Price Competition, Mergers and Structural Estimation in Oligopoly

Alberto Salvo

Submitted for the degree of Ph.D.

London School of Economics

May 2005

To Fernanda,
for her unconditional support

Abstract

This thesis examines the exercise of market power by oligopolistic firms.

The first part deals with a phenomenon that has important implications for market power: horizontal mergers. I seek to uncover why the pattern of equilibria in sequential merger games of a certain type is similar across a fairly wide class of models studied in the literature. By developing general conditions characterising each element of the set of possible equilibria, I show that the solution to models that satisfy a certain sufficient condition will be restricted to the same subset of equilibria. This result is of empirical relevance in that the pattern of equilibria obtained for this class of models is associated with mergers happening, not in isolation, but rather in bunches. I extend the results to the analysis of cross-border mergers, studying two standard models that satisfy my sufficient condition: Sutton's (1991) vertically-differentiated oligopoly and Perry and Porter's (1985) fixed-supply-of-capital model.

The second part is concerned with the structural inference of market power, a central theme in empirical Industrial Organisation. I demonstrate that when an industry faces potential entry and this threat of entry constrains pre-entry prices, cost and conduct cannot be identified from the comparative statics of equilibrium. In such a setting, the identifying assumption behind the well-established technique of relying on exogenous demand perturbations to distinguish empirically between alternative hypotheses of conduct is shown to fail. The Brazilian cement industry, where the threat of imports restrains market outcomes, provides an empirical illustration. In particular, price-cost margins estimated using this established technique are biased heavily downwards, underestimating the degree of market power. I propose a test of conduct, adapted to this constrained setting, which suggests that outcomes in the industry are collusive and characterised by market division.

Robustness of this result is verified along several dimensions: by considering simple dynamic multimarket games which in equilibrium give rise to market division; by reviewing the spatial competition literature; and by resorting to a gravity model to statistically analyse shipments.

Keywords: Mergers; Sequential mergers; Cross-border investment; Technology transfer; Conduct; Multimarket competition; Market division; Limit pricing; Cement

JEL classification: D43, F14, F23, G34, L13, L41, L70

Contents

List of Figures

Acknowledgements

This thesis, as well as its author, has benefited to an unquantifiable extent from the help, encouragement and generosity of many people.

My main intellectual debt lies with my supervisors, John Sutton and Peter Davis. I have learnt much from John's scientific rigour, integrity and professionalism, teaching me the wisdom of constantly challenging one's own views, searching for competing hypotheses and submitting one's theses to honest robustness tests. I cannot think of a meeting with John from which I walked away not having gained something substantial. And I recall him stating repeatedly that the IO field stood only to benefit from its researchers visiting plants, offices and stores more often. John has also provided invaluable organisational support, for which I am very grateful. Peter's arrival at the LSE in the second year of my Ph.D. opened up a whole new world of exciting research, marked by the transition from the game-theoretic oligopoly models of Chapter 1 to the structural estimation using industry-specific data of Chapter 2. Through Peter I realised that there was a pay-back to investing a tremendous amount of effort learning about the institutional details of a single industry and collecting original data. I recall his strategic advice early on that this was a sound bet for an IO graduate student hoping to make a contribution and go on the U.S. job market. Peter's approachability cannot be exaggerated, and I thank him for his friendship.

At the LSE and STICERD I have been fortunate to interact with people from different walks of life. Among colleagues and fellow graduate students, I thank Nick Bloom, Antonio Buttà, Fernando Cabello-Astolfi, Jad Chaaban, Lapo Filistrucchi, Kaisa Kotakorpi, Barbara Luppi, Maria Luisa Mancusi, Steve Redding, Helder Vasconcelos, among others, and above all Thomas Buettner and Cristian Huse for the regular exchange of ideas and mutual support. Frank Cowell was particularly encouraging in my early days of research. Mark Schankerman's supportive comments following one of my first and tough-going Ph.D. seminars are gratefully remembered. Martin Pesendorfer kindly agreed to meet on occasion to talk over different parts of my thesis. Friday lunches, one-on-one meetings and dinners – hosted and organised by John Sutton and Peter Davis – provided the invaluable opportunity to benefit on a personal level from the visit of outside seminar speakers. My long-time friend, and job-market veteran, Bernardo Guimarães was a particular source of encouragement in this final year. And I cannot forget to acknowledge the invaluable administrative support provided during the Ph.D. years and the job market experience most notably by Leila Alberici, Jane Dickson, Nora Lippincott, Emma Taverner, Gill Wedlake and Mark Wilbor. Nor do I forget Nic Warner's and Irina Verkhova's IT support.

Regarding feedback on my research, I would further like to thank Roman Inderst and Massimo Motta for their comments on Chapter 1. For comments on chapters 2 and 3 I am also grateful to seminar audiences at Università Bocconi, Fundação Getulio Vargas (EESP and EPGE), Institute for Fiscal Studies, Kellogg School of Management, Northeastern University, Universitat Pompeu Fabra, PUC Rio, UBC Sauder School of Business, NYU Stern School of Business and University of Warwick, in addition to seminar audiences at the 2004 Meetings of the Brazilian Econometric Society and the Brazilian Antitrust Authorities (CADE and SDE).

I gratefully acknowledge financial support in the course of my graduate studies, namely from the British Foreign and Commonwealth Office, from the LSE Economics Department and from the Brazilian Ministry of Education (CAPES). I am also grateful to STICERD for providing research facilities.

