < bk_n(1-k_n)/(1+bk_n) < 1$ is a sufficient condition for $\frac{\partial L_n}{\partial \varepsilon} < 0$, i.e., for sufficiently low values of ε , the incentive constraint of the smallest firm to stick to the

collusive agreement is also relaxed by the above-mentioned (marginal) transfer of output share. This completes the proof of Proposition 3. ■

Hence, by accepting a distortion in the optimal allocation rule, firms can enlarge the set of discount factors for which perfect collusion can be credibly enforced.

Two remarks are in order before closing this section. First, notice that by transferring output share from the larger to the smallest firms in the market, this exercise is lessening asymmetries between market participants, thus contributing to increase collusion possibilities. In this way, the previous Proposition confirms and extends the general intuition developed in the previous section. Second, it should be noted that the (extended) proportional two-phase penal code used in the proof of Proposition 3 is not globally optimal. It does not drive *all* firms to their minmax payoff if a deviation occurs, but should drive the largest and smallest firms in the agreement close to the (zero) minmax payoff. Studying alternative output allocation rules as well as proposing other types of optimal punishments within a more general framework will be dealt with in future research.

2.5 Conclusion

This chapter has explored the relationship between the distribution of a tangible industry asset which affects firms' marginal costs and the scope for collusion. In particular, we have found the conditions under which a (perfect) collusive outcome can be enforced when an infinitely repeated game is played between cost asymmetric firms which produce a homogeneous good and adopt optimal punishments in the style of Abreu (1986, 1988) that guarantee a prospective deviant zero profits (the lowest profits consistent with individual

rationality) in the event a deviation occurs. The results obtained embody some important insights for practical application of competition policy.

First, we show that the sustainability of perfect efficient collusion crucially depends on the asset holdings of the firms involved in the agreement. In particular, it has been found that the smallest (and, hence, most inefficient) firm in the agreement, being the one which is allotted the lowest share in the collusive aggregate output, represents the main obstacle for collusion to be enforced because this share may be too low with respect to its optimal deviation output. On the other hand, if the punishment is started, then the largest firm is the one which is proportionally most penalized in the first (severe) phase of the punishment. Therefore, this firm faces the greatest incentives to deviate from the first period of the punishment strategy.

Second, it is shown that if firms accept to diverge from the joint-profit maximization behavior, then they can enhance collusion possibilities. Perfect inefficient collusion can be enforced within a larger set of discount factors. Therefore, in our setting, there exists a trade-off between efficiency and sustainability of perfect collusive agreements.

A distinct issue is also addressed, which is the impact of changes in the distribution of firms' asset holdings on the likelihood of collusion. Some important policy implications can be derived from the results. In particular, they shed some light on the complex problem of assessing the potential *joint dominance* (pro-collusive) effects induced by a merger. It turns out that when asymmetric firms co-exist in the industry the impact of a merger depends on which firms it involves. More specifically, it is shown that a merger can induce two distinct effects. First, if firms in the market were already colluding before the merger,

then a merger either has no effect on the possibility of collusion or it harms that possibility. The latter case will happen when the merger affects the size of the largest firm in the market. This result stresses the fact that although the number of competitors is reduced with the merger, which tends to facilitate collusion, this effect is more than compensated for by a more asymmetric post-merger industry configuration. Second, if before the merger collusion is not feasible, then a merger might make collusion possible afterwards. This will happen when the merger involves very small and inefficient firms that are not able to credibly participate in a collusive scheme before the merger takes place.

This analysis thus suggests that a systematic analysis of market shares and concentration indexes, such as the Herfindahl-Hirshman Index (HHI), does not always provide a reliable guide to assess potential effects on the level of competition in the market induced by a horizontal merger. Once post-merger concentration appears to be high, then, among other things, firms' cost conditions and asset holdings' distribution must be an important part of the analysis, since asymmetries may offset any increased risk of post-merger anti-competitive behavior.

It remains to be seen whether our results are robust to changes in the model assumptions: different, or more general, functional forms for cost and demand, and firm heterogeneity. It would also be interesting to analyze under which conditions a subset of large firms could reach a collusive agreement without involving small firms (partial collusion). All these questions seem to deserve further research.

2.A The Cournot equilibrium

Consider the general case in which inverse demand is given by (2.43) and total cost function of firm i is represented by eq. (2.44). Firm i chooses q_i to maximize profits:

$$\pi_i(q_1, \dots, q_n) = (a - bQ)q_i - \left(cq_i + \frac{q_i^2}{2k_i}\right). \quad (2.82)$$

$\pi_i(\cdot)$ is strictly concave in q_i . Hence, the choice of q_i results from the following first order condition:

$$p - c = \left(b + \frac{1}{k_i}\right) q_i. \quad (2.83)$$

Define $\beta_i = \frac{bk_i}{1+bk_i}$ and $B = \sum_{i=1}^n \beta_i$. Now, from (2.83), it can be shown that $\frac{q_i}{q_j} = \frac{\beta_j}{\beta_i}$, for $i, j = \{1, \dots, n\}$, $i \neq j$. Hence, the following results can be easily derived:

$$q_i^n = \left(\frac{a-c}{b}\right) \frac{\beta_i}{1+B}, \quad (2.84)$$

$$Q^n = \left(\frac{a-c}{b}\right) \left(\frac{B}{1+B}\right), \quad (2.85)$$

$$p = \frac{a+cB}{1+B}, \quad (2.86)$$

$$s_i^n \equiv \frac{q_i}{Q} = \frac{\beta_i}{B}, \quad (2.87)$$

$$\pi_i^n(k_i) = \frac{1}{2} (a-c)^2 \beta_i \frac{2k_i b - \beta_i}{b^2 (1+B)^2 k_i}. \quad (2.88)$$

Chapter 3

Towards a Characterization of the Upper Bound to Concentration in Endogenous Sunk Cost Industries

3.1 Introduction

The purpose of this study is to provide empirically testable implications regarding the relationship between market size and concentration in *endogenous sunk cost industries* (Sutton (1991, 1998)), that is, industries in which firms are involved in research and development (henceforth R&D) activities with the aim of enhancing consumer's willingness to pay for the products they offer.

Following Sutton (1998), this study assumes that the market in which firms operate encompasses a (discrete) set of feasible technological trajectories along which firms can develop their capabilities. Firms are confined to operate along a *single* distinct technological trajectory. In addition, each trajectory is assumed to be associated with a different group of products or submarket. A firm spending more on R&D outlays will be able to improve the quality of the product it offers within the submarket associated with the trajectory along which it is operating.⁷⁰

A fixed cost function, proposed by Sutton (1991, 1998), is introduced. This function maps a spectrum of quality of products offered into an *endogenous* fixed cost which must

⁷⁰ For the sake of simplicity, each firm's technological capability is assumed to be summarized by a single quality attribute.

be incurred and sunk before production begins. Hence, in this setting, each firm's choice set comprises not only the quantity to be produced, but also the quality which the firm aims to achieve along the chosen R&D trajectory.

Sutton (1991,1998) has shown that, under very general conditions, a lower bound exists to the equilibrium level of concentration in *endogenous sunk cost industries*, independently of how large the market is. As pointed out by Bresnahan (1992) and Scherer (2000), a question not addressed there, however, is whether or not an upper bound to the degree of concentration can be characterized in this kind of industries.

In this chapter, this open question is addressed by posing the strategic behavior of firms in terms of a four-stage noncooperative game. At the first stage, firms decide whether or not to enter the industry. Those firms which decided to enter must incur a fixed (sunk) cost of entry ϵ . At the second stage, firms that have entered might form coalitions. All firms that have chosen the same coalition then merge. At the third stage, the newly formed coalitions simultaneously choose the quality level of the product they offer and incur an associated fixed cost. Finally, at the fourth and last stage of the game, firms/coalitions compete by choosing output levels in accordance with the usual Cournot oligopoly model. If a merger occurs, then the involved firms share the final stage profit and also the cost associated with the investment in fixed outlays. Therefore, in this setting, firms have a double incentive to participate in a merger. Apart from reducing competition in the market, a merger allows the involved firms to realize a cost advantage over the unmerged rivals. When firms merge, they afford a cost reduction through the elimination of duplication efforts in their R&D investments and, hence, the incentive for merger is reinforced. As pointed out by Kamien and

Zang (1993), "despite the systematization of research and development (R&D), there still remains a significant component of trial and error. Common research activity allows for more efficient weeding out of fruitless approaches and concentration on the most promising." (p. 24).

As far as the coalition formation game is concerned, we adopt a (coalitional) stability concept which differs from standard stability concepts like Nash stability in two ways. First, it also requires stability with respect to deviations by a group of players/firms. Second it rules out the assumption that once a coalition deviation occurs, it cannot be followed by subsequent deviations,⁷¹ an assumption which is unwarranted. By allowing players/firms to look far ahead, the proposed stability concept addresses two related questions: (i) whether a deviation is vulnerable to further deviations; and (ii) which is the *ultimate* result a deviation will lead to. Obviously, an *ultimate* result must, by definition, be immune to further deviations. In addition, it must be part of the solution set of the coalition formation game. This study thus relates to a strand of the literature on farsighted stability. A sample of recent works on this area, though in different contexts, includes the papers by Chwe (1994), Xue (1998) and Diamantoudi and Xue (2001). We also confront the predictions obtained applying the Nash stability concept with the ones obtained when using the proposed concept where players are assumed to be endowed with foresight. By so doing, one concludes that Nash stability fails to rule out a large set of coalition structures due to players' myopic behavior embedded in this notion.

⁷¹ Nash stability, as well as other standard stability concepts, considers only one-step deviations by firms/players.

Analysis of the (refined) subgame perfect Nash equilibria (henceforth, (refined) SPE) of this game discloses that,⁷² when comparing this model with a basic model in which firms are not allowed to merge, the anticipation of the possibility to form coalitions gives rise to a pattern of excessive entry in the earliest stage of the game. The industry is subsequently whittled down by the (endogenous) mergers formation game, through which firms cooperate in their R&D efforts to enhance products' quality. In particular, it is shown that there exists a threshold value for the parameter measuring the degree of product substitutability above which a merger to monopoly is the equilibrium outcome even in the limit when market size goes to infinity. Therefore, this study shows that in *endogenous sunk cost industries*, an upper bound to concentration exists and is independent of the size of the market. Interestingly, in some equilibria where a merger to monopoly arises as the unique equilibrium outcome, it turns out that firms belonging to the monopoly coalition structure earn strictly positive profits even under the threat of entry. In addition, it is also shown that if, instead, the degree product substitutability is below the above mentioned threshold value but still product substitutability is not too low, then only duopoly coalition structures composed of sufficiently size asymmetric coalitions can be sustained in a (refined) SPE. The intuition that underlies this result is simple. The larger the size of the smaller coalition in the duopoly coalition structure is, the more likely is that firms in that specific coalition will be willing to participate in a (joint) deviation towards complete monopolization of the industry, therefore ruling out the stability of the coalition structure they belong to.

⁷² Remember that, when solving the game backwards, we apply at the second stage of the game the proposed farsighted stability concept instead of the standard Nash stability concept.

To the best of our knowledge, the only paper that addresses the question of whether there exists an upper bound to concentration in *endogenous sunk cost industries* is Nocke (2000). His predictions are, however, not entirely satisfactory because of the following reasons. First, the coalition formation game is modelled as a noncooperative open membership game (see Yi (1997)) and, therefore, firms are implicitly assumed to be myopic when taking their mergers decisions. Second, and most importantly, due to several reasons explained in the paper, Nocke is not able to solve for all equilibria at the investment (third) stage of his four-stage game. As a result, instead of solving the overall game initially proposed, the paper analyzes a “constrained game” in which firms are *not* allowed to merge. By so doing, the chapter contributes to the characterization of the *lower-bound* to the one firm concentration ratio that can arise in equilibrium in *endogenous sunk costs industries*.

The rest of the chapter is organized as follows. Section 2 presents the general model and the coalition formation game. This section also characterizes and discusses the outcomes that can arise in a (refined) SPE of the overall game. Finally, section 3 offers some concluding comments.

3.2 The basic model

Following Sutton (1998), let all consumers have the same utility function defined over n substitute goods (or n varieties of the same product), and a separate outside good whose price is fixed exogenously at unity. Let the utility function take the form

$$U = \sum_k \left(x_k - \frac{x_k^2}{u_k^2} \right) - 2\sigma \sum_k \sum_{l < k} \frac{x_k x_l}{u_k u_l} + M, \quad (3.89)$$

where x_k is the quantity and u_k is the quality level of good k , M denotes the consumption of the outside good and the parameter σ , $0 \leq \sigma \leq 1$, measures the degree of substitution between the goods. When $\sigma = 0$, product varieties are independent in demand, whereas as $\sigma \rightarrow 1$, existent varieties become closer and closer substitutes.

For the general utility function (3.89), the individual consumer's inverse demand for good k becomes,

$$p_k = 1 - \frac{2x_k}{u_k^2} - \frac{2\sigma}{u_k} \sum_{l \neq k} \frac{x_l}{u_l}. \quad (3.90)$$

Notice that as $u_k \rightarrow 0$, demand falls to zero for any given $p_k \geq 0$. On the other hand, an increase in u_k , shifts the (linear) demand outward, that is, enhances the consumers' willingness to pay for good k .

Assuming that there are S identical consumers in the market, market demand for good k is just a multiple S of individual demand. Simplifying notation, use x_k to denote the per-capita output of good k so that total output of good k is written as Sx_k .

3.2.1 The game

Consider now a four-stage game as follows. At the first stage, a sufficiently large number of *ex-ante* identical firms N_0 simultaneously decide whether or not to enter the market. The entry cost is denoted by ε .⁷³ We assume that the number of potential entrants N_0 at stage one is large enough so that at least one player will decide not to enter the market at this stage. At the second stage, firms that have decided to enter at the preceding stage, simultaneously choose which "coalition" to join. All firms that have decided to join the

⁷³ Of course, we assume in what follows that the market is viable, in the sense that the entry cost ε is no larger than the monopoly profits.

same coalition then merge. At the third stage, the newly formed coalitions simultaneously choose a quality level u_i for the *single* product they offer and incur a fixed cost

$$F(u_i) = u_i^\beta, \quad (3.91)$$

where β is the elasticity of the investment function with respect to the quality level u_i . We restrict qualities to the interval $[1, \infty)$ and $\beta > 2$. Hence, a firm (say, firm i) must incur a minimum level of cost of one to enter the industry with a product of (minimal) quality level $u_i = 1$. Increases in spending in R&D beyond this minimal level are associated with increases in the quality of the product offered by the firm. Finally, there is a fourth stage in which firms compete *à la* Cournot.