Every empirical IO researcher is acquainted with the sheer difficulty of gathering private industry data, especially in the absence of research funds. For reasons of confidentiality (and brevity) I cannot list all the people who have helped me in putting together different datasets. On the cement project which forms the core of my thesis, an acknowledgement to numerous anonymous sources is stated in Appendix B. I am indebted to Professor José Vicente Caixeta Filho of ESALQ, at the University of São Paulo, for providing an extract of the SIFRECA database on freight prices. I would also like to thank Ricardo Fort at the Coca-Cola Company and Bruno Gouvêa at AC Nielsen for access to a valuable dataset on soft drinks which, in spite of not making it into my thesis, is exciting work in progress.

My final words of acknowledgement go to my family in general and to four people in particular. My dear parents, Segundo (who passed away in August 2003) and Maria Angeles, for doing all they could to ensure that I received the best possible education, in the widest sense of the word. My son Tomás, for the joy of his birth only five months ago and for providing an added sense of perspective. My beloved wife Fernanda, to whom I dedicate this thesis, for all her affection, understanding and selflessness through the years.

Introduction

This thesis revolves around the exercise of market power by firms in an oligopolistic industry. Its chapters are of both an applied theoretical and an empirical nature, falling into either of two parts. The first part, Chapter 1, deals with mergers between firms in a particular form of theoretic game, that where different groups of firms within an industry make merger decisions in sequence. It is well known among researchers and policymakers that mergers have important implications on the exercise of market power. The second part, contained in Chapters 2 and 3, is concerned with the measurement of market power. It analyses the performance of a well-established empirical methodology in a general setting where the ability of an industry to set prices is constrained by the threat of entry. This analysis is first presented theoretically and is then tested empirically by reference to a real-world industry, the Brazilian cement industry. While the first part takes the model of firm pricing behaviour – or firm conduct – as given to study the profitability and prevalence of mergers, the second part deals with the econometrician's task of ascertaining the model of firm conduct from observable market data.

Following the somewhat puzzling result by Salant, Switzer and Reynolds (1983) that an exogenous merger in a homogeneous-goods Cournot oligopoly tends to be unprofitable, in view of the free-riding by firms not participating in the merger, several papers were published showing that this "merger unprofitability" result was not robust to the relaxation of different assumptions in Salant et al's model. By either introducing some differentiation among products, making marginal cost rise in output, adding capacity constraints, or switching to price-setting firms with upward-sloping reaction functions, such free-riding by non-participants to the merger may be contained or eliminated (see, for example, Perry and Porter 1985, and Deneckere and Davidson 1985). More recently, a string of papers has shown that by allowing firms to merge in sequence, the profitability of mergers can be further enhanced (see, for example, Kamien and Zang 1993, and Fauli-Oller 2000).

As it happens, in a certain type of sequential merger game studied in the literature a similar pattern of equilibria is recurring under fairly broad industry conditions, including different demand or cost setups, or different merger technologies. In this pattern of equilibria, only a strict subset of the set of possible equilibria is obtained and this subset is such that when a merger occurs along the equilibrium path of the game it occurs alongside other mergers. In other words, the equilibrium outcomes of the sequential merger game across different standard models is such that either *all* groups of firms that are allowed to merge choose to do so or that *none* of the groups of firms choose to merge. In particular, outcomes where only some groups of firms choose to merge while other groups choose not to are not supported in equilibrium. Chapter 1 seeks to uncover why this is so. I show that the solution to models that satisfy a certain sufficient condition will be restricted to the same subset of equilibria that characterises the "all-or-none" merger result. I motivate this sufficient condition by resorting to the way non-participating firms react to the merger of rival firms. In addition to examples of this type of sequential merger game already studied in the literature, I develop two other standard examples – Sutton's (1991) vertically-differentiated Cournot oligopoly and Perry and Porter's (1985) quadratic-cost (fixed-supply-of-capital) model – in which the sufficient condition holds. As such, the same pattern of equilibria obtains. I note that this "all-or-none" finding may be of empirical relevance in that mergers are commonly observed to "bunch" together, where periods of high merger activity are followed by periods of low activity.

Industrial organisation researchers and antitrust practitioners have long been concerned with measuring the degree of market power enjoyed by firms in an industry. When marginal cost is observed, market power – or firm conduct – can be inferred from the distance between price and cost, i.e. the price-cost margin. It is usually the case, however, that marginal cost is not observed. A standard methodology has evolved in the past decades that attempts to infer conduct, and cost, from the comparative statics of equilibrium (see, for example, Bresnahan 1989). Stated simply, the way equilibrium prices vary as demand conditions move exogenously reveal the degree of market power. Consider two extreme examples by way of providing intuition. In a competitive industry, firms set output at the point where price equals marginal cost. At the other extreme, a collusive industry, or a cartel, changes prices such that marginal revenue equals marginal cost.

Chapter 2 revisits this standard methodology to show that in a more general setting, where the threat of entry constrains the ability of an industry to set prices, the methodology does not work, in the sense of consistently estimating the true price-cost margins. One can motivate the threat of entry constraining pre-entry prices by reference to a domestic industry facing potential competition from high-cost imports. (An alternative motivation, among others, can be a credible threat of investigation by an antitrust authority that imposes downward pressure, in unobserved ways, on an oligopoly's prices.) Intuitively, the response of prices to fluctuations in demand is no longer distinct according to whether firms enjoy market power or not because the threat of imports, or entry more generally, acts to constrain the ability of firms with market power to equate marginal revenue to marginal cost. Why is this important? I show that a researcher attempting to infer the degree of market power who overlooks the threat of entry when this threat constrains market outcomes (on a non-trivial subset of his observations), will underestimate the degree of market power – finding more competition when there is less. This is of high relevance to antitrust. It is also increasingly relevant in a world where trade barriers are being pulled down.