Assume all the 'quality goods' are produced at some constant marginal cost c ,⁷⁴ which we normalize to zero. Assume also that each firm operates along a distinct R&D trajectory and produces one product, each trajectory being occupied by just one firm.⁷⁵ In the Nash equilibrium in quantities (Cournot equilibrium), the profit function in a market of size S in which N firms (products) have positive sales at equilibrium is given by:⁷⁶

$$S\pi(u_i | N, \bar{u}) = S \frac{1}{2} \left\{ \frac{u_i + \frac{\sigma}{2-\sigma} N (u_i - \bar{u})}{2 + (N-1)\sigma} \right\}^2, \quad (3.92)$$

where u_i denotes the quality of firm i 's product and \bar{u} denotes the average quality of all products that have positive sales at equilibrium, $\bar{u} = (\sum_i u_i) / N$.⁷⁷

⁷⁴ Following Sutton (1991, 1998), this assumption implies that the burden of quality improvement falls primarily on fixed cost.

⁷⁵ Hence, σ is thought of as a measure of the degree of substitution between products associated with *different* technological trajectories (submarkets, product groups).

⁷⁶ For the derivation of this reduced form profit function, see Appendix 2.2 in Sutton (1998).

⁷⁷ As can be seen from (3.92), the profit function increases with u_i^2 . Hence, the restriction $\beta > 2$ ensures that $F(u_i)$ rises with u_i at least as rapidly as profit.

The merger formation

In the merger formation stage, stage 2, each of the N firms ($1 \leq N \leq N_0 - 1$) that have committed to enter the market at stage 1, decides whether or not to form coalitions. In particular, each firm $i \in \{1, \dots, N\}$ simultaneously announces a list of players (including itself) with whom it is willing to form a coalition. The firms that announce *exactly* the same list of firms form a coalition. This coalition formation game was first proposed by Hart and Kurz (1983).

In formal terms, each firm i 's strategy is to choose a set of firms S^i , which is a subset of the set of entrants $\{1, \dots, N\}$ and includes firm i . The set of strategies for firm i is, therefore, $\Sigma^i = \{S \subset \{1, \dots, N\} | i \in S\}$. Given firms's announcements $\alpha \equiv (S^1, \dots, S^N)$, the resulting coalition structure is $C = \{C_1, \dots, C_T\}$, where T denotes the number of *different* lists chosen by the N entrants. $C_i \cap C_j = \emptyset$ for $i \neq j$ and $\cup_{i=1}^T C_i = \{1, \dots, N\}$. Firms i and j belong to the same coalition C_k if and only if $S^i = S^j$. Notice, however, that S^i (respectively, S^j) is the largest set of firms firm i (respectively, firm j) would be willing to be associated with in the same coalition. As a result, the coalition C_k may in general be different from S^i (respectively, S^j). A coalition, in this setting, corresponds to an equivalent class, with respect to equality of strategies.

Each of the N_0 firms is assumed to maximize its own profit. The same applies to the merged entities at the quantity setting stage. The members of each coalition share both the final stage profit and the investments in R&D outlays. Since all firms are symmetric, we assume that profit and investment outlays are equally shared. That is, any side payments are ruled out with respect to membership decisions.

3.2.2 Equilibrium analysis

In this section we look for a (refined) SPE of this game, following the usual backward induction procedure, starting with the last stage of the game where the merged entities compete in quantities.

Quantity-setting stage

Suppose that at the end of stage two, we have N_2 firms/coalitions of firms, where $N_2 \leq N < N_0$. Hence, from (3.92) one has that the profit earned by a firm offering quality u_i in the final stage sub-game is given by

$$S\pi(u_i | N_2, \bar{u}) = S \frac{1}{2} \left\{ \frac{u_i + \frac{\sigma}{2-\sigma} N_2 (u_i - \bar{u})}{2 + (N_2 - 1) \sigma} \right\}^2, \quad (3.93)$$

where $\bar{u} = (\sum_i u_i) / N_2$.

Investment stage

Armed with the above expression for profits, it is now possible to look at firm's quality choice. Notice that we are only interested in a symmetric outcome in which, at stage 3, $\forall i \in \{1, \dots, N_2\}$, $u_i = v$.⁷⁸ Besides, N_2 is taken as a parameter since it is the outcome of firms' choices made at stage 2.

As pointed out by Sutton (1998), the equilibrium involves two regimes that correspond to low and high values of the market size parameter S . In particular, at stage 2, the equilibrium outcome takes one of two forms, according as whether an internal or corner

⁷⁸ Notice that fixed costs apart, coalitions of firms operate as independent entities which are similar in all other aspects. Therefore, it seems natural to focus on a symmetric equilibrium at the investment stage.

solution applies, that is:

$$S \frac{\partial \pi(u_i | N_2, \bar{u})}{\partial u_i} \Big|_{u_i=v=1} \leq \frac{\partial F(u_i)}{\partial u_i} \Big|_{u_i=v=1} \quad (3.94)$$

or otherwise.

Hence, if market size (as measured by S) is sufficiently small, then condition (3.94) holds and, therefore, at equilibrium all firms that have entered the market at stage 1 will choose $v = 1$. If, instead, market size is sufficiently large, then the above condition fails to hold and at equilibrium all the N_2 coalitions of firms formed at stage 2 will choose a value of $v > 1$ determined by the following condition

$$S \frac{\partial \pi(u_i | N_2, \bar{u})}{\partial u_i} \Big|_{u_i=v} = \frac{\partial F(u_i)}{\partial u_i} \Big|_{u_i=v}. \quad (3.95)$$

The critical level of S is defined by writing condition (3.94) as an equality. In what follows, we assume that S is above this critical value so that $v > 1$ at equilibrium. From (3.93), on writing $N_2 \bar{u} = u_i + (N_2 - 1)v$, one obtains that

$$S \frac{\partial \pi(u_i | N_2, \bar{u})}{\partial u_i} \Big|_{u_i=v} = S \frac{v(2 + \sigma(N_2 - 2))}{(2 + \sigma(N_2 - 1))^2(2 - \sigma)}, \quad (3.96)$$

or, equivalently

$$S \frac{\partial \pi(u_i | N_2, \bar{u})}{\partial u_i} \Big|_{u_i=v} = S \frac{1}{v} \pi(v | N_2, v) \frac{2(2 + \sigma(N_2 - 2))}{(2 - \sigma)}. \quad (3.97)$$

On the other hand, from (3.91) it is straightforward to show that

$$\frac{\partial F(u_i)}{\partial u_i} \Big|_{u_i=v} = \frac{1}{v} \beta F(v). \quad (3.98)$$

Now, making use of expressions (3.97) and (3.98) to write (3.95) in explicit form, one obtains

$$S \pi(v | N_2, v) \frac{2(2 + \sigma(N_2 - 2))}{(2 - \sigma)} = \beta F(v), \quad (3.99)$$

where this equation defines the level of fixed outlays incurred by firms/coalitions as a function of the number of active firms/coalitions at the end of stage 2, N_2 ($N_2 \leq N < N_0$). In what follows, let $F^*(N_2; S, \sigma)$ denote the value of F implicitly defined by equation (3.99). The following Lemma summarizes the equilibrium behavior at the investment stage.

Lemma 1 *At a symmetric equilibrium, in which $\forall i \in \{1, \dots, N_2\}$, $u_i = v$,*

$$v = \left\{ \frac{S}{\beta} \frac{2 + \sigma(N_2 - 2)}{(2 - \sigma)(2 + (N_2 - 1)\sigma)^2} \right\}^{\frac{1}{\beta-2}},$$

where v decreases in N_2 , the number of active firms/coalitions at the end of stage 2.

Proof. Notice that one can use (3.91) and (3.93) to express (3.99) in explicit form and obtain a solution for the equilibrium value of v as a function of the parameters of the model.

By so doing, one gets

$$v = \left\{ \frac{S}{\beta} \frac{2 + \sigma(N_2 - 2)}{(2 - \sigma)(2 + (N_2 - 1)\sigma)^2} \right\}^{\frac{1}{\beta-2}}. \quad (3.100)$$

As mentioned above, N_2 is a parameter at this stage. It results from firms' decisions at stage 2. One can now use (3.100) to study how the equilibrium value of v reacts to a change in the number of firms/coalitions active at the end of stage 2, N_2 .

$$\frac{\partial v}{\partial N_2} = \frac{-\sigma(2 + \sigma(N_2 - 3))}{(2 + \sigma(N_2 - 2))(2 + \sigma(N_2 - 1))(\beta - 2)} \left(\frac{S(2 + \sigma(N_2 - 2))}{\beta(2 + \sigma(N_2 - 1))^2(2 - \sigma)} \right)^{\frac{1}{\beta-2}} < 0. \quad (3.101)$$

Hence, an increase in the number of firms/coalitions (as measured by N_2) induces a decrease in the (symmetric) equilibrium quality level v . ■

It is worth remarking at this point that from (3.100), one can see that increases in the market size S are associated with an increasing (common) level of the quality offered

at equilibrium by firms/coalitions in the market at stage 3.⁷⁹ This feature of the model will play a central role later on in explaining why in endogenous sunk costs industries the industry structure does not converge to a fragmented outcome, no matter how large the market becomes (as shown in Sutton (1998)).

Merger formation stage

We can now consider firms' merger decisions at the second stage of the game. By participating in a coalition, firms share both investments in R&D (therefore, avoiding duplication of efforts) and the profits obtained at the quantity competition final stage.

The nonfragmentation result

Notice that since we are seeking a symmetric perfect equilibrium of this game, it follows that, at stage 3, all firms will choose the same level of quality and, therefore, incur the same level of fixed outlays $F^*(N_2; S, \sigma) \geq 1$. Besides, all coalitions active at the end of stage 2 expect to earn nonnegative profits, which implies that N_2 must satisfy the following viability condition

$$S\pi(v|N_2, v) - F(v) \geq 0. \quad (3.102)$$

Now, combining equations (3.99) and (3.102), one finds an upper bound for the equilibrium number of firms/coalitions (lower bound to concentration) at the end of stage 2

$$N_2 \leq \frac{\beta(2 - \sigma) - 4(1 - \sigma)}{2\sigma} \equiv \bar{N}_2(\beta, \sigma). \quad (3.103)$$

This is the *nonfragmentation result*. It shows that, as long as σ is strictly positive, so that there is some degree of substitution among products, the lower bound to the one

⁷⁹ Remember that we have assumed that market size is sufficiently high so that investment in R&D outlays is profitable (see discussion of eqs. (3.94) and (3.95)).

firm concentration ratio is bounded away from zero in *endogenous sunk cost industries*, no matter how large the market is (see Sutton (1991)).⁸⁰ The intuition underlying this result is the following. Consider a configuration in which, in the limit, as market size becomes indefinitely large, the number of firms/coalitions rises without bound. If this is the case, then firms/coalitions will all have very small market shares and their revenues, relative to market size S , will be small as well. Now, from (3.102), each firm/coalition's revenue should at least cover its fixed outlays. As a consequence, in such a configuration, each firm/coalition's fixed outlay would be very small. This could not be an equilibrium. Some firm/coalition would find it optimal to break this fragmented outcome by increasing its fixed outlays (and, hence, quality), outspending its rivals, thereby capturing a large share of the market. As a result, the outcome of this process is that, in the limit, as market size increases, firms/coalitions' investment outlays rise proportionally with market size and the number of firms/coalitions which can be supported in equilibrium becomes limited and independent of market size.

One can now use (3.103) to carry out a simple exercise of comparative statics with respect to the substitution parameter σ and to the cost parameter β , obtaining

$$\frac{\partial \bar{N}_2(\beta, \sigma)}{\partial \sigma} = -\frac{\beta - 2}{\sigma^2} < 0, \quad (3.104)$$

$$\frac{\partial \bar{N}_2(\beta, \sigma)}{\partial \beta} = \frac{2 - \sigma}{2\sigma} > 0. \quad (3.105)$$

Hence, as pointed out by Sutton (1998), for a given value of β , a fall in the degree of substitution between products associated with different technological trajectories shifts

⁸⁰ Notice that if $\sigma \rightarrow 0$, the upper bound for the equilibrium number of firms/coalitions at the end of stage 2 rises without bound ($\bar{N}_2(\beta, \sigma) \rightarrow \infty$).

the incentives away from an escalation of spending on R&D outlays towards the introduction of new varieties. Therefore, the lower bound to the one firm concentration ratio, $\underline{C}_1 = 1/(\overline{N}_2(\beta, \sigma))$, increases in the substitutability parameter σ . On the other hand, for any fixed value of σ , a decrease in the parameter β leads to an increase in the effectiveness of R&D investments since it lowers the cost of improving technical performance (see equation (3.91)). This encourages escalation investments on R&D to improve technical performance and, hence, tends to increase concentration levels at equilibrium. In other words, \underline{C}_1 decreases in the effectiveness of R&D parameter β .

In what follows, we study the individual firms' incentives towards the participation in coalitions. In particular, we characterize the stable coalition structures that can arise under two alternative stability notions. First, we analyze the coalitions that can form at stage 2 by applying the well known Nash stability concept. We then proceed by introducing an alternative (coalition) stability notion that amends the firms' myopic behavior embedded in the Nash stability notion, by assuming that players are *farsighted*. It turns out that the latter stability concept enables us to make much sharper predictions than the former one with regards to the coalitions that are likely to form in equilibrium.

Nash stability

Let $\Pi(C_i; C)$ denote the *per member* equilibrium payoff of a coalition C_i under the coalition structure C . In addition, for a given coalition structure C , let $C_j(i)$ denote the set of coalitions to which firm i could migrate through an individual deviation. More formally, if initially $i \in C_k \subset C$, $C_j(i) = \{C_j \in C, C_j \neq C_k | \forall t \in C_j, i \in S^t\}$.

The following definition turns out to play a central role in the identification of equilibrium coalition structures.

Definition 1 *Stand-alone stability (cf. Yi (1997)):* $C = \{C_1, \dots, C_T\}$ is *stand-alone stable* if and only if $\forall k \in \{1, \dots, T\}, \forall i \in C_k, \Pi(C_k; C) \geq \Pi(\{i\}; C'_i)$, where $C'_i = C \setminus C_k \cup \{C_k \setminus \{i\}, i\}$.

A coalition structure C is stand-alone stable if and only if no firm can unilaterally improve its payoff by forming a singleton coalition, holding the rest of the coalition structure (including its former coalition) fixed.

Definition 2 *Nash stability:* A coalition structure $C = \{C_1, \dots, C_T\}$ is the outcome of a Nash equilibrium if there do not exist $i \in C_k, k \in \{1, \dots, T\}$, and $S \in C_j(i) \cup \{\emptyset\}$ such that $\Pi(S \cup \{i\}; C) \geq \Pi(C_k; C^N)$, where $C^N = C \setminus \{C_k, S\} \cup \{C_k \setminus \{i\}, S \cup \{i\}\}$.