I illustrate by reference to the Brazilian cement industry, where potential (high-cost) imports restrain market outcomes. That is, imports impose a price ceiling which binds at the equilibrium, such that no or few imports are observed. I take on the role of this researcher and begin by overlooking the latent effect of imports, employing the standard methodology to estimate conduct and cost. The supply estimates I obtain are indicative of domestic competition, with estimated price-cost margins centred around zero. It turns out that one reason why I use the cement industry is that its simple technology allows me to *observe* marginal cost so that I can check these estimates. Cement is a given amount of limestone, a given amount of thermal energy to fire up the kiln, a given amount of electrical energy to grind the intermediate product and in my data I observe the flow of cement from each plant to each market, allowing me to calculate the cost of freight. I thus compute actual marginal cost from engineering coefficients, the observed flow of cement and observed factor prices. The true price-cost margins are far from competitive, amounting to around 50% of producer prices. This empirically confirms my theoretical proposition that the standard methodology in such settings yields biased estimates in the direction of more competition.

This result delivers a message of caution. That one must employ the standard methodology more cautiously than has arguably been done in the past. I then address the following question. If the standard methodology fails in a setting where market outcomes are constrained, how can we then learn about conduct, even if we are prepared to require more information? I propose a test of conduct for such constrained settings against a standard oligopoly benchmark, the Cournot solution. I illustrate this test again by reference to the Brazilian cement industry. I find that outcomes in the industry are more collusive than the Cournot benchmark, and can be characterised by tacit market division. This is identified despite the threat of imports restraining prices. In addition to illustrating the proposed test of conduct, such an exercise reiterates the rejection of competitive domestic behaviour, as suggested by the standard methodology when the latent effect of entry is overlooked.

Chapter 3's recurring theme is the testing of the robustness of the preceding chapter's findings. I begin by considering, in light of the different local market structures observed in the Brazilian cement industry, simple dynamic multimarket games which in equilibrium give rise to the pattern of market division identified previously. The aim in providing examples of such dynamic games is to indicate the rationality of a strategy of market division in an industry where firms meet in different (geographic or product) markets. I then turn to the spatial competition literature to ensure the validity of the theoretical framework of Chapter 2. The chapter then outlines several robustness and specification tests of the structural modelling estimates of Chapter 2. Finally, to complement the earlier structural estimation approach, I fit a gravity model to statistically analyse the flow of cement between plants and local markets. The distance from plants to markets, market size and plant ownership, among other variables, are highly correlated with – or "explain" to a considerable degree – the observed shipments of cement from plants to local markets. Yet there are interesting outliers, which I explore. One can tentatively attribute such residual variation to firm behaviour which differs from an "average" pattern of industry behaviour. Importantly, the exercise illustrates a point which, to the best of my knowledge, has not been made in the trade literature's use of gravity models in explaining the flow of trade and investment between countries or regions: that estimated distance effects may also be proxying for firms' strategies of dividing (geographic) markets, as is the case in the Brazilian cement industry. This is yet another manifestation of the classic identification problem in the estimation of supply: distinguishing firm costs from firm conduct.

The spatial dimension to price competition is present throughout both parts of the thesis. The analysis of sequential merger games in the first part is extended to the study of cross-border mergers, by way of two standard oligopoly models (which satisfy the sufficient condition of the chapter) that are cast in an international (or inter-regional) setting. In both, increasing horizontal product differentiation in the form of trade costs serves to limit the output expansion by non-participating firms that adversely affects the participants to a merger. The spatial dimension is again present in the second part by modelling the threat of entry facing the (domestic) oligopoly (whose conduct the econometrician wishes to infer) by way of the competitive supply of high-cost imported product, which sets a ceiling on the pricing outcomes of the industry, such that in equilibrium this ceiling binds and no foreign entry is observed. And in the structural estimation of the cement industry one must carefully control for the effects of distance, given that transportation and location are important components of the industry's cost structure and the strategic interaction among firms.

Another topic common to both parts of the thesis relates to the implications for antitrust policy and how these cut across national (or regional) boundaries. The cross-border setting in which the two examples of Chapter 1 are framed highlights an aspect of "acquisition investment" which is not present in traditional studies of foreign investment in the "greenfield" form. The latter tend to highlight the tariff-jumping (or trade-cost avoidance) rationale of foreign investment as against foreign trade (exports). But on top of such a driving force, by merging or buying their way into foreign markets firms can reduce the number of rivals and enjoy muted competition, a feature which is absent in simple exports-versus-greenfield investment studies. As for the theoretical framework and the empirical illustration of the second part, the competitive threat of foreign supply impinges on domestic prices and consumer welfare, in addition to having important implications for the inference of the degree of market power, a core element of an antitrust practitioner's toolkit. The way in which cement producers divide markets again illustrates that firms' competitive strategies cut across geographic boundaries.

I postpone further remarks on the two parts of the thesis to each chapter individually – these include thorough and self-contained introductory sections providing an overview and contributions of each chapter.

Chapter 1

A General Analysis of Sequential Merger Games, with an Application to Cross-Border Mergers

1.1 Introduction

Several papers have studied the interdependence of horizontal mergers by modelling a particular type of sequential game of merger formation (Fauli-Oller 2000, Matsushima 2001, Motta and Vasconcelos 2003). In this type of game, merger decisions are made in sequence by different and exogenous groups of firms, ending in a product market competition stage. Despite employing different setups, such as different demand or cost conditions, or different merger technologies, the pattern of equilibria in the sequential game is similar across these models. As parameter values vary, only elements from the same subset of the set of possible equilibria is obtained. The equilibrium outcomes corresponding to the equilibria obtained in all of these papers is such that either *all* groups of firms that are allowed to merge choose to do so or that *none* of the groups of firms choose to merge. In particular, outcomes where only some groups of firms choose to merge while other groups choose not to are not supported in equilibrium. As will be noted, this "all-or-none" finding may be of empirical relevance in that mergers are commonly observed to "bunch" together, where periods of high merger activity are followed by periods of low activity.