Hence, a coalition structure C is said to be the outcome of a Nash equilibrium if no player has incentives to either (individually) migrate to another coalition also in C whose members permit him to do so or to stay alone in a new singleton coalition.

Proposition 1 *Let $C = \{C_1, \dots, C_T\}$ be the coalition structure induced by the vector of firms' announcements $\alpha \equiv (S^1, \dots, S^N)$. If $\forall k \in \{1, \dots, T\}, \forall i \in C_k, S^i = C_k$, then the coalition structure C is the outcome of a Nash equilibrium of the (endogenous) coalition game if and only if C is stand-alone stable.*

Proof. First, suppose the coalition structure C is stand-alone stable. Since we have assumed that $\forall k \in \{1, \dots, T\}, \forall i \in C_k, S^i = C_k$, no firm $i \in C_k$ can join an alternative coalition $C_{k'}$ in $C, k' \neq k$, by unilaterally deviating, i.e., $C_j(i) = \{\emptyset\}$. The only feasible individual deviation for a firm is to leave its own coalition and form a singleton coalition, holding the rest of the coalition structure unchanged. However, C is assumed to be stand-alone stable and, therefore, no player finds it optimal to form a one-firm coalition. Next, suppose the coalition structure C is not stand-alone stable. If this is the case, there exists at least one firm (say, firm j) which can benefit by leaving its coalition in order to form a singleton coalition, holding the rest of the coalition structure fixed. Let firm j belong to coalition C_l , i.e., $S^j = C_l$, where $l \in \{1, \dots, T\}$. If j leaves coalition C_l , by changing its announcement to $\{j\}$, then the induced coalition structure is $C'_j = C \setminus C_l \cup \{C_l \setminus \{j\}, j\}$. Now, firm j is better off after its unilateral deviation, because, by hypothesis, $\Pi(\{j\}; C'_j) > \Pi(C_l; C)$. This completes the proof. ■

The next Corollary shows the existence of an unmerged Nash equilibrium for the coalition formation game, that is, an equilibrium in which there are no mergers.

Corollary 1 *For every number of entrants $N < N_0$, the coalition structure $\{\{1\}, \{2\}, \dots, \{N\}\}$ constitutes a Nash equilibrium to the coalition formation game.*

Proof. By definition, the coalition structure $\{\{1\}, \{2\}, \dots, \{N\}\}$ is stand-alone stable. Now, applying Proposition 1, one has that $\{\{1\}, \{2\}, \dots, \{N\}\}$ is the outcome of a Nash equilibrium in the coalition formation game. ■

Hence, this framework encompasses as a special case the degenerate coalition structure $\{\{1\}, \{2\}, \dots, \{N\}\}$ in which each coalition is composed of a single member (firm).

In what follows, for simplicity of analysis, let us focus on the case in which the industry is partitioned into coalitions of the same size (say, m members, where $2 \leq m \leq N$). When this is the case, according to the stand-alone stability requirement, merger behavior in the second stage will depend on the trade-off between the profits that each player anticipates to earn by participating in one of the N/m coalitions with the ones he would obtain as the lone holdout. Remember that firms/coalitions are assumed to operate along a single technological trajectory, on the one hand, and we are seeking a symmetric SPE, in which at the investment stage 3, $\forall i \in \{1, \dots, N_2\}$, $u_i = v$, on the other. These assumptions imply that the per-member equilibrium profit $\Pi(C_i; C)$ of a coalition C_i under coalition structure C will, in this setting, be characterized by the number and the size of the coalitions in the market. In particular, let $C^s = \{\{1, \dots, m\}, \{m+1, \dots, 2m\}, \dots, \{N-m+1, \dots, N\}\}$. Then, $\Pi(\{1, \dots, m\}; C^s) = \Pi(\{m+1, \dots, 2m\}; C^s) = \dots = \Pi(\{N-m+1, \dots, N\}; C^s) = \frac{1}{m} [S\pi(v | \frac{N}{m}, v) - F^*(\frac{N}{m}; S, \sigma)] - \varepsilon$, while $\Pi(\{i\}; C_i^{s'}) = S\pi(v | \frac{N}{m} + 1, v) - F^*(\frac{N}{m} + 1; S, \sigma) - \varepsilon$, where, for all $i \in \{1, \dots, N\}$, $i \in C_k \subset C^s$ and $C_i^{s'} = C^s \setminus C_k \cup \{C_k \setminus \{i\}, i\}$. Hence, applying Definition 1, the coalition structure C^s is said to be stand-alone stable if and only if the following incentive compatibility constraint holds

$$\frac{1}{m} \left[S\pi \left(v \left| \frac{N}{m}, v \right. \right) - F^* \left(\frac{N}{m}; S, \sigma \right) \right] \geq S\pi \left(v \left| \frac{N}{m} + 1, v \right. \right) - F^* \left(\frac{N}{m} + 1; S, \sigma \right), \quad (3.106)$$

or, equivalently

$$\frac{1}{m} \left(\frac{S\pi \left(v \left| \frac{N}{m}, v \right. \right)}{F^* \left(\frac{N}{m}, S, \sigma \right)} - 1 \right) \geq \frac{F^* \left(\frac{N}{m} + 1; S, \sigma \right)}{F^* \left(\frac{N}{m}, S, \sigma \right)} \left(\frac{S\pi \left(v \left| \frac{N}{m} + 1, v \right. \right)}{F^* \left(\frac{N}{m} + 1, S, \sigma \right)} - 1 \right). \quad (3.107)$$

Now, from (3.99), one has that

$$\frac{S\pi(v|N_2, v)}{F^*(N_2; S, \sigma)} = \frac{\beta(2 - \sigma)}{2(2 + \sigma(N_2 - 2))}. \quad (3.108)$$

Combining (3.107) and (3.108), one can rewrite the incentive compatibility constraint as follows

$$\frac{1}{m} \left(\frac{\beta(2 - \sigma)}{2(2 + \sigma(\frac{N}{m} - 2))} - 1 \right) \geq \frac{F^*(\frac{N}{m} + 1; S, \sigma)}{F^*(\frac{N}{m}; S, \sigma)} \left(\frac{\beta(2 - \sigma)}{2(2 + \sigma(\frac{N}{m} - 1))} - 1 \right). \quad (3.109)$$

In what follows, we investigate whether, in *endogenous sunk cost industries*, arbitrarily concentrated outcomes, namely monopoly or a symmetric duopoly, can be sustained as outcomes of a Nash equilibrium of the coalition formation game. In particular, next Proposition shows that complete monopolization of the industry at stage 2 can only occur in equilibrium as long as the effectiveness of R&D investments (as measured by β) is sufficiently high or there is *no excessive entry* at the entry first stage of the game.

Proposition 2 *A coalition structure composed of a single ‘grand’ coalition $\{1, \dots, N\}$ is the outcome of a Nash equilibrium of the coalition formation game if and only if:*

- $2 < \beta \leq 4/(2 - \sigma)$, or
- $\beta > 4/(2 - \sigma)$ and $N \leq \overline{N}(\beta, \sigma)|_{N=m}$, where

$$\overline{N}(\beta, \sigma)|_{N=m} \equiv \left(\frac{(2 + \sigma)^2(2 - \sigma)}{8} \right)^{\frac{\beta}{\beta-2}} \frac{2(\beta - 2)}{\beta(2 - \sigma) - 4}.$$

Proof. In the case of a merger to monopoly, $m = N$ and $\forall i \in \{1, \dots, N\}$, $S^1 = \dots = S^N = \{1, \dots, N\}$. Hence, $\forall i \in \{1, \dots, N\}$, $C_j(i) = \{\emptyset\}$. The only feasible deviation for any firm i is to leave the ‘grand’ coalition and form a singleton coalition.

Applying Proposition 1 and re-evaluating condition (3.109) for the case in which $m = N$, some algebra shows that the 'grand' coalition is sustainable as a Nash equilibrium of the coalition formation game if and only if

$$\frac{1}{N}(\beta - 2) \geq \frac{F^*(2; S, \sigma)}{F^*(1; S, \sigma)} \left(\frac{\beta(2 - \sigma) - 4}{2} \right). \quad (3.110)$$

Two different cases should now be considered. First, consider the case in which $2 < \beta \leq 4/(2 - \sigma)$. When this is the case, the l.h.s. of condition (3.110) is always non-negative, while the r.h.s. is always non-positive.⁸¹ Hence, condition (3.110) is trivially satisfied. Second, if $\beta > 4/(2 - \sigma)$, then some algebra shows that condition (3.110) holds if and only if

$$N \leq \frac{F^*(1; S, \sigma)}{F^*(2; S, \sigma)} \left(\frac{2(\beta - 2)}{\beta(2 - \sigma) - 4} \right). \quad (3.111)$$

One can now use (3.91) and (3.100) to express $F^*(1; S, \sigma)/F^*(2; S, \sigma)$ in explicit form. By doing so, it turns out that

$$\frac{F^*(1; S, \sigma)}{F^*(2; S, \sigma)} = \left(\frac{(2 + \sigma)^2(2 - \sigma)}{8} \right)^{\frac{\beta}{\beta - 2}}. \quad (3.112)$$

Thus, combining (3.111) and (3.112), one concludes that, when $\beta > 4/(2 - \sigma)$, a merger to monopoly can be supported in equilibrium if and only if

$$N \leq \left(\frac{(2 + \sigma)^2(2 - \sigma)}{8} \right)^{\frac{\beta}{\beta - 2}} \frac{2(\beta - 2)}{\beta(2 - \sigma) - 4} \equiv \bar{N}(\beta, \sigma)|_{N=m}. \quad (3.113)$$

This completes the proof. ■

Figure 1 shows the relevant viability regions in the (σ, β) space and contributes to the understanding of the intuition which underlies the result in Proposition 1.

⁸¹ Remember that $F^*(N_2; S, \sigma) \geq 1$.

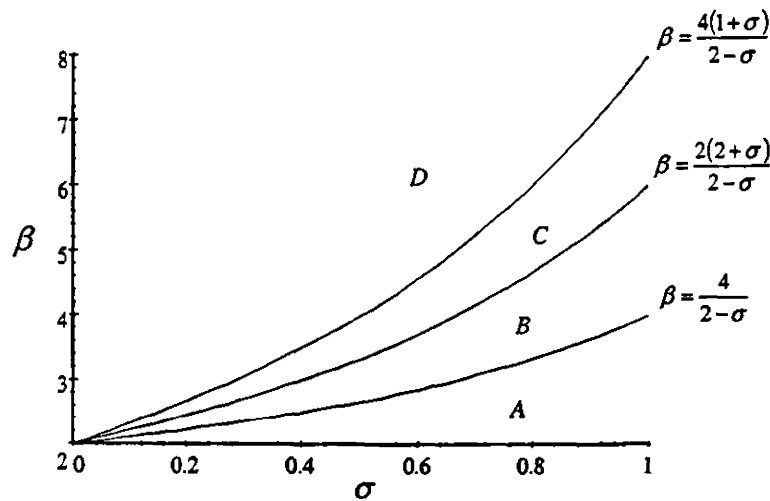


Figure 1 - Viability Regions

The different solid lines were constructed making use of eq. (3.103). For each region between these lines, one can identify the maximum number of coalitions of firms which are viable (i.e., earn non-negative payoffs) at equilibrium. In region *A*, one can have at most one coalition earning non-negative profits at equilibrium. In regions *B* and *C*, the maximum number of viable coalitions at equilibrium is two and three, respectively.

Let us start by considering the case in which $2 < \beta \leq 4/(2 - \sigma)$ (region *A* in Figure 1). If more than one firm enter the market at stage one, then a single 'grand' coalition is formed at stage 2. Two different effects, implied by the viability condition, justify this result. First, a merger to monopoly is the only rational action for firms in the market at stage 2. Both in the situation where there are no mergers and in the case where mergers occur but more than one coalition is formed, firms/coalitions in the market end up earning negative payoffs. Second, and as a consequence, once the 'grand' coalition is formed, no firm has incentives to unilaterally deviate from its strategy. Free-riding on the $N - 1$ merging rivals

would never be a best reply for a firm since the deviator would earn negative profits at equilibrium.

Consider now the case where $\beta > 4/(2 - \sigma)$. A merger to monopoly leads to the formation of a single 'grand' coalition with N member firms. By sticking to the strategy of joining this 'grand' coalition, a firm belonging to the initial wave of N entrants will get a share $1/N$ of the coalition overall profit, whereas by free-riding on its $N - 1$ merging rivals it can obtain the duopoly profit (net of the respective fixed costs). As a consequence, the higher the number of entrants at stage 1, N , the greater the incentives to free-ride, which in turn implies that the merger to monopoly is more difficult to sustain as a Nash equilibrium of the coalition formation game. In other words, complete monopolization is found to be infeasible when the number of firms entering at stage 1 becomes sufficiently large.^{82,83}

Figure 1 can also be used to illustrate the comparative statics results provided by eqs. (3.104) and (3.105). First, take some fixed $\beta > 2$ and allow σ to vary in its domain $[0, 1]$. If goods associated with different trajectories become closer substitutes in consumption, then escalation is profitable. A firm spending heavily in R&D outlays along one trajectory will be able to capture sales in submarkets other than its own ("business stealing effect"). Hence, if σ is high, at equilibrium, only a limited number of firms will survive. Concentration cannot be low. Second, for a given value of $\sigma \in [0, 1]$, an increase in the effectiveness of R&D (i.e., a decrease in the cost parameter β) induces a decrease in the upper bound to the

⁸² In the limiting case where the cost parameter $\beta \rightarrow \infty$, so that it becomes arbitrarily expensive to raise the quality level above unity (*exogenous sunk cost industries*), condition (3.113) discloses that no merger to monopoly will occur in an SPE as long as $\sigma < 0.82843$.

⁸³ Properties of the upper bound to the number of entrants $\bar{N}(\beta, \sigma)|_{N=m}$, given by (3.113), will be studied later on (see Figure 2).

number of active firms at the end of stage 2 of the game, $\overline{N}_2(\beta, \sigma)$. In particular, as already mentioned, if, for a given $\sigma \in [0, 1]$, $2 < \beta \leq 4/(2 - \sigma)$, then only *one* coalition will be active at the end of stage 2 no matter how large the market becomes and, hence, the one firm concentration ratio attains its maximum.