Fauli-Oller (2000) considers cost-asymmetric firms facing linear demand under Cournot competition. In the merger stages, prior to the final market competition stage, low-cost firms are allowed to bid in sequence for their high-cost counterparts. In a spatial context, Matsushima (2001) studies sequential mergers by placing firms symmetrically around a unit circle, with firms facing linear demand at each consumer location and engaging in Cournot competition. Sequential pairwise mergers between firms located diametrically opposite one another are then allowed. Motta and Vasconcelos (2003) consider an industry with a fixed supply of capital distributed among its firms, with firms' marginal costs being a decreasing function of the capital they own[1]. Exogenous groups of firms decide in sequence on whether to merge, leading up to a market competition stage. What is of interest to me in the current chapter is that, despite the different setups of each model, the equilibria obtained are of underlying similarity, in the sense that only a similar subset of the set of possible equilibria of the sequential merger game is obtained.

This chapter makes two contributions. The first aim is to uncover this underlying similarity in the pattern of equilibria across these models studied in the literature. By writing conditions on firms' reduced-form profit functions, I characterise each possible equilibrium of the sequential merger game. I explain why only a similar subset of the set of possible equilibria is obtained in these models by reference to a sufficient condition which holds across these models. When this condition is satisfied, a common subset of the set of possible equilibria is ruled out, considerably simplifying the solution to the game. I motivate this sufficient condition by resorting to the way non-participating firms react to the merger of rival firms, as is well understood in the literature (see Salant, Switzer and Reynolds 1983 in the case of Cournot competition, and Deneckere and Davidson 1985 in the case of Bertrand competition).

[1]This feature of their model is thus similar to the second example presented in this chapter, based on Perry and Porter (1985).

The second aim is to illustrate that the class of models for which this similar pattern of equilibria is obtained is fairly broad. To this end, I develop two other standard examples – Sutton's (1991) vertically-differentiated Cournot oligopoly and Perry and Porter's (1985) quadratic-cost (fixed-supply-of-capital) model – in which the sufficient condition holds. As such, the same pattern of equilibria as that in, say, Fauli-Oller (2000) obtains. By introducing an additional parameter (namely a trade cost between two countries), I frame the sequential merger game in both examples in a cross-border context. I thus extend the results to the analysis of cross-border mergers.

Adapting a vertically-differentiated Cournot oligopoly model due to Sutton (1991) and later applied in a trade context by Motta (1992, 1994), I allow high-quality firms located in one country to merge with low-quality firms located in the second country. The merger technology is such that if a merger is undertaken, the level of quality offered by the merged (multinational) firm is the higher of the qualities offered by its constituent firms. Further, goods produced in one country but sold in another (exports) incur (linear) trade costs. Thus cross-border mergers enable firms to transfer technology.

The second example is an adaptation of Perry and Porter (1985)'s quadratic-cost model. This model is of particular interest for the following reason. The majority of models (with single-product firms) in the merger literature considers a merged firm to be "about" the same size (bar the presence of merger synergies) as each of its constituent firms and the non-participating firms; there is no notion of assets or firm size. In contrast, in the Perry and Porter model, a merger results in a "new firm that is 'larger' than the others" (pp. 219). They model this by assuming a fixed factor of production (say capital) whose total supply is fixed to the industry; what distinguishes firms is the amount they own of this factor[2]. In the adaptation, I again allow mergers between firms located in the different countries, where the capital stock of the merged (multinational) firm is the sum of the capital stock of

[2] In the Perry and Porter (1985) model, on which my example is based, marginal cost is linear in output; by increasing its capital stock, the firm reduces the slope of its marginal cost curve. In Motta and Vasconcelos (2003), marginal cost is flat in output; a higher capital stock lowers marginal cost, regardless of output.

its constituents. Thus, the "larger" multinational firm has a lower marginal cost than either of its constituents at a given level of output.

This all-or-none merger result is of empirical relevance in that the pattern of equilibria found for this fairly wide – as I claim – class of models is associated with mergers happening, not in isolation, but rather in bunches. Empirical studies have tested for bunching, or wave-like behaviour, using aggregate industry data (see, for example, Town 1992, Golbe and White 1993, and Barkoulas et al 2001) while other authors have argued that this phenomenon is also observed within individual industries (Mueller 1989). Casual observation suggests, moreover, that this kind of bunching phenomenon appears to be of relevance also in the case of cross-border merger activity (UNCTAD 2000, Knickerbocker 1973). Caves (1991) offers two lines of reasoning to explain why mergers may bunch together. One line attributes this to an unprecedented shift in an exogenous industry- or economy-wide parameter suddenly making mergers profitable[3]. The other line looks at the strategic interaction between firms' merger decisions within an industry: under certain conditions firms will merge only if rivals merge. It is the latter line of reasoning which sequential merger games have sought to address.

Using similar notation and structure to the sequential merger game laid out in Nilssen and Sørgard (1998), in the following section I spell out general conditions – in terms of firms' reduced-form profit functions – characterising each possible equilibrium. I show that if a certain sufficient condition holds then a subset of the possible equilibria can be ruled out. I briefly illustrate the general analysis by reference to Fauli-Oller (2000). I then turn to this chapter's application, developing two further standard examples in which the sufficient condition holds, and showing that the general analysis can be extended to the case of cross-border mergers. Section 1.3 lays out and solves the sequential merger game in the vertically-differentiated Cournot oligopoly. Results are discussed in light of the international trade and investment literature, where foreign investment has been modelled mostly in the form of greenfield investment, as opposed to mergers and acquisitions. I argue that

[3]On this first line of reasoning, Stigler (1950) and Bittlingmayer (1985) analyse the effects of changes in competition policy. Van Wegberg (1994) studies how the business cycle affects the profitability of mergers.

modelling foreign investment in the form of cross-border mergers suggests just how widespread such activity may be. Considerations are also made on the robustness of results to the number of merger stages and on the specification of the extensive form of the merger game (i.e. sequential moves in counterpoint to merger decisions being made simultaneously). Section 1.4 considers the fixed-capital-stock model of Perry and Porter (1984), to show that similar results hold. In Section 1.5 I revisit the vertically-differentiated oligopoly model to introduce fixed costs associated with implementing a merger. In so doing I attempt to address an inherent weakness in the results of Sections 1.3 and 1.4 as to their usefulness in explaining the typically-observed flurry of merger activity as countries integrate their markets: if, as the results in these earlier sections suggest, cross-border mergers occur along the equilibrium path for a "large" region of parameter space, then why were these mergers not undertaken prior to market integration? Section 1.6 briefly concludes.

1.2 General conditions for equilibria in sequential merger games

I begin by defining the structure of the sequential merger game, following Nilssen and Sørgard (1998)[4]. In brief, disjoint and exogenously-given groups of firms make sequential merger decisions, leading up to a final market competition stage, where payoffs of the merged and independent firms are conditional on market structure resulting from the earlier merger stages. Bearing in mind the later aim of extending the results to a cross-border context, I adapt the notation and structure to a two-country setting. (I also point out how the analysis easily collapses to the single-country case.)

Consider two countries, $l \in \{A, B\}$, each initially with n^l firms, where $n^l \geq 2$. Denote the set of firms located in country A as $\mathcal{A}$ and the set of B-country firms as $\mathcal{B}$; thus $n^A = |\mathcal{A}|$, $n^B = |\mathcal{B}|$ and the global industry consists of $n := n^A + n^B \geq 4$

[4]Where convenient, I use similar notation and structure to that in Nilssen and Sørgard (1998)'s general discussion of sequential merger decisions. This is done for the benefit of the reader who is acquainted with their paper and also to point out how this section builds on their work. Notice that there is no repeated product market interaction as in Kamien and Zang (1990, 1993).

firms. Now consider the following form of sequential (cross-border) merger games. Assume that two cross-border mergers[5] can take place in sequence in this industry: label the set of firms that first decide whether to merge $\mathcal{M}_1$, and the set of firms that subsequently decide whether to merge $\mathcal{M}_2$. Label the merged firms possibly arising from the first and second merger decisions M_1 and M_2 respectively. The extensive form of the merger game in consideration satisfies the following assumptions:

- At least two firms take part in each merger decision, i.e. $|\mathcal{M}_i| \geq 2$, $i \in \{1, 2\}$.
- At least one firm from each country takes part in each merger decision, i.e. $\mathcal{A} \cap \mathcal{M}_i \neq \emptyset$ and $\mathcal{B} \cap \mathcal{M}_i \neq \emptyset$, $i \in \{1, 2\}$.
- Each firm participates in at most one merger decision, $\mathcal{M}_1 \cap \mathcal{M}_2 = \emptyset$, so that $|\mathcal{M}_1| + |\mathcal{M}_2| \leq n$.

An example is in order: country A has three firms, labelled 1, 2 and 3, while country B has two firms, labelled 4 and 5. The sequential merger game given by firms 1 and 4 first deciding whether to merge, followed by firms 2 and 5 deciding whether to merge, would satisfy the above assumptions. If both mergers are carried through, the market structure is given by two (multinational) merged firms M_1 and M_2, where $\mathcal{M}_1 = \{1, 4\}$ and $\mathcal{M}_2 = \{2, 5\}$, and an independent firm located in country A, firm 3.

Coming out of the merger stages, four situations (market structures) are possible (see Figure 1-1). Denote the situation where neither merger is undertaken by s_0. If the firms in $\mathcal{M}_1$ do not merge but those in $\mathcal{M}_2$ do, call this situation s_1; in contrast, let situation s_2 depict a favourable merger decision by firms in $\mathcal{M}_1$ followed by a no-merger decision by firms in $\mathcal{M}_2$. Finally, situation s_3 denotes the market structure arising from two favourable merger decisions[6].

Finally, the firms which have merged or remained independent compete on the world market; goods produced in one country can be exported for consumption

[5]The choice of *two* mergers is done for ease of exposition, as in Nilssen and Sørgard (1998). In the application of Section 1.3 I consider a set-up where three mergers can take place, in addition to the case with two mergers, and discuss the implications of allowing greater numbers of mergers to take place.

[6]Note that situations s_1 and s_2 are reversed as compared to the use of notation by Nilssen and Sørgard (1998).

in the other country. This completes the description of the type of sequential cross-border merger games being considered[7].