Notice that, in order to further characterize the stable coalition structures in our setting, we have to consider now the cases in which mergers do not lead to a complete monopolization of the industry. In particular, we will now turn to the analysis of the conditions under which a stable coalition structure composed of two individual coalitions of the same size (merger to a symmetric duopoly) can arise as the outcome of a Nash equilibrium of the coalition formation game. In what follows, firms are allowed to adopt any strategy profile which supports a symmetric duopoly as an outcome.⁸⁴ Hence, apart from stand-alone stability, one has to check whether there are incentives for firms to (unilaterally) migrate to the other coalition in the same coalition structure, when that is feasible.

The next Lemma puts forward the conditions in order for a duopoly coalition structure composed of equal sized coalitions to be stand-alone stable, whereas Proposition 3 checks whether there are incentives for firms organized in such a coalition structure to migrate between coalitions, when they are allowed to do so by firms in the welcoming coalition.

Notice that the analysis that follows obviously assumes that $\beta \geq 4/((2 - \sigma))$ (i.e., we are now considering regions B, C and D of Figure 1).⁸⁵

⁸⁴ In particular, we are not restricting the analysis to the strategy profiles considered in Proposition 1, which, as has been shown, were the only type of strategy profile that could support a single 'grand' coalition as an equilibrium outcome (see Proposition 2).

⁸⁵ Otherwise, two coalitions at the end of stage 2 would not be viable and, therefore, studying mergers to

Lemma 2 *A coalition structure composed of two coalitions of the same size $N/2$ is stand-alone stable if and only if:*

- $4/((2 - \sigma)) \leq \beta \leq 2(2 + \sigma)/(2 - \sigma)$, or
- $\beta > 2(2 + \sigma)/(2 - \sigma)$ and $N \leq \bar{N}(\beta, \sigma)|_{N=2m}$, where

$$\bar{N}(\beta, \sigma)|_{N=2m} \equiv \left(\frac{8(1 + \sigma)^2}{(2 + \sigma)^3} \right)^{\frac{\sigma}{\beta - 2}} \frac{(2 + \sigma)(\beta(2 - \sigma) - 4)}{\beta(2 - \sigma) - 2(2 + \sigma)}.$$

Proof. Take the case in which the industry is composed of two coalitions of the same size, $m = N/2$. Now, re-evaluating condition (3.109) for the case in which $m = N/2$, one concludes that this coalition structure is stand-alone stable if and only if:

$$\frac{1}{N}(\beta(2 - \sigma) - 4) \geq \frac{F^*(3; S, \sigma)}{F^*(2; S, \sigma)} \left(\frac{\beta(2 - \sigma) - 2(2 + \sigma)}{(2 + \sigma)} \right). \quad (3.114)$$

Now, two different cases should be considered. First, if $4/((2 - \sigma)) \leq \beta \leq 2(2 + \sigma)/(2 - \sigma)$, then the l.h.s. of condition (3.114) is always non-negative, while the r.h.s. of the same condition is always non-positive. Therefore, condition (3.114) is trivially satisfied for this range of parameter values. Second, if $\beta > 2(2 + \sigma)/(2 - \sigma)$, then some algebra shows that condition (3.114) holds if and only if

$$N \leq \frac{F^*(2; S, \sigma)(2 + \sigma)(\beta(2 - \sigma) - 4)}{F^*(3; S, \sigma)\beta(2 - \sigma) - 2(2 + \sigma)}. \quad (3.115)$$

We can now combine eqs. (3.91) and (3.100) to express $F^*(2; S, \sigma)/F^*(3; S, \sigma)$ in explicit form. By doing so, we obtain

$$\frac{F^*(2; S, \sigma)}{F^*(3; S, \sigma)} = \left\{ \frac{8(1 + \sigma)^2}{(2 + \sigma)^3} \right\}^{\frac{\sigma}{\beta - 2}}. \quad (3.116)$$

duopoly for that range of parameter values would not make any sense.

Therefore, combining (3.115) and (3.116), one concludes that, whenever $\beta > 2(2 + \sigma) / (2 - \sigma)$, a coalition structure composed of two equal sized coalitions is stand-alone stable if and only if

$$N \leq \left(\frac{8(1 + \sigma)^2}{(2 + \sigma)^3} \right)^{\frac{2}{\beta - 2}} \frac{(2 + \sigma)(\beta(2 - \sigma) - 4)}{\beta(2 - \sigma) - 2(2 + \sigma)} \equiv \bar{N}(\beta, \sigma)|_{N=2m}. \quad (3.117)$$

This completes the proof. ■

Two remarks are in order at this point. First, notice that the intuition behind this Lemma is very similar to the one underlying the result in Proposition 2. If $4/((2 - \sigma)) \leq \beta \leq 2(2 + \sigma) / (2 - \sigma)$ (Region *B* in Figure 1), then at most two coalitions are viable in equilibrium, which in turn implies that whatever the number of coalition members is, an unilateral deviation towards the formation of a singleton coalition is never going to be profitable. If, instead, $\beta > 2(2 + \sigma) / (2 - \sigma)$ (Regions *C* and *D* in Figure 1), then by sticking the strategy of forming a symmetric duopoly, each of the N entrants will earn a share $2/N$ of the net duopoly profit, while by free-riding on the $N - 1$ rivals, it earns the triopoly profits. As a result, an upper bound for the number of firms in the initial wave of entrants is called for so that firms do not find it profitable to free-ride.⁸⁶ Second, as already mentioned, for a symmetric duopoly to be the outcome of a Nash equilibrium of the coalition formation game, then apart from stand-alone stability, it is also necessary that no firm will have incentives to unilaterally deviate by migrating to the rival coalition in the

⁸⁶ Notice that $\lim_{\beta \rightarrow \infty} \bar{N}(\beta, \sigma)|_{N=2m} = 8 \frac{(1 + \sigma)^2}{(2 + \sigma)^2}$ and $\forall \sigma \in [0, 1]$, $8 \frac{(1 + \sigma)^2}{(2 + \sigma)^2} < 4$, the minimum number of firms that should enter at stage 1 in order for a merger to a symmetric duopoly to be feasible at the second stage. Therefore, the stand-alone stability requirement fails to hold, which in turn implies that no merger to (symmetric) duopoly can occur in *exogenous sunk cost industries*. This is so since in such industries firms' only incentive to merge is to reduce competition in the market. Firms in this type of industries do not raise the quality level above unit and, therefore, affording a cost reduction through the elimination of duplication efforts in R&D is obviously not feasible.

same coalition structure. The next Proposition shows that this latter stability requirement is always satisfied in our framework and, therefore, stand-alone stability turns out to be a *necessary and sufficient* condition for a coalition structure composed of two equal sized individual coalitions to be the outcome of a Nash equilibrium of the coalition formation game.

Proposition 3 *Let $C = \{C_1, C_2\}$ be the coalition structure induced by firms' announcements $\alpha \equiv (S^1, \dots, S^N)$. If coalitions C_1 and C_2 are equal sized, then C is the outcome of a Nash equilibrium if and only if C is stand-alone stable.*

Proof. From Definition 2, one has that for a coalition structure C to be the outcome of a Nash equilibrium, no firm should have incentives to either unilaterally move between the two coalitions C_1 and C_2 when that is feasible or to stay alone in a new singleton coalition. Hence, in order to show that stand-alone stability is a necessary and sufficient condition for $C = \{C_1, C_2\}$ to be the outcome of a Nash equilibrium, all we have to show is that unilateral migrations between C_1 and C_2 are not profitable deviations. Consider a strategy profile in which $S^1 = \dots = S^{N/2} = \{1, \dots, N/2\}$, and $S^{N/2+1} = \dots = S^N = \{i, N/2 + 1, \dots, N\}$, where $i \in \{1, \dots, N/2\}$. The outcome of this strategy profile is $C = \{C_1, C_2\}$, where $C_1 = \{1, \dots, N/2\}$ and $C_2 = \{N/2 + 1, \dots, N\}$. Given this strategy profile, firm i can leave its actual coalition and join coalition C_2 by an individual deviation, $C_j(i) = C_2$. However, it is straightforward to show that firm i will not find this feasible deviation profitable. By sticking to the above-mentioned strategy profile, firm i earns a share $2/N$ of the coalition C_1 net profit $S\pi(v|2, v) - F^*(2; S, \sigma)$, whereas by changing its an-

nouncement to $S^i = \{i, N/2 + 1, \dots, N\}$ firm i will be worse-off. Its new payoff is given by a (lower) share $2/(N + 2)$ of the same profit $S\pi(v|2, v) - F^*(2; S, \sigma)$. This completes the proof. ■

The results given in Propositions 2 and 3 may be illustrated using Figure 2.

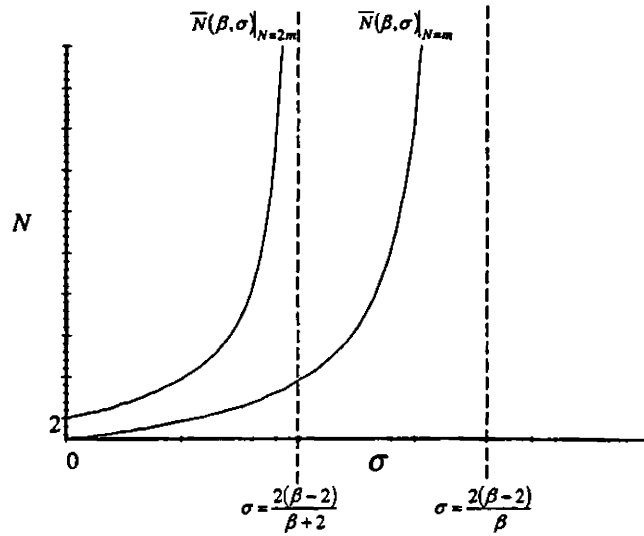


Figure 2 - Nash Stability Conditions

Figure 2 is constructed for a given value of the cost parameter β . The regions between the dashed vertical lines correspond to the ones in Figure 1, but now they are split in terms of the values assumed by σ . In addition, the two upward sloping curves $\bar{N}(\beta, \sigma)|_{N=m}$ and $\bar{N}(\beta, \sigma)|_{N=2m}$ are related to conditions (3.113) and (3.117), respectively. In particular, when $N < \bar{N}(\beta, \sigma)|_{N=m}$ (respectively, $N < \bar{N}(\beta, \sigma)|_{N=2m}$), then condition (3.113) (respectively, condition (3.117)) holds. Hence, given the number of entrants N , a coalition structure composed of a single 'grand' coalition (respectively, two equal sized duopoly coalitions) is stand-alone stable for sufficiently high values of σ . The intuition behind this result is that, the higher the degree of substitution between product varieties (the higher the

value of σ), the tougher the price competition will be at the last stage of the game. This in turn implies that the higher σ is, the lower profit margins are and, therefore, the less profitable will a deviation towards the creation of a new singleton coalition be. Appendix 3.A, studies the properties of the curves $\bar{N}(\beta, \sigma)|_{N=m}$ and $\bar{N}(\beta, \sigma)|_{N=2m}$ which justify the pattern presented in Figure 2.

Notice that in Region A, i.e., for $\sigma > 2(\beta - 2)/\beta$, the unique prediction of the coalition formation game is a merger to monopoly, whatever the number of firms that decided to enter at stage 1 is. This is so because, as explained above, in this region one can have at most one coalition earning non-negative profits at equilibrium (see discussion of Figure 1 above). In the remaining regions, given the number of entrants N , the prediction will depend on the specific value assumed by σ . For any pair (N, σ) below the $\bar{N}(\beta, \sigma)|_{N=m}$ curve, a single 'grand' coalition can be supported as a Nash equilibrium outcome of the coalition formation game. Likewise, for any pair (N, σ) below the $\bar{N}(\beta, \sigma)|_{N=2m}$ curve, a duopoly coalition structure composed of two equal sized coalitions is stand-alone stable and, hence, can be supported as a Nash equilibrium outcome.

Note that applying the Nash stability concept to the coalition formation game does not give very sharp predictions as to the coalitions that are likely to form in equilibrium. As shown by Figure 2, when $\sigma \leq 2(\beta - 2)/\beta$, all the region below the $\bar{N}(\beta, \sigma)|_{N=m}$ curve is a region where multiple Nash equilibria exist for the coalition formation game. In addition, it is also straightforward to show that for all pairs (N, σ) in Region B, any duopoly coalition

structure can be supported as the outcome of a Nash equilibrium of the coalition formation game.⁸⁷

To try and obtain sharper predictions about stable coalition structures which can arise at equilibrium, the next section will introduce a new stability concept which differs from Nash stability in two respects. First, stability will also be required with respect to deviations by a *group* of players (from the same coalition or from different ones), holding the strategy of the other players fixed. Second, firms are assumed to be *farsighted*. When making a decision, each coalition will take into account the future reactions of other coalitions to such a decision. This foresight is assumed to be common knowledge.

Farsighted stability

The analysis in this section will be performed in two steps. First, firms' incentives to participate in group deviations are studied. We then introduce a stability concept which not only requires stability with respect to individual as well as group deviations, but also assumes players are not myopic when considering deviations from the status quo coalition structure they belong to.

Let us start by considering the possibility that a group of players (from the same coalition or from different coalitions) forms a new coalition. Two different checks are in order at this point: (1) analyzing whether a group of players have incentives to create a new coalition whose size is smaller than the size of the coalition(s) they are departing from; and

⁸⁷ If we focus on strategies of the type considered in Proposition 1, then stand-alone stability turns out to be a necessary and sufficient condition for a coalition structure to be Nash stable. In addition, even though Lemma 2 considers only the case of a duopoly composed of two equal sized coalitions, it is straightforward to extend its proof and show that *any* duopoly coalition structure would be stand-alone stable in Region *B*. This is so for the simple reason that 2 is the maximum number of coalitions that can earn non-negative profits at equilibrium.

(2) investigating whether firms initially organized in a duopoly coalition structure have incentives to jointly deviate towards complete monopolization of the industry. The two next Lemmas analyze these issues.

Lemma 3 *Let $C = \{C_1, C_2\}$ be the coalition structure induced by firms' announcements $\alpha \equiv (S^1, \dots, S^N)$. If C is stand-alone stable and C_1 and C_2 are viable, then there is no profitable deviation towards the creation of new coalition(s) whose size is no larger than the actual size of C_1, C_2 .*

Proof. Consider the strategy profile in which $S^1 = \dots = S^k = \{1, \dots, k\}$, and $S^{k+1} = \dots = S^N = \{k+1, \dots, N\}$, where $1 \leq k \leq N/2$. The outcome of this strategy profile is $C = \{C_1, C_2\}$, where $C_1 = \{1, \dots, k\}$ and $C_2 = \{k+1, \dots, N\}$. Suppose that C is stand-alone stable. First, consider the case in which a deviation gives rise to the creation of a new coalition whose size is no larger than $\min\{k, N-k\}$, holding the rest of the coalition structure fixed. Let h players included in C_1 and l players belonging to C_2 (jointly) deviate towards the creation of a new third coalition of size $h+l$, where $0 \leq h < k$, $0 \leq l < (N-k)$ and $1 \leq h+l \leq \min\{k, N-k\}$. The non-deviator members of coalition C_i , $i = 1, 2$, stay together in a smaller coalition than the initial one since they announce exactly the same list of firms. Each of the deviators earns a share $1/(h+l)$ of the new coalition net profit $S\pi(v|3, v) - F^*(3; S, \sigma)$. Notice, however, that by individually leaving its coalition and forming a one-firm coalition, a deviator would earn the whole profit $S\pi(v|3, v) - F^*(3; S, \sigma)$. But, since C is stand-alone stable by hypothesis, none of these deviations is profitable.

Second, let us consider the case in which there is a deviation by a group of players towards the creation of several smaller coalitions. From the deviators point of view, the best deviation scenario would be the creation of two new singleton coalitions. Let j players included in C_1 and t players belonging to C_2 jointly deviate towards the creation of two singleton coalitions, where $0 \leq j \leq 2$, $0 \leq t \leq 2$ and $j + t = 2$. When this is the case, each deviator will earn a payoff of $S\pi(v|4, v) - F^*(4; S, \sigma)$. Notice, however, using (3.108) one concludes that

$$\frac{S\pi(v|N_2, v) - F^*(N_2; S, \sigma)}{F^*(N_2; S, \sigma)} = \frac{1}{2} \frac{\beta(2 - \sigma) - 2(2 + \sigma(N_2 - 2))}{2 + \sigma(N_2 - 2)}. \quad (3.118)$$

Now, from (3.118) one has that

$$\frac{\partial}{\partial N_2} \left(\frac{S\pi(v|N_2, v) - F^*(N_2; S, \sigma)}{F^*(N_2; S, \sigma)} \right) = -\frac{1}{2} \frac{\sigma\beta(2 - \sigma)}{(2 + \sigma(N_2 - 2))^2} < 0. \quad (3.119)$$

In addition, notice that $\frac{\partial F^*(N_2; S, \sigma)}{\partial N_2} = \frac{dF(v)}{dv} \frac{\partial v}{\partial N_2}$, where v , the (symmetric) equilibrium quality level, is given by (3.100). Now, since, on the one hand, from (3.91), $\frac{dF(v)}{dv} > 0$ and, by Lemma 1 $\frac{\partial v}{\partial N_2} < 0$, on the other, one concludes that $\frac{\partial F^*(N_2; S, \sigma)}{\partial N_2} < 0$. Combining this result with eq. (3.119) and applying the implicit function theorem, one concludes that, if C_1 and C_2 are viable coalitions, then $\frac{\partial(S\pi(v|N_2, v) - F^*(N_2; S, \sigma))}{\partial N_2} = \frac{\partial}{\partial N_2} \left[\left(\frac{S\pi(v|N_2, v) - F^*(N_2; S, \sigma)}{F^*(N_2; S, \sigma)} \right) F^*(N_2; S, \sigma) \right] < 0$. When the group deviation involves the creation of two singleton coalitions, each deviator gets a lower profit than in the case in which it unilaterally deviates towards the creation of an unique singleton coalition. This result, combined with the stand-alone stability of C , rules out the incentives by a group of players towards the creation of several singleton coalitions. This completes the proof. ■

The next Lemma provides the conditions which should be satisfied in order for a joint deviation by all players (initially organized in a duopoly coalition structure) towards complete monopolization of the industry not to be profitable.

Lemma 4 *Let $C = \{C_1, C_2\}$ be the coalition structure induced by firms' announcements $\alpha \equiv (S^1, \dots, S^N)$. Let the sizes of coalitions C_1 and C_2 be given by k and $N - k$, respectively, where $1 \leq k \leq N/2$. Then, a joint deviation by all players towards the formation of a monopoly is not profitable if and only if $N \geq \underline{N}^{DM_k}$, where*

$$\underline{N}^{DM_k} \equiv k \frac{2(\beta - 2)}{\beta(2 - \sigma) - 4} \left(\frac{(2 + \sigma)^2(2 - \sigma)}{8} \right)^{\frac{\theta}{\beta - 2}}.$$

Proof. If initially there are two coalitions C_1 and C_2 of sizes k and $N - k$, respectively, $1 \leq k \leq N/2$, then firms will have *no* incentive to jointly deviate towards complete monopolization of the industry if and only if the following 2 conditions *simultaneously* hold:

$$\frac{1}{k} [S\pi(v|2, v) - F^*(2; S, \sigma)] \geq \frac{1}{N} [S\pi(v|1, v) - F^*(1; S, \sigma)], \quad (3.120)$$

and

$$\frac{1}{N - k} [S\pi(v|2, v) - F^*(2; S, \sigma)] \geq \frac{1}{N} [S\pi(v|1, v) - F^*(1; S, \sigma)]. \quad (3.121)$$

Now, since firms in coalition C_1 are obviously the ones who have less incentives to participate in such a (group) deviation, incentive constraint (3.120) is the one that will matter for our analysis. Hence, rewriting (3.120) one obtains

$$\frac{1}{k} \left[\frac{S\pi(v|2, v)}{F^*(2; S, \sigma)} - 1 \right] \geq \frac{1}{N} \left[\frac{S\pi(v|1, v)}{F^*(1; S, \sigma)} - 1 \right] \frac{F^*(1; S, \sigma)}{F^*(2; S, \sigma)}. \quad (3.122)$$

Now, making use of (3.118), one has that:

$$\frac{S\pi(v|1, v)}{F^*(1; S, \sigma)} - 1 = \frac{1}{2}\beta - 1,$$

$$\frac{S\pi(v|2, v)}{F^*(2; S, \sigma)} - 1 = \frac{1}{4}\beta(2 - \sigma) - 1.$$

Hence, one can rewrite condition (3.122) as follows:

$$\frac{1}{k} \left(\frac{1}{4}\beta(2 - \sigma) - 1 \right) \geq \frac{1}{N} \left(\frac{1}{2}\beta - 1 \right) \frac{F^*(1; S, \sigma)}{F^*(2; S, \sigma)}. \quad (3.123)$$

Now, using (3.112), after some algebra, the previous condition can be rewritten as:

$$N \geq k \frac{2(\beta - 2)}{\beta(2 - \sigma) - 4} \left(\frac{(2 + \sigma)^2(2 - \sigma)}{8} \right)^{\frac{\beta}{\beta - 2}} \equiv \underline{N}^{DM_k} \quad (3.124)$$

This completes the proof. ■

The intuition behind this result is as follows. The higher the number of firms N that have entered at stage 1, the lower the share $(1/N)$ that each firm will get if a joint deviation towards monopolization of the industry takes place. In addition, the k firms in the smaller coalition C_1 are obviously the ones with less incentives to participate in such a deviation.⁸⁸ This in turn implies that if the number of firms that have decided to enter at stage 1 is sufficiently large, then firms in the smallest coalition C_1 will not agree to form a single 'grand' coalition.

Let us now turn to the introduction of a new (coalition) stability concept which not only requires stability with respect to group deviations of the type just described, but also amends the firm's myopic behavior embedded in the Nash stability notion described in the previous section. It will be assumed that players are fully farsighted, which, as pointed out by Chwe (1994), means that "a coalition considers the possibility that, once it acts, another

⁸⁸ Firms in coalition C_1 are the ones which get the highest share of the duopoly profits (net of fixed costs).

coalition might react, a third coalition might in turn react, and so on, without limit.” (p. 300) Hence, we assume an environment of coalition formation in which deviations can be followed by further deviations. This being the case, then what matters is not only to analyze whether a deviation by some player(s) is vulnerable to further deviations, but also, and most importantly, to identify the *ultimate* result of such a deviation. This *ultimate* result, being, by definition, immune to further deviations, must be in the solution set of the coalition formation game.

Assume that coalition structure $C = \{C_1, \dots, C_T\}$ is under consideration (C is the status quo). Now, a coalition $D \subset \{1, \dots, N\}$ may form and “object to” C . Let D_i denote the subset of deviant firms belonging to coalition C_i , where $i = 1, \dots, T$ and $D = \cup_{i=1}^T D_i$. After the deviation by coalition D and before other players eventually regroup, the resulting coalition structure is

$$C' = \{D\} \cup \{C_i \setminus D_i, i = 1, \dots, T \mid C_i \in C \text{ and } C_i \setminus D_i \neq \{\emptyset\}\}; \quad (3.125)$$

in this case we write $C \xrightarrow{D} C'$. This means that if C is the status quo coalition structure, coalition D can make C' the new status quo. This does not imply that C' can be enforced by coalition D . As will be shown, once C' is the new status quo, then another coalition can form and “object to” C' , and so on.

Definition 3 Coalitional preferences: Let $C = \{C_1, \dots, C_T\}$ be the status quo coalition structure. Let D_i denote the (possibly empty) set of deviant firms belonging to coalition C_i , $i = 1, \dots, T$ and $D = \cup_{i=1}^T D_i$. Assume that $D \subset \{1, \dots, N\}$ is formed, giving rise to the coalition structure C' , defined by eq. (3.125). Coalition D strictly prefers C' to C ,

denoted $C' \succ_D C$, if for all $i \in D_i$, $D_i \neq \{\emptyset\}$, where $i = 1, \dots, T$, $\Pi(D; C') > \Pi(C_i; C)$.

In words, the acting coalition D strictly prefers the induced coalition structure C' to the status quo coalition C if all its members are better off after D is formed (and stays formed until the next move). However, as stressed above, at any point in time, some firm(s) might form a new coalition and change the (new) status quo coalition structure. The coalition formation game is over when an *ultimate* coalition structure is reached, i.e., when no firm has incentives to further deviate from this coalition structure. The next definitions formalize this idea of foresight in the coalition formation game.

Definition 4 Dominance: Take two coalition structures C and C' . C' is said to dominate C , $C' > C$, if there exists a sequence of coalition structures C^1, C^2, \dots, C^k such that $C^1 = C$ and $C^k = C'$ and a sequence of coalitions D^1, D^2, \dots, D^k such that $C^j \xrightarrow{D^j} C^{j+1}$ and $C^{j+1} \succ_{D^j} C^j$, for all $j = 1, \dots, k-1$.

Definition 5 Farsighted stability: The coalition structure C is said to be farsighted stable if there does not exist another coalition structure C' such that $C' > C$.

In what follows, we will apply this concept of farsighted coalitional stability to our setting in order to obtain predictions about the coalitions that are likely to form in equilibrium. Two notes are in order at this point. First, we will focus our analysis on regions A and B of the previous Figures. In other words, our aim is to predict the equilibrium structure of the industry at hand for sufficiently high values of the substitutability parameter σ , i.e.,

for $\sigma \in [2(\beta - 2)/(\beta + 2), 1]$.⁸⁹ This is for model tractability reasons.⁹⁰ Second, as the previous discussion suggests, when solving the overall game and applying the concept of farsighted stability, we have to take into account the possible eventual deviations as follows. When players are initially organized in a monopoly coalition structure, then the only feasible deviation is an individual deviation towards the creation of a singleton coalition (monopoly stand-alone stability). If, instead, the status quo coalition structure is a duopoly one, then players can not only deviate towards the formation of a singleton coalition, but also jointly deviate towards the complete monopolization of the industry. All other kinds of deviations have been shown to be unfeasible or unprofitable.

The study of the farsighted stable coalition structures that can arise in equilibrium will be done jointly with the determination of the equilibrium number of entrants at the first stage of the game. This analysis will be developed in the next section and will give rise to the main result of this chapter.

Entry stage

We are now in position to analyze firms' entry decisions at the first stage of the game and so determine the equilibrium structure of the industry. By so doing, our objective is twofold: (i) determine the number of pre-merger entrants at stage 1, and (ii) characterize the coalition structure that is likely to occur at equilibrium at the second stage, assuming

⁸⁹ Obviously, requiring that products are sufficiently good substitutes, where the minimum threshold value for σ depends on β , is equivalent to require that investment in R&D is sufficiently effective, i.e., $\beta \in (2, 2(2 + \sigma)/(2 - \sigma)]$.

⁹⁰ If cases in which $\sigma < 2(\beta - 2)/(\beta + 2)$ were considered, then one would have to also study the stability of coalition structures composed of more than two coalitions and the associated feasible deviations of firms involved in such coalitions.

that players are endowed with foresight in that they look at many steps ahead and consider only credible outcomes of the coalition formation game of stage 2.

As already mentioned, we are restricting the analysis the cases in which $\sigma \in [2(\beta - 2)/(\beta + 2), 1]$ - Regions *A* and *B* of the previous figures. We know that in region *A*, *B*, the maximum number of coalitions earning non-negative profits at equilibrium is one and two, respectively. Hence, let us start by computing the upper bounds to the number firms that can enter at stage 1 without incurring in losses, anticipating that a monopoly or a duopoly coalition structure is going to be formed at stage 2, respectively.

If firms anticipate that a monopoly coalition structure is going to be formed in equilibrium at stage 2, then at the first stage of the game firms will enter up to a point at which N is the (largest integer) value satisfying:

$$\frac{1}{N} (S\pi(v|1, v) - F^*(1; S, \sigma)) \geq \varepsilon, \quad (3.126)$$

or, equivalently,

$$\frac{1}{N} \left(\frac{S\pi(v|1, v) - F^*(1; S, \sigma)}{F^*(1; S, \sigma)} \right) \geq \frac{\varepsilon}{F^*(1; S, \sigma)}. \quad (3.127)$$

For the sake of simplicity, write N as a continuous variable for the moment. Then, from (3.118), one has that $(S\pi(v|1, v) - F^*(1; S, \sigma))/F^*(1; S, \sigma) = (\beta - 2)/2$. Hence, eq. (3.127) can now be solved for the equilibrium number of entrants at stage 1, N , obtaining:

$$N \leq \left(\frac{\beta - 2}{2\varepsilon} \right) F^*(1; S, \sigma). \quad (3.128)$$

Finally, combining (3.91) and (3.100), one has that $F^*(1; S, \sigma) = \left(\frac{S}{4\beta}\right)^{\beta/(\beta-2)}$. This in turn implies that the previous condition can be rewritten in the following way:

$$N \leq \left(\frac{\beta-2}{2\varepsilon}\right) \left(\frac{S}{4\beta}\right)^{\frac{\beta}{\beta-2}} \equiv \overline{N}_\varepsilon^M. \quad (3.129)$$

Notice that, as expected, $\overline{N}_\varepsilon^M$ increases with the market size S and is a decreasing function of the entry cost ε . In particular, $\overline{N}_\varepsilon^M \rightarrow \infty$ as $S \rightarrow \infty$ or $\varepsilon \rightarrow 0$.