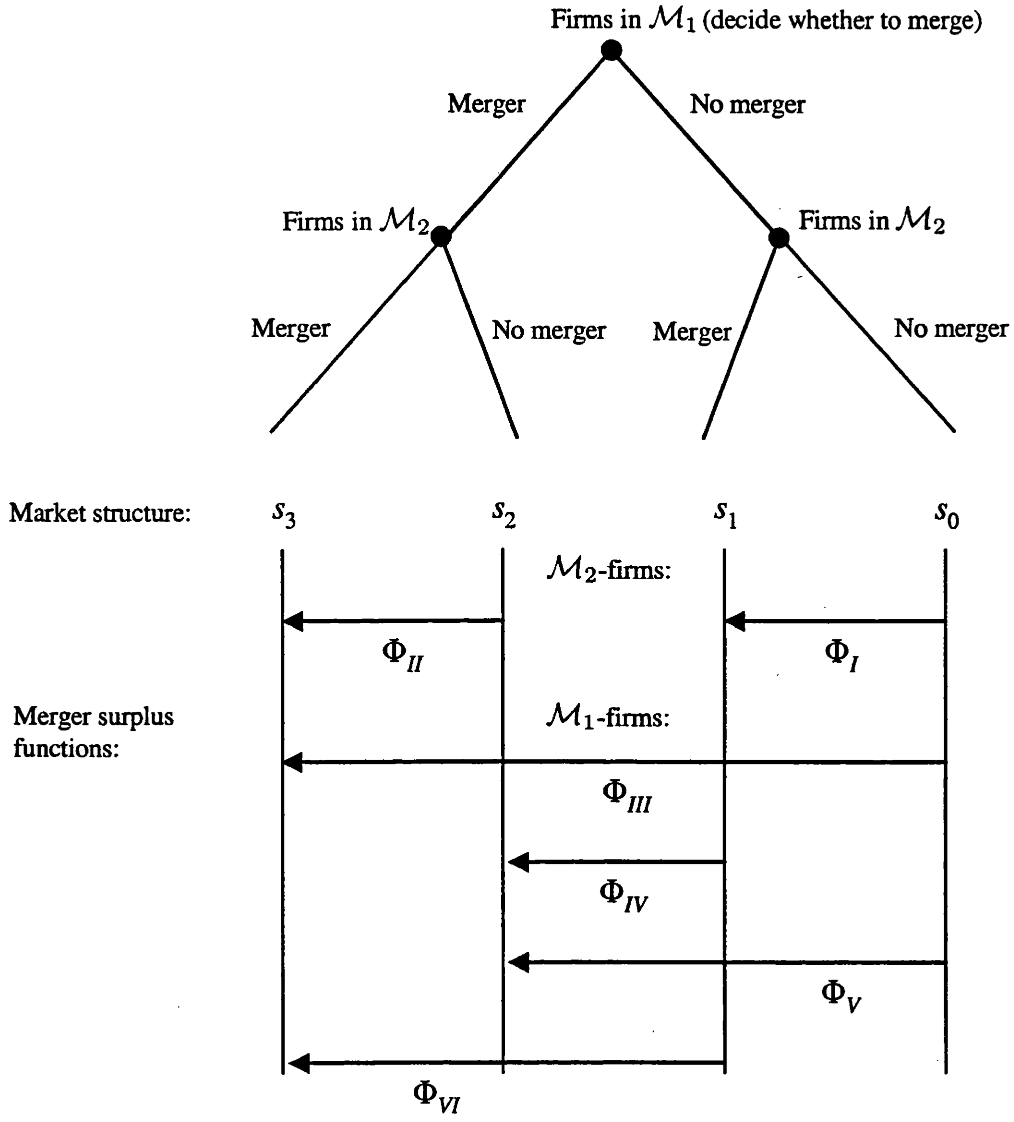


Figure 1-1: Sequential merger game and merger surplus functions

[7]As I comment in footnote 16, if there is no geographic dimension to the game and a single-country setting is being considered, the second assumption, namely that at least one firm from each country takes part in each merger decision, is dropped. There is no longer a distinction between firms located in one country or the other and the relevant set of firms is $\mathcal{I} := \mathcal{A} \cup \mathcal{B}$.

1.2.1 Reduced-form profit functions and conditions

Denote the equilibrium payoff to firm i under market structure s by $\Pi_{i,s}$, where $s \in \{s_0, s_1, s_2, s_3\}$ and $i \in \mathcal{S} \subset \mathcal{A} \cup \mathcal{B} \cup \{M_1, M_2\}$, with $\mathcal{S}$ being the set of (merged and/or independent) firms competing under market structure s. Note that $\Pi_{i,s}$ is in reduced form, indexed by a vector of "primitives" of the specific model in question, such as demand and supply-side parameters[8]. The merger surplus functions are then defined in terms of these reduced-form profit functions.[9]

Definition 1.1 *(Merger surplus functions) Define six Φ functions as follows:*

$\Phi_I := \Pi_{M_2,s_1} - \sum_{i \in \mathcal{M}_2} \Pi_{i,s_0}$ $\qquad$ $\Phi_{IV} := \Pi_{M_1,s_2} - \sum_{i \in \mathcal{M}_1} \Pi_{i,s_1}$

$\Phi_{II} := \Pi_{M_2,s_3} - \sum_{i \in \mathcal{M}_2} \Pi_{i,s_2}$ $\qquad$ $\Phi_V := \Pi_{M_1,s_2} - \sum_{i \in \mathcal{M}_1} \Pi_{i,s_0}$

$\Phi_{III} := \Pi_{M_1,s_3} - \sum_{i \in \mathcal{M}_1} \Pi_{i,s_0}$ $\qquad$ $\Phi_{VI} := \Pi_{M_1,s_3} - \sum_{i \in \mathcal{M}_1} \Pi_{i,s_1}$

Figure 1-1 illustrates the six merger surplus functions. To exemplify, Φ_{III} is the difference between the profit of merged firm M_1 under market structure s_3 and the sum of profits of its constituent (independent) firms under market structure s_0. As shown below, these six functions capture the surplus behind all the relevant merger decisions by either firms in $\mathcal{M}_1$, in the first stage, or firms in $\mathcal{M}_2$, in the second stage.

It follows from the consideration of two sequential merger stages that there are eight possible sets of Nash strategies by the players. These are illustrated in Figure 1-2[10]. In sets (a), (b), (g) and (h), the early firms' merger decision has no effect on the subsequent firms' merger decision. By way of example, for set (a)

[8] Formally, the reduced-form payoff function is $\Pi_{i,s}(\theta)$, where $\theta \in \mathcal{P}$ is the vector of parameters capturing relevant features of the model and $\mathcal{P}$ is the space of parameters. For ease of exposition, I suppress the argument θ of the functions.

[9] The notation here departs from that in Nilssen and Sørgard (1998) for the following reason. While I define six functions, for the purposes of their paper they define eight "profitability of merger" functions for merger participants, in addition to four such functions for non-participating parties to a merger. For the sake of comparability, note that Φ_I, Φ_{II}, Φ_{III}, Φ_{IV}, Φ_V and Φ_{VI} in this chapter are equivalent to Δ_2^1, Δ_2^2, Δ_1^3, Δ_1^4, Δ_1^1 and Δ_1^2 respectively, in their paper. To simplify the interpretation of results, I choose to refer the reader to Figure 1-1 (and thus number the Φ functions in such a way) rather than ask him to recall the logic behind Nilssen and Sørgard's use of subscripts and superscripts.

[10] To avoid clutter, I rely on Figure 1-2 to define each of the eight sets of Nash strategies, rather than doing so in words. In any case, this is standard. For example, set (c) would be defined as follows. $\mathcal{M}_1$-firms: 'merger'; $\mathcal{M}_2$-firms: 'merger' only if $\mathcal{M}_1$-firms merge, otherwise 'no merger'.

to be an equilibrium, firms in $\mathcal{M}_2$ need to find it profitable to merge irrespective of $\mathcal{M}_1$-firms' decision whether to merge or not. Also, firms in $\mathcal{M}_1$ need to find it profitable to merge in anticipation of $\mathcal{M}_2$-firms' subsequent merger decision.

In sets (of strategies) (c), (d), (e) and (f), on the other hand, the subsequent firms' merger decision does depend on the early firms' merger decision. If set (c) characterises an equilibrium, then firms in $\mathcal{M}_2$ find it profitable to merge *only if* firms in $\mathcal{M}_1$ merge. (To complete the characterisation of this equilibrium, firms in $\mathcal{M}_1$ must find it profitable to merge, anticipating how the subsequent merger decision depends on their own decision.) As in set (a), mergers occur along the equilibrium path of the game yet here $\mathcal{M}_1$-firms' merger decision has a bearing on (and takes place anticipating) $\mathcal{M}_2$-firms' merger decision[11].

Proposition 1.1 states necessary and sufficient conditions for each of these sets of strategies to form an equilibrium (unique for every combination of parameter values)[12].

Proposition 1.1 *(NSC for equilibria) Necessary and sufficient conditions for each of Nash-strategy sets (a) to (h), as depicted in Figure 1-2, to be supported as the unique equilibrium (for a given combination of parameter values) are:*

Set (a) $\Phi_I \geq 0$, $\Phi_{II} \geq 0$ *and* $\Phi_{VI} \geq 0$

Set (b) $\Phi_I \geq 0$, $\Phi_{II} \geq 0$ *and* $\Phi_{VI} < 0$

Set (c) $\Phi_I < 0$, $\Phi_{II} \geq 0$ *and* $\Phi_{III} \geq 0$

Set (d) $\Phi_I < 0$, $\Phi_{II} \geq 0$ *and* $\Phi_{III} < 0$

Set (e) $\Phi_I \geq 0$, $\Phi_{II} < 0$ *and* $\Phi_{IV} \geq 0$

[11] Some authors, such as Fauli-Oller (2000), have related results (c) and (d) to the merger-wave phenomenon, in the sense that an early decision to merge (either along the equilibrium path of the game, as in (c), or as a deviation from it, as in (d)) triggers subsequent mergers.

[12] Proposition 1.1 is consistent with Proposition 1 in Nilssen and Sørgard (1998; pp.1689). One must be careful, however, in comparing the conditions in their Propostion 2 and Table 1 (pp.1690-1) with mine. For example, they state that "the occurrence of a Fat Cat strategy implies a sequence of two decisions to merge" (pp.1691) yet this stands at odds with my conditions (and their Proposition 1) for a sequence of two mergers to be supported as an equilibrium outcome. They refer to a "Fat Cat" strategy as $\Delta_2^1 < 0 < \Delta_2^2$ and $\Delta_{-2}^2 > 0$. Yet these same conditions could also lead to a sequence of two decisions *not* to merge as long as Δ_1^1 is sufficiently negative such that $\Delta_1^3 = \Delta_1^1 + \Delta_{-2}^2 < 0$. If $\Delta_2^1 < 0 < \Delta_2^2$ (i.e. $\Phi_I < 0 < \Phi_{II}$ in my Proposition 1.1) then $\Delta_1^3 > 0$ (i.e. $\Phi_{III} > 0$) is the necessary and sufficient condition supporting "a sequence of two decisions to merge".