Take now the alternative case in which firms, at stage 1, anticipate that in the second stage of the overall game, the equilibrium outcome of the coalition formation game is coalition structure $C = \{C_1, C_2\}$, where the sizes of coalitions C_1, C_2 are given by $k, N - k$, respectively, $1 \leq k \leq N/2$. If this is the case, then the number of firms that decide to enter at stage 1 will be the (largest integer) value of N for which each member of the larger coalition C_2 will make non-negative profits:

$$\frac{1}{N-k} (S\pi(v|2, v) - F^*(2; S, \sigma)) \geq \varepsilon, \quad (3.130)$$

or, equivalently

$$\frac{1}{N-k} \left(\frac{S\pi(v|2, v) - F^*(2; S, \sigma)}{F^*(2; S, \sigma)} \right) \geq \frac{\varepsilon}{F^*(2; S, \sigma)}. \quad (3.131)$$

Now, since from (3.118), $(S\pi(v|2, v) - F^*(2; S, \sigma)) / F^*(2; S, \sigma) = (\beta(2 - \sigma) - 4) / 4$, one can reevaluate the previous inequality and solve for N , obtaining:

$$N \leq \left(\frac{\beta(2 - \sigma) - 4}{4\varepsilon} \right) F^*(2; S, \sigma) + k. \quad (3.132)$$

But, since from (3.91) and (3.100), $F^*(2; S, \sigma) = \left(\frac{S}{\beta(2-\sigma)(2+\sigma)^2}\right)^{\frac{\beta}{\beta-2}}$, one finally has that:

$$N \leq \left(\frac{\beta(2 - \sigma) - 4}{4\varepsilon} \right) \left(\frac{S}{\beta(2 - \sigma)(2 + \sigma)^2} \right)^{\frac{\beta}{\beta-2}} + k \equiv \overline{N}_\varepsilon^{D_k}. \quad (3.133)$$

In Appendix 3.B, properties of the upper bound to the number of firms $\overline{N}_\epsilon^{D_k}$ are studied. It is interesting to note, in particular, that while \overline{N}_ϵ^M is independent of σ , $\overline{N}_\epsilon^{D_k}$ is decreasing in the substitution parameter. The intuition for this is as follows. If a monopoly is formed at the second stage of the game, then the unique grand coalition will operate along a single R&D trajectory and produce a single product. This explains why \overline{N}_ϵ^M is independent of the substitution parameter σ . As far as $\overline{N}_\epsilon^{D_k}$ is concerned, two reasons justify why this upper bound to the number of entrants is a decreasing function of σ . First, as already mentioned, the higher the substitutability between product varieties in the market (the higher the value of σ) is, the tougher competition will be at the last stage of the game. This implies that the higher σ is, the lower profit margins are and, therefore, the lower the number of firms that can enter and join a duopoly coalition structure without incurring in losses. Second, the higher σ is, the less variety is offered in the market and, hence, the lower will consumers' expenditure be on the goods offered by this industry, for given prices. This will also imply that increases in σ will induce a decrease in the maximum number of firms that can enter earning a non-negative profit in case a duopoly coalition structure is formed at the second stage of the game.

For each value of k , $k \in \{1, 2, \dots, N/2\}$, equation (3.133) defines a decreasing schedule in the (N, σ) space. Figure 3 shows this family of schedules and also the upper bound to the number of firms \overline{N}_ϵ^M , given by eq. (3.129). In addition, Figure 3 presents another family of schedules, each of which for a given value of k . This family of schedules is given by eq. (3.124) and represents a lower bound to the number of entrants above which a duopoly coalition structure of the type $C = \{C_1, C_2\}$, where k and $N - k$ denote, respectively, the

sizes of C_1 and C_2 , is stable against group deviations towards complete monopolization of the industry. The intuition for the upward sloping of these curves is as follows. As already mentioned, for high values of σ , competition in the product market stage 4 is expected to be fierce. Therefore, the higher the value of σ is, the more likely is that firms will have incentives to jointly deviate towards complete monopolization of the industry in order to suppress competition at the product market stage 4. In other words, for a given number of pre-merger entrants N , the higher σ is, the less likely is that condition (3.124) holds.

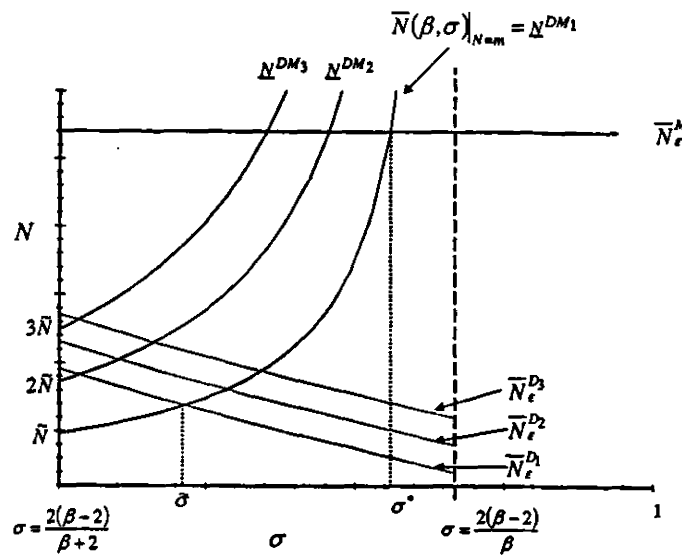


Figure 3 - Farsighted Stability Conditions

Let us now turn to the discussion of the equilibrium structure of the industry. As explained before, this will be done by applying a refinement of subgame perfection to the overall game, where it is assumed that at the second stage players are endowed with foresight. The analysis is going to be performed with the help of Figure 3, which summarizes the relevant stability requirements which we impose on permissible individual and group deviations as well as the upper bounds to the number of entrants earning non negative

profits in monopoly and duopoly coalition structures. In what follows, let $\sigma^*(\beta; S, \varepsilon)$ denote the value of σ implicitly defined by the condition $\bar{N}_\varepsilon^M = \bar{N}(\beta, \sigma)|_{N=m}$, where \bar{N}_ε^M and $\bar{N}(\beta, \sigma)|_{N=m}$ are given by eqs. (3.129) and (3.113), respectively. In addition, let $\tilde{\sigma}(\beta; S, \varepsilon)$ denote the value of σ implicitly defined by condition $\bar{N}_\varepsilon^{D_1} = \bar{N}(\beta, \sigma)|_{N=m}$, where $\bar{N}_\varepsilon^{D_1}$ is given by eq. (3.133), for the specific case in which $k = 1$.⁹¹

The following Proposition puts forward the main result of this chapter.

Proposition 4 *Let $\sigma \in [2(\beta - 2) / (\beta + 2), 1]$. Then, there exists a (refined) SPE of the four stage game. Let \bar{N} denote the equilibrium number of pre-merger entering firms in a (refined) SPE. Then,*

1. *If $\sigma \geq \sigma^*$, $\bar{N} = \bar{N}_\varepsilon^M$ and a monopoly coalition structure is formed at the second stage of the game, $N_2 = 1$;*
2. *If $\sigma \in (\tilde{\sigma}, \sigma^*)$, $\bar{N} = \bar{N}(\beta, \sigma)|_{N=m}$ and a monopoly coalition structure is formed at the second stage of the game, $N_2 = 1$;*
3. *Otherwise, $\bar{N} \in \{ \bar{N}_\varepsilon^{D_k} \mid k^* \in \{k \in \{1, 2, \dots, N/2\} \mid \bar{N}^{DM_k} \leq \bar{N}_\varepsilon^{D_k} \} \}$ and a duopoly coalition structure is formed at the second stage of the game, $N_2 = 2$.*

Proof. Two preliminary remarks are in order. First, from Proposition 2 and Lemma 2, one has that a monopoly coalition structure and a duopoly coalition structure are stand-alone stable in Region *A* and *B*, respectively, of Figures 1 - 3. Second, take the cases in which a group of N entrants form a coalition structure of the type $C = \{C_1, C_2\}$ at stage

⁹¹ In what follows, for the sake of simplicity of notation, we will refer to $\sigma^*(\beta; S, \varepsilon)$ and $\tilde{\sigma}(\beta; S, \varepsilon)$, as σ^* and $\tilde{\sigma}$, respectively.

2, where k and $N - k$ denote the sizes of C_1 , C_2 , respectively, $1 \leq k \leq N/2$. When this is the case, due to symmetry and, therefore, without loss of generality, it will be assumed in what follows that $C_1 = \{1, \dots, k\}$ and $C_2 = \{k + 1, \dots, N\}$. Let us now turn to the analysis of each of the three cases specified above for $\sigma \in [2(\beta - 2)/(\beta + 2), 1]$.⁹² In the first case, i.e., when $\sigma \geq \sigma^*$, two different scenarios should be considered. First, if $\sigma \in [2(\beta - 2)/\beta, 1]$ (Region A), as mentioned above, one can have at most one coalition earning non-negative profits at equilibrium. As a consequence, firms anticipate that a monopoly coalition structure is going to be formed at stage 2. Since, \bar{N}_ϵ^M represents the maximum number of firms in a monopoly coalition structure earning non negative profits at equilibrium, one must have that $\bar{N} = \bar{N}_\epsilon^M$ and $N_2 = 1$. Second, consider the alternative scenario in which $\sigma \in [\sigma^*, 2(\beta - 2)/\beta)$. Let us analyze the ensuing subgame for each possible value of N . If $N \leq \bar{N}_\epsilon^M$, a monopoly coalition structure would give non-negative profits to all firms and is stand-alone stable since $\bar{N}_\epsilon^M \leq \bar{N}(\beta, \sigma)|_{N=m}$. In addition, any duopoly coalition structure, although stand-alone stable, would not be stable against group deviations towards complete monopolization of the industry, $\bar{N}_\epsilon^M < \underline{N}^{DM_k}$, $k \in \{1, 2, \dots, N/2\}$. If, instead, $\bar{N}_\epsilon^M < N < \underline{N}^{DM_1}$, then a monopoly coalition structure, although stand-alone stable ($N < \bar{N}(\beta, \sigma)|_{N=m} = \underline{N}^{DM_1}$), implies that all its members will earn negative profits at equilibrium ($N > \bar{N}_\epsilon^M$). In addition, any duopoly coalition structure is not stable towards complete monopolization of the industry, $N < \underline{N}^{DM_k}$, $k \in \{1, 2, \dots, N/2\}$, and we know that in a monopoly firms end up earning negative prof-

⁹² Actually, as already mentioned, the proof of Lemma 2 is developed for the case in which a duopoly coalition structure is composed of coalitions of the same size. None the less, it is straightforward to extend that proof and show that any duopoly coalition structure is stand-alone stable in Region B.

its. Lastly, if $N > \underline{N}^{DM_1}$, a monopoly coalition structure is not stand-alone stable. Any coalition structure of the type $C = \{C_1, C_2\}$, where k and $N - k$ denote the sizes of C_1, C_2 , respectively and where $N > \{\bar{N}_\epsilon^{D_k}, \underline{N}^{DM_k}\}$, is stand-alone stable and, therefore, farsighted stable due to Lemma 3 and Definition 5, but it implies that firms in the larger coalition C_2 will earn negative profits, $N > \bar{N}_\epsilon^{D_k}$. As a result, one concludes that if $\sigma \geq \sigma^*$, $\bar{N} = \bar{N}_\epsilon^M$ and $N_2 = 1$.

Next, let us turn to the analysis of the second case, $\sigma \in (\tilde{\sigma}, \sigma^*)$. As before, let us start by analyzing the ensuing subgame for each possible value of N . First, if $N \leq \bar{N}(\beta, \sigma)|_{N=m}$ a monopoly coalition structure gives strictly positive profits to all its members ($N < \bar{N}_\epsilon^M$) and is stand-alone stable ($N \leq \bar{N}(\beta, \sigma)|_{N=m}$). Moreover, any duopoly coalition structure, although stand-alone stable, would not be stable against group deviations towards complete monopolization of the industry since $\bar{N}(\beta, \sigma)|_{N=m} \leq \underline{N}^{DM_k}$, for $k \in \{1, 2, \dots, N/2\}$. Second, consider any $\bar{N}(\beta, \sigma)|_{N=m} < N \leq \bar{N}_\epsilon^M$. When this is the case, then a monopoly coalition structure gives non negative profits to all its members, but it is not stand-alone stable. A single grand coalition would, therefore, induce a duopoly coalition structure of the type $C' = \{C_1, C_2\}$, where the sizes of C_1 and C_2 are, respectively, 1 and $N - 1$. Such a coalition structure is stand-alone stable and, therefore, from Lemma 3 and Definition 5, farsighted stable, but would imply that the $N - 1$ firms included in C_2 all earn negative profit, $N > \bar{N}_\epsilon^{D_1}$. It should be also noted that when $\bar{N}(\beta, \sigma)|_{N=m} < N \leq \bar{N}_\epsilon^M$, other duopoly coalition structure might be feasible. In particular whenever $N > \max\{\bar{N}_\epsilon^{D_k}, \underline{N}^{DM_k}\}$, $k \in \{2, \dots, N/2\}$, a coalition structure of the type $C = \{C_1, C_2\}$, where k and $N - k$ denote the sizes of C_1, C_2 , respectively, although stand-

alone stable and, thus, farsighted stable due to Lemma 3 and Definition 5, would imply that members in the larger coalition would earn negative profits in equilibrium, $N > \bar{N}_\epsilon^{D^*}$. Lastly, consider any $N > \bar{N}_\epsilon^M$. In this situation, a monopoly is not stand alone stable ($N > \bar{N}(\beta, \sigma)|_{N=m}$). A monopoly coalition structure would induce a duopoly coalition structure of the type $C' = \{C_1, C_2\}$, where the sizes of C_1 and C_2 are, respectively, 1 and $N - 1$. Such a coalition, as explained before for the case $\bar{N}(\beta, \sigma)|_{N=m} < N \leq \bar{N}_\epsilon^M$, is farsighted stable but implies that the $N - 1$ firms in C_2 earn negative profits. In addition, and again as in the case where $\bar{N}(\beta, \sigma)|_{N=m} < N \leq \bar{N}_\epsilon^M$, other duopoly coalition structures might be feasible but they imply that members in the larger coalition earn negative profits. Having said this, one concludes that if $\sigma \in (\tilde{\sigma}, \sigma^*)$, $\bar{N} = \bar{N}(\beta, \sigma)|_{N=m}$ and $N_2 = 1$.

Finally, let us analyze the situation where σ , although in region B does not belong to the intervals specified in the 2 previous cases. For this to be the case, one must have that $\tilde{\sigma} > 2(\beta - 2)/(\beta + 2)$.⁹³ First, for all values of $N \leq \bar{N}(\beta, \sigma)|_{N=m}$, a monopoly coalition structure is stand-alone stable and firms earn non negative profits. However, in equilibrium one must have more than $\bar{N}(\beta, \sigma)|_{N=m}$ firms entering the market. This is so because there exists a nonempty set $\{N | \underline{N}^{DM_1} \leq N \leq \bar{N}_\epsilon^{D_1}\}$. For all values of N in this set, a duopoly coalition structure $C' = \{C_1, C_2\}$, where the sizes of C_1 and C_2 are, respectively, 1 and $N - 1$ is stand-alone stable and, therefore, due to Lemma 3 and Definition 5, farsighted stable. In addition, since in that set $N \leq \bar{N}_\epsilon^{D_1}$, all members in the coalition structure C' earn non negative profits. Moreover, if $\bar{N}_\epsilon^{D_1} + 1$ firms enter the market, then, since $\bar{N}_\epsilon^{D_1} + 1 >$

⁹³ This will happen for sufficiently small values of ϵ . When ϵ is small enough, then $\lim_{\sigma \rightarrow \frac{2(\beta-2)}{\beta+2}} \bar{N}_\epsilon^{D^*} > \tilde{N}$, which in turn implies that $\tilde{\sigma} > 2(\beta - 2)/(\beta + 2)$ (see eq. (3.146)).

$\max \{ \bar{N}_\epsilon^{D_1}, \underline{N}^{DM_1} \}$, for the reasons explained above, a duopoly coalition structure whose coalitions sizes are 1 and $\bar{N}_\epsilon^{D_1}$ is stand-alone stable and, therefore, farsighted stable as well but members of the larger coalitions end up earning a negative profit. Knowing this, the initial $\bar{N}_\epsilon^{D_1}$ entrants would only accept the additional entrant in the larger coalition (earning a negative profit), whose actual members are firms $\{2, \dots, \bar{N}_\epsilon^{D_1}\}$. This will discourage entry by this additional firm. Notice as well that other duopolies might arise in equilibrium, giving rise to a multiple equilibria scenario. In particular, for any N such that $N > \bar{N}(\beta, \sigma)|_{N=m}$, if $N(k) = \{N > \bar{N}(\beta, \sigma)|_{N=m} | \underline{N}^{DM_k} \leq N \leq \bar{N}_\epsilon^{D_k}, k \in \{2, \dots, N/2\}\}$, a coalition structure $C'' = \{C_1, C_2\}$, where k and $N(k) - k$ denote the sizes of C_1 , C_2 , respectively, is farsighted stable. In addition, $\max_N N(k) = \bar{N}_\epsilon^{D_k}$ and, a very similar argument to the one used for the case $k = 1$ shows that an additional firm trying to enter the market and join one of the coalitions in C , will only be allowed by incumbents to join coalition C_2 and ends up earning negative profits. This justifies why $\bar{N} \in \{ \bar{N}_\epsilon^{D_{k^*}} | k^* \in \{k \in \{1, 2, \dots, N/2\} | \underline{N}^{DM_k} \leq \bar{N}_\epsilon^{D_k}\} \}$ and $N_2 = 2$ in case 3. ■

This result has two central empirical implications. First, notice that independently of the market size and the level of entry costs, arbitrarily concentrated outcomes may be supported in equilibrium. Even if $S \rightarrow \infty$ (or $\epsilon \rightarrow 0$), a monopoly coalition structure can be supported in equilibrium. As shown by eqs. (3.129) and (3.133), both $\bar{N}_\epsilon^M \rightarrow \infty$ and $\bar{N}_\epsilon^{D_k} \rightarrow \infty$ as $S \rightarrow \infty$ (or $\epsilon \rightarrow 0$). Nonetheless, the previous proposition shows that for every $\sigma \in [\tilde{\sigma}, 1]$, a merger to monopoly is sustained as the unique outcome of a (refined) SPE of the game. Hence, not even in large markets is it generally possible to exclude the occurrence of arbitrarily high concentration ratios. Therefore, in *endogenous sunk cost*

industries, an upper bound to concentration exists and is independent of the size of the market and the level of the entry costs. Second, notice that for $\sigma \in (\tilde{\sigma}, \sigma^*)$, a monopoly coalition structure where firms belonging to it earn strictly positive profits can be sustained in equilibrium even under the threat of entry. The intuition which underlies this result is that additional entry would rule out the monopoly stand-alone stability and induce a very sharp fall in profits. As shown in the proof of the previous Proposition, whenever $\sigma \in (\tilde{\sigma}, \sigma^*)$ and $N > \bar{N}(\beta, \sigma)|_{N=m}$, in any (farsighted) stable duopoly coalition structure firms belonging to the larger coalition end up earning negative profits. This suffices to deter entry by additional firms and maintain a monopoly outcome in which firms earn strictly positive profits.

The next Lemma further characterizes the equilibrium structure of the industry by identifying duopoly coalition structures which cannot be sustained in a (refined) SPE.

Lemma 5 *Let $\sigma \in [2(\beta - 2) / (\beta + 2), 2(\beta - 2) / \beta]$. Then, a coalition structure $C = \{C_1, C_2\}$, where k and $N - k$ denote the sizes of C_1, C_2 , respectively, cannot be sustained in a (refined) SPE if*

$$\frac{k}{N} > \frac{1}{\tilde{N}},$$

$$\text{where } \tilde{N} = \frac{1}{2} \frac{16^{\frac{\beta}{\beta-2}} \left(\beta^{\frac{\beta}{\beta-2}} \right)^2 (\beta+2)}{\left((\beta+2)^{\frac{\beta}{\beta-2}} \right)^3}.$$

Proof. From eq. (3.142), we have that $\lim_{\sigma \rightarrow \frac{2(\beta-2)}{\beta+2}} \underline{N}^{DM_k} = k\tilde{N}$, where \tilde{N} is given by eq. (3.135) and $1 \leq k \leq N/2$. Now, since, on the other hand, we know that $(d\underline{N}^{DM_k}/d\sigma) > 0$ (see eq. (3.144)), one concludes that, for all $\sigma \in [2(\beta - 2) / (\beta + 2), 2(\beta - 2) / \beta]$, a

coalition structure $C = \{C_1, C_2\}$, where k and $N - k$ denote the sizes of C_1 , C_2 , respectively, is not stable against a joint deviation towards complete monopolization of the industry if

$$N < \lim_{\sigma \rightarrow \frac{2(\beta-2)}{\beta+2}} N^{DM_k} = k\tilde{N},$$

or equivalently

$$\frac{k}{N} > \frac{1}{\tilde{N}}.$$

This completes the proof. ■

Hence, when $\sigma \in [2(\beta - 2) / (\beta + 2), 2(\beta - 2) / \beta]$ (Region *B* of Figures 1-3), a duopoly coalition structure can only be supported in a (refined) SPE if composed of sufficiently size asymmetric coalitions. The intuition behind this result is as follows. The higher the (relative) size of the smaller coalition (as measured by k/N) is, the lower will be the individual share of each firm belonging to that coalition on the coalition overall (duopoly) profit. This in turn implies that the higher k/N is, the more likely is that firms in the smaller coalition will be willing to participate in a joint deviation towards complete monopolization of the industry, therefore ruling out the (farsighted) stability of the duopoly coalition structure to which they initially belong.

3.3 Conclusion

This chapter investigates whether an upper bound to concentration exists and can be characterized in industries where firms are involved in R&D activities with the aim of enhancing the perceived quality of their products - *endogenous sunk cost industries* (Sutton (1991, 1998)).

Merger decisions were investigated in a setting where firms have a double incentive to participate in a merger. In our framework, a merger leads both to a reduction in the product market competition and to a realization of a cost reduction. By merging, firms share the (endogenous and fixed) cost associated with the investment in R&D activities. Hence, through a merger firms can pursue common research activities, avoid eventual duplication efforts in their R&D investments and, as a result, afford a cost advantage over the unmerged rivals.

A coalition formation game is proposed to capture the endogenous (horizontal) merger formation. The adopted coalitional stability concept differs from Nash stability in two respects. First, stability is also required with respect to deviations by a group of players (from the same coalition or from different ones), holding the strategy of the other players fixed. Second, this stability concept assumes firms are endowed with foresight when making merger decisions. Firms take into account the future reactions of other coalitions and, therefore, anticipate the ultimate result of their actions. This stability concept is shown to provide much sharper predictions than the standard Nash stability concept with regards to the coalition structures that can arise in equilibrium.

The analysis shows that, in this kind of industries, arbitrarily concentrated outcomes can be attained in equilibrium for *any* market size. In addition, when products are sufficiently good substitutes (or, when investment in R&D is sufficiently effective), coalition structures composed of sufficiently size symmetric coalitions cannot arise in equilibrium. These results, therefore, complement those of Sutton (1991, 1998).

3.A Stability requirements

In this section, we study in detail the properties of the stability requirements which are relevant for the determination of the stable coalition structures which may emerge as equilibria outcomes at the second stage of the game. These conditions are given by eqs. (3.113), (3.117) and (3.124). The following properties of these stability checks justify the pattern of the curves presented in Figures 2 and 3.

3.A.1 Stand-alone stability of a monopoly

As shown in Proposition 2, a monopoly is always stand-alone stable in region *A* of Figure 1, i.e., when $2 < \beta \leq 4/(2 - \sigma)$ (or, equivalently $2(\beta - 2)/\beta \leq \sigma \leq 1$). However, if, instead, $\beta > 4/(2 - \sigma)$ (or, equivalently, $0 \leq \sigma < 2(\beta - 2)/\beta$), then a monopoly will only be stand-alone stable whenever condition (3.113) holds.

First, let us study of the limiting properties of $\bar{N}(\beta, \sigma)|_{N=m}$. Making use of eq. (3.113), some algebra shows that

$$\lim_{\sigma \rightarrow 0} \bar{N}(\beta, \sigma)|_{N=m} = 1, \quad (3.134)$$

$$\lim_{\sigma \rightarrow \frac{2(\beta-2)}{\beta+2}} \bar{N}(\beta, \sigma)|_{N=m} = \frac{1}{2} \frac{16^{\frac{\beta}{\beta-2}} \left(\beta^{\frac{\beta}{\beta-2}}\right)^2 (\beta + 2)}{\left((\beta + 2)^{\frac{\beta}{\beta-2}}\right)^3} \equiv \tilde{N}, \quad (3.135)$$

$$\lim_{\sigma \rightarrow \frac{2(\beta-2)}{\beta}} \bar{N}(\beta, \sigma)|_{N=m} = +\infty. \quad (3.136)$$

Second, let us use (3.113) to carry out a simple comparative statics exercise with respect to the substitution parameter σ . By doing so, one concludes that

$$\frac{d \bar{N}(\beta, \sigma)|_{N=m}}{d\sigma} = \frac{4\beta (\beta (2 - \sigma)^2 + \sigma^2 + 6\sigma - 8) \left(\frac{1}{8} (2 + \sigma)^2 (2 - \sigma)\right)^{\frac{\beta}{\beta-2}}}{(\beta (2 - \sigma) - 4)^2 (2 - \sigma) (2 + \sigma)}, \quad (3.137)$$

which turns out to be always positive because of the following reason. A sufficient condition for $d \bar{N}(\beta, \sigma)|_{N=m} / d\sigma$ to be positive is that $\beta(2 - \sigma)^2 + \sigma^2 + 6\sigma - 8 > 0$, which in turn implies that $\beta > (8 - \sigma^2 - 6\sigma) / (2 - \sigma)^2$. Notice, however, that in region *B* (Figures 1-3) this is always the case since there $\beta \geq 4 / (2 - \sigma)$ and $\frac{4}{2 - \sigma} - \frac{8 - \sigma^2 - 6\sigma}{(2 - \sigma)^2} = \sigma \frac{2 + \sigma}{(2 - \sigma)^2} > 0$. The same will be true for regions *C* and *D* since in those regions β assumes even higher values.

3.A.2 Stand-alone stability of a duopoly

As is shown in Lemma 2, if we consider a coalition structure composed of two equal sized coalitions, then this duopoly coalition structure is always stand-alone stable in region *B* of Figure 1, i.e., if $4 / ((2 - \sigma)) \leq \beta \leq 2(2 + \sigma) / (2 - \sigma)$ (or equivalently, if $2(\beta - 2) / (\beta + 2) \leq \sigma \leq 2(\beta - 2) / \beta$). If, instead, $\beta > 2(2 + \sigma) / (2 - \sigma)$ (or, equivalently, $0 \leq \sigma < 2(\beta - 2) / (\beta + 2)$), regions *C* and *D* in Figures 1 and 2), then a duopoly coalition structure will only be stand-alone stable as long as condition (3.117) is satisfied. Let us focus on this latter case and start by studying the limiting properties of $\bar{N}(\beta, \sigma)|_{N=2m}$. Making use of (3.117), simple algebra shows that

$$\lim_{\sigma \rightarrow 0} \bar{N}(\beta, \sigma)|_{N=2m} = 2, \quad (3.138)$$

$$\lim_{\sigma \rightarrow \frac{2(\beta-2)}{\beta+2}} \bar{N}(\beta, \sigma)|_{N=2m} = +\infty. \quad (3.139)$$

Second, notice that from (3.117), one can also show that

$$\begin{aligned} \frac{d \bar{N}(\beta, \sigma)|_{N=2m}}{d\sigma} &= -2\beta \times 8^{\frac{\beta}{\beta-2}} \left(\frac{(1 + \sigma)^2}{(2 + \sigma)^3} \right)^{\frac{\beta}{\beta-2}} \times \\ &\quad \frac{-\beta^2 (2 - \sigma)^2 + \beta(1 - \sigma)(8 - \sigma)(2 + \sigma) + 2\sigma(7\sigma + \sigma^2 + 2) - 16}{(\beta - 2)(1 + \sigma)(2(2 + \sigma) - \beta(2 - \sigma))^2}, \end{aligned} \quad (3.140)$$

which turns out to be positive since $\forall \beta > 2 \forall \sigma \in \left[0, \frac{2(\beta-2)}{\beta+2}\right)$,
 $(-\beta^2(2-\sigma)^2 + \beta(1-\sigma)(8-\sigma)(2+\sigma) + 2\sigma(7\sigma + \sigma^2 + 2) - 16) < 0$.