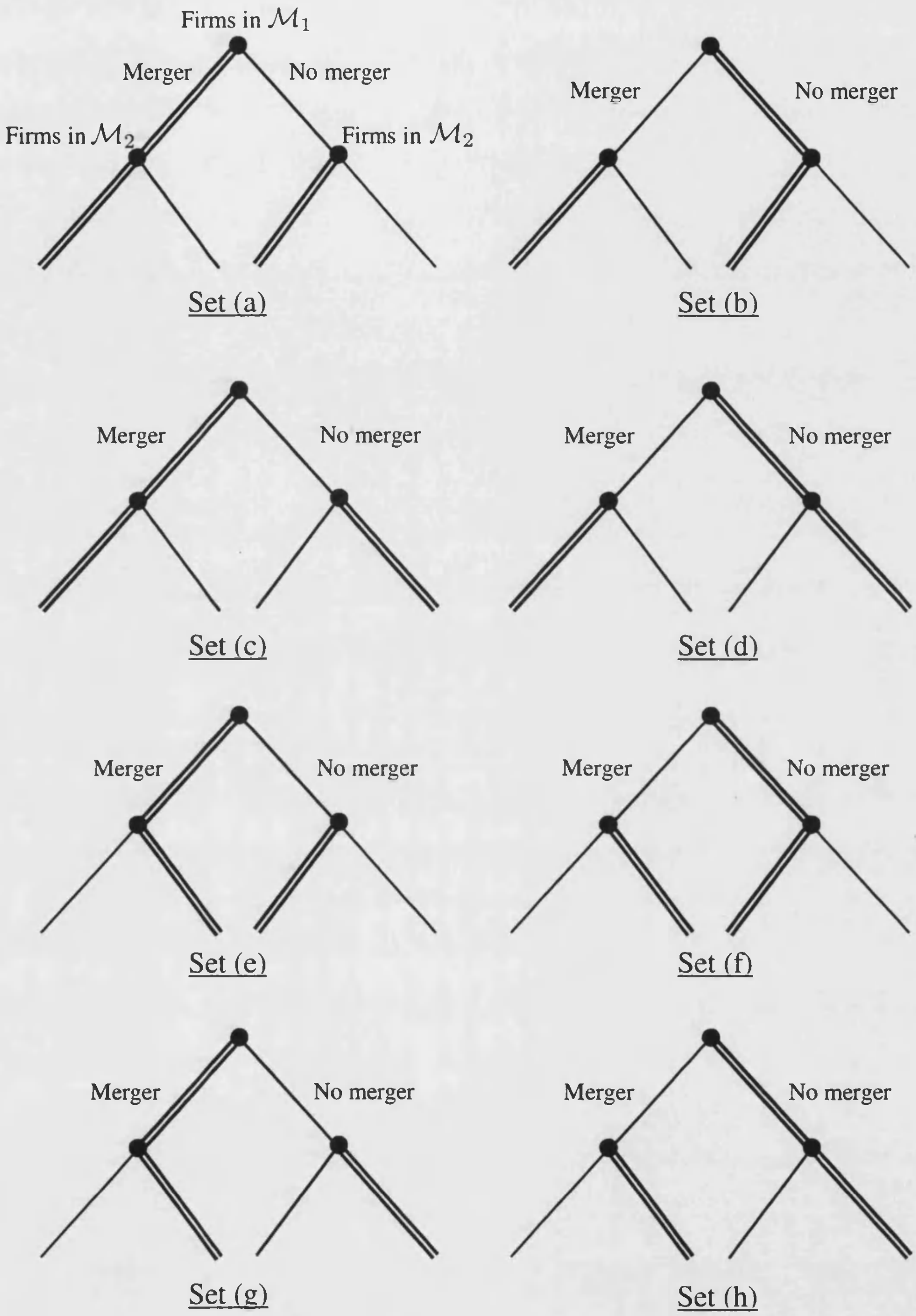


Figure 1-2: Possible equilibrium sets of Nash strategies. From each node, left depicts "merger" and right depicts "no merger"

Set (f) $\Phi_I \geq 0$, $\Phi_{II} < 0$ *and* $\Phi_{IV} < 0$

Set (g) $\Phi_I < 0$, $\Phi_{II} < 0$ *and* $\Phi_V \geq 0$

Set (h) $\Phi_I < 0$, $\Phi_{II} < 0$ *and* $\Phi_V < 0$

Proof. This follows by backward induction. Begin by the conditions for set (a) to be the equilibrium. In the second stage, following a decision by $\mathcal{M}_1$-firms not to merge, $\mathcal{M}_2$-firms will find it profitable to merge if and only if $\Phi_I = \Pi_{M_2,s_1} - \sum_{i\in\mathcal{M}_2} \Pi_{i,s_0} \geq 0$. Conditional on a favourable merger decision by $\mathcal{M}_1$-firms, that $\mathcal{M}_2$-firms find it profitable to merge requires $\Phi_{II} = \Pi_{M_2,s_3} - \sum_{i\in\mathcal{M}_2} \Pi_{i,s_2} \geq 0$. In the first stage, anticipating $\mathcal{M}_2$-firms' subsequent merger regardless of their merger decision, $\mathcal{M}_1$-firms will find it profitable to merge if and only if $\Phi_{VI} = \Pi_{M_1,s_3} - \sum_{i\in\mathcal{M}_1} \Pi_{i,s_1} \geq 0$. Turn to the conditions for (b). As in (a), $\mathcal{M}_2$-firms' strategy is optimal if and only if $\Phi_I \geq 0$ and $\Phi_{II} \geq 0$. Yet for firms in $\mathcal{M}_1$ choosing not to merge to be optimal implies $\Phi_{VI} < 0$. For (c) to be the equilibrium, $\mathcal{M}_2$-firms finding it profitable to merge in the second stage only if $\mathcal{M}_1$-firms choose to merge earlier requires $\Phi_I < 0$ and $\Phi_{II} \geq 0$. In the first stage, anticipating how $\mathcal{M}_2$-firms' merger decision depends on their own decision, $\mathcal{M}_1$-firms will merge if and only if $\Phi_{III} = \Pi_{M_1,s_3} - \sum_{i\in\mathcal{M}_1} \Pi_{i,s_0} \geq 0$. In (d), as in (c), $\mathcal{M}_2$-firms' strategy is optimal if and only if $\Phi_I < 0$ and $\Phi_{II} \geq 0$. Yet firms in $\mathcal{M}_1$ will find it optimal not to merge if and only if $\