3.A.3 The 'no-deviation to monopoly' requirement

Eq. (3.124) identifies a lower bound for the number of entrants in order for firms initially organized in a duopoly coalition structure not to find it profitable to jointly deviate towards the complete monopolization of the industry. Let us start by studying the limiting properties of this lower bound \underline{N}^{DM_k} :

$$\lim_{\sigma \rightarrow 0} \underline{N}^{DM_k} = k, \quad (3.141)$$

$$\lim_{\sigma \rightarrow \frac{2(\beta-2)}{\beta+2}} \underline{N}^{DM_k} = k \frac{1}{2} \frac{(\beta+2) 16^{\frac{\beta}{\beta-2}} \left(\beta^{\frac{\beta}{\beta-2}}\right)^2}{\left((\beta+2)^{\frac{\beta}{\beta-2}}\right)^3} = k \tilde{N}, \quad (3.142)$$

$$\lim_{\sigma \rightarrow \frac{2(\beta-2)}{\beta}} \underline{N}^{DM_k} = +\infty, \quad (3.143)$$

where \tilde{N} is given by eq. (3.135) and, as mentioned above, k denotes the size of the smallest coalition in the duopoly coalition structure, $1 \leq k \leq N/2$.

Next, notice that

$$\frac{d\underline{N}^{DM_k}}{d\sigma} = 4k \frac{\beta(\beta(2-\sigma)^2 + \sigma^2 + 6\sigma - 8) \left(\frac{1}{8}(2+\sigma)^2(2-\sigma)\right)^{\frac{\beta}{\beta-2}}}{(2-\sigma)(2+\sigma)(4-\beta(2-\sigma))^2}, \quad (3.144)$$

which is always positive since, as explained above (see discussion of eq. (3.137)), whenever $\beta \geq 4/(2-\sigma)$ (regions B , C and D in Figures 1-3), then one has that $(\beta(2-\sigma)^2 + \sigma^2 + 6\sigma - 8) > 0$, which is a sufficient condition for $d\underline{N}^{DM_k}/d\sigma > 0$.

3.B Upper bound to the number entrants anticipating a duopoly coalition structure

In this section, we briefly describe the properties of the upper bound to the number of firms $\bar{N}_\varepsilon^{D_k}$, given by eq. (3.133). These properties, justify the pattern of the family of curves $\bar{N}_\varepsilon^{D_1}, \bar{N}_\varepsilon^{D_2}, \dots$ presented in Figure 3.

First, from (3.133), simple algebra shows that

$$\lim_{\sigma \rightarrow 0} \bar{N}_\varepsilon^{D_k} = \left(\frac{(\beta - 2)}{2\varepsilon} \left(\frac{S}{4\beta} \right)^{\frac{\beta}{\beta-2}} + k \right) = \bar{N}_\varepsilon^M + k, \quad (3.145)$$

where \bar{N}_ε^M is given by eq. (3.129). In addition,

$$\lim_{\sigma \rightarrow \frac{2(\beta-2)}{\beta+2}} \bar{N}_\varepsilon^{D_k} = \left(\frac{(\beta - 2) S^{\frac{\beta}{\beta-2}} \left((\beta + 2)^{\frac{\beta}{\beta-2}} \right)^3}{\left(64^{\frac{\beta}{\beta-2}} \right) \left(\beta^{\frac{\beta}{\beta-2}} \right)^3 (\beta + 2) \varepsilon} + k \right), \quad (3.146)$$

$$\lim_{\sigma \rightarrow \frac{2(\beta-2)}{\beta+2}} \bar{N}_\varepsilon^{D_k} = k. \quad (3.147)$$

Next, notice that

$$\frac{d\bar{N}_\varepsilon^{D_k}}{d\sigma} = -\frac{1}{2}\beta \left(2 \frac{S}{\beta(2-\sigma)(2+\sigma)^2} \right)^{\frac{\beta}{\beta-2}} \frac{(2-\sigma)^2 \beta + \sigma^2 + 6\sigma - 8}{(2+\sigma)(2-\sigma)(\beta-2)\varepsilon}, \quad (3.148)$$

which turns out to be always negative since, as mentioned above (see discussion of eq. (3.137)), $(\beta(2-\sigma)^2 + \sigma^2 + 6\sigma - 8) > 0$, whenever $\beta \geq 4/(2-\sigma)$ (regions B, C and D in Figures 1-3).

Bibliography

- Abreu, D. (1986), "Extremal Equilibria of Oligopolistic Supergames", *Journal of Economic Theory* 39, 191-225.
- Abreu, D. (1988), "On the Theory of Infinitely Repeated Games with Discounting", *Econometrica*, Vol. 56, No. 2, 383-396.
- Abreu, D., Pearce, D. and Stacchetti, E. (1986), "Optimal Cartel Equilibria with Imperfect Monitoring", *Journal of Economic Theory* 39, 251-269.
- Barros, P. P. (1998), "Endogenous Mergers and Size Asymmetry of Merger Participants", *Economics Letters* 60, 113-119.
- Bernhardt, I. and Jung, B. S. (1979), "The interpretation of Least Squares Regression with Interaction or Polynomial Terms", *Review of Economics and Statistics* Volume 61, Number 3, August 1979, 481-483.
- Bernheim, B. D. (1984), "Strategic Deterrence of Sequential Entry into an Industry", *Rand Journal of Economics* 15, 1-11.
- Bernheim, B. D., Peleg, B. and Whinston, M. (1987), "Coalition-Proof Nash Equilibria: I. Concepts", *Journal of Economic Theory* 42, pp. 1-12.
- Brander, J. A. and Spencer, B. J. (1985), "Tacit Collusion, Free Entry and Welfare", *The Journal of Industrial Economics*, Vol. XXXIII, No. 3, 277-294.
- Bresnahan, T. (1992), "Sutton's Sunk Costs and Market Structure: Price Competition, Advertising and the Evolution of Concentration", *RAND Journal of Economics*, Vol. 23, No. 1, pp. 137-152.
- Briggs III, Hugh (1996), "Optimal Cartel Trigger Strategies and the Number of Firms", *Review of Industrial Organization* 11, 551-561.
- Caffarra, C. and Kühn, K.-U. (1999), "Joint Dominance: The CFI Judgment on Gencor/Lonrho", *European Competition Law Review*, 20/7, 355-359.
- Chwe, M. S.-Y. (1994), "Farsighted Coalitional Stability", *Journal of Economic Theory*, Vol. 63, pp. 299-325.

- Compte, O., Jenny, F. and Rey, P. (1997), "Capacity Constraints, Mergers and Collusion", *European Economic Review* (forthcoming).
- Dhrymes, P. J. (1986), "Limited Dependent Variables", Ch. 27 in Z. Griliches and M. D. Intriligator (Eds.) *Handbook of Econometrics*, Vol. III North Holland: Amsterdam.
- Diamantoudi, E. and Xue, L. (2001), "Farsighted Stability in Hedonic Games", *mimeo*
- Eaton, C. and Eswaran, M. (1998), "Endogenous Cartel Formation", *Australian Economic Papers*, Vol. 37, No. 1, 1-13.
- Ellison, G. (1994), "Theories of Cartel Stability and the Joint Executive Committee", *Rand Journal of Economics*, Vol. 25, No. 1, 37-57.
- Fisher, F. M. (1987), "Horizontal Mergers: Triage and Treatment", *Journal of Economic Perspectives*, Vol. 1, Number 2, 23-40.
- Friedman, J. W., (1971), "A Non-cooperative Equilibrium for Supergames", *Review of Economic Studies* 28, 1-12.
- Fudenberg, D. and Maskin, E., (1986) "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information.", *Econometrica*, Vol. 54, No. 3, 533-554.
- Fudenberg, D. and Tirole, J. (1989), "Noncooperative Game Theory for Industrial Organization: an Introduction and Overview", Ch.5 in: R. Schmalensee and R. Willig (Eds.), *Handbook of Industrial Organization*, Vol. I, North-Holland: Amsterdam.
- Fudenberg, D. and Tirole, J. (1991), *Game Theory*, The MIT Press.
- Geroski, Paul, Gilbert, Richard J. and Jacquemim, Alexis (1990), *Barriers to Entry and Strategic Competition*, Fundamentals of Pure and Applied Economics, Volume 41, Harwood Academic Publishers GmbH.
- Geroski, Paul (1995), "What Do We Know about Entry?", *International Journal of Industrial Organization* 13, 421-440.
- Green, E. J. and Porter, R. H. (1984), "Noncooperative Collusion under Imperfect Price Information", *Econometrica*, Vol. 52, No.1, 87-100.
- Green, W. H.. (1987), *Econometric Analysis*, 3d. Ed., Prentice Hall.
- Hajivassiliou, V. (1997), "Testing Game-Theoretic Models of Price Fixing Behaviour", LSE-STICERD Discussion Paper No. EM/97/324.

- Hart, S. and Kurz, M. (1983), "Endogenous Formation of Coalitions", *Econometrica*, Vol. 51, No. 4, pp. 1047-1064.
- Harrington, J. E. (1989), "Collusion and Predation under (almost) Free Entry", *International Journal of Industrial Organization* 7, 381-401.
- Harrington, J. E. (1991a), "The Joint Profit Maximum as a Free-Entry Equilibrium Outcome", *European Economic Review* 35, 1087-1101.
- Harrington, J. E. (1991b), "The Determination of Price and Output Quotas in a Heterogeneous Cartel", *International Economic Review*, Vol. 32, No. 4, 767-792.
- Kamien, M. and Zang, I. (1990), "The Limits of Monopolization through Acquisition", *Quarterly Journal of Economics*, Vol. 105 (2), pp. 465-499.
- Kamien, M. and Zang, I. (1991), "Competitively Cost Advantageous Mergers and Monopolization", *Games and Economic Behavior*, Vol. 3 (3), pp. 323-338.
- Kamien, M. and Zang, I. (1993), "Competing Research Joint Ventures", *Journal of Economics & Management Strategy*, Vol. 2, Number 1, pp. 23-40.
- Kühn, K.-U. and Motta, M. (1999), "The Economics of Joint Dominance", *mimeo*.
- Lambson, V. E. (1988), "Aggregate Efficiency, Market Demand, and the Sustainability of Collusion", *International Journal of Industrial Organization* 6, 263-271.
- Lee, L.-F. and Porter, R. H. "Switching Regression Models with Imperfect Sample Separation Information - with an Application on Cartel Stability", *Econometrica*, Vol. 52, No. 2, 391-418.
- MacAvoy, P. W. (1965), *The Economic Effects of Regulation*, Cambridge: MIT. Press.
- MacLeod, W. B. (1987), "Entry, Sunk Costs, and Market Structure", *Canadian Journal of Economics* 20, 140-151.
- MacLeod, W. B., Norman, G. and Thisse, J.-F. (1987), "Competition, Tacit Collusion and Free Entry", *The Economic Journal*, Vol. 97, 189-198.
- Martin, S. (1988), *Industrial Economics - Economic Analysis and Public Policy*, Macmillan Publishing Company, New York.
- Martin, S. (1990), "Fringe Size and Cartel Stability", Working Paper ECO 90/16, European University Institute.

- Maskin, E. and Tirole, J. (1988), "A Theory of Dynamic Oligopoly, II: Price Competition, Kinked Demand Curves, and Edgeworth Cycles", *Econometrica*, Vol. 56, No. 3, 571-599.
- Mason, C. F., Phillips, O. R. and Nowell, (1992), "Duopoly Behavior in Asymmetric Markets: an Experimental Evaluation", *Review of Economics and Statistics*, 662-670.
- McAfee, R. P. and Williams, M. A. (1992), "Horizontal Mergers and Antitrust Policy", *Journal of Industrial Economics* Vol. XL, No. 2, 181-187.
- McFadden, D. L. (1984), "Econometric Analysis of Qualitative Response Models", Ch. 24 in Z. Griliches and M. D. Intriligator (Eds.) *Handbook of Econometrics*, Vol. II North Holland: Amsterdam.
- Motta, M. (2000), "EC Merger Policy, and the Airtours case", *European Competition Law Review*, 21/4 (April 2000), 199-207.
- Nocke, V. (2000), "Monopolisation and Industry Structure", *mimeo*, Nuffield College, Oxford.
- Nti, K. O. and Shubik, M. (1981), "Noncooperative Oligopoly with Entry", *Journal of Economic Theory* 24, 187-204.
- Osborne, D. K. (1976), "Cartel Problems", *American Economic Review* 66, 835-844.
- Park, J. Y. and Phillips, P. C. B. (1999), "Nonstationary Binary Choice", *Econometrica* (forthcoming).
- Perry, M. and Porter, R. H. (1985), "Oligopoly and the Incentive for Horizontal Merger", *American Economic Review* 75, 219-227.
- Porter, R. H. (1983a), "Optimal Cartel Trigger Price Strategies", *Journal of Economic Theory* 29, 313-338.
- Porter, R. H. (1983b), "A Study of Cartel Stability: The Joint Executive Committee, 1880-1886", *Bell Journal of Economics*, Vol. 14, 301-314.
- Porter, R. H. (1985), "On the Incidence and Duration of Price Wars", *Journal of Industrial Economics*, Vol. XXXIII, No. 4, 415-426.
- Porter, R. H. (1986), "A Note on Tacit Collusion under Demand Uncertainty", *Canadian Journal of Economics*, Vol. XIX, No. 3, 556-558.

- Rothschild, R. (1999), "Cartel Stability When Costs Are Heterogeneous", *International Journal of Industrial Organization* 17, 717-734.
- Scherer, F. M. (1980), *Industrial Market Structure and Economic Performance*, Rand McNally College Publishing Company.
- Schmalensee, R. (1987), "Horizontal Merger Policy: Problems and Changes.", *Journal of Economic Perspectives*, Volume 1, Number 2, 41-54.
- Shaffer, S. (1995), "Stable Cartels with a Cournot Fringe", *Southern Economic Journal* 61, 744-754.
- Stenbacka, L. R. (1990), "Collusion in Dynamic Oligopolies in the Presence of Entry Threats", *Journal of Industrial Economics*, Vol. XXXIX, No.2, 147-154.
- Stigler, G. J. (1964), "A Theory of Oligopoly", *Journal of Political Economy*, Vol. 72, 44-61.
- Sutton, J. (1991), *Sunk Costs and Market Structure*, MIT Press.
- Sutton, J. (1998), *Technology and Market Structure*, MIT Press.
- Tirole, Jean (1988), *The Theory of Industrial Organization*, The MIT Press.
- Verboven, F. (1995), "Corporate Restructuring in a Collusive Oligopoly", *International Journal of Industrial Organization* 13, 335-354.
- Vives, X. (2000), *Oligopoly Pricing: Old Ideas and New Methods*, The MIT press.
- Wenders, J. T. (1971), "Collusion and Entry", *The Journal of Political Economy*, No. 79 (6), 1258-1277.
- Xue, L. (1998), "Coalitional Stability under Perfect Foresight", *Economic Theory* 11, 603-627.
- Yi, S.-S. (1997), "Stable Coalition Structures with Externalities", *Games and Economic Behavior* 20, 201-237.

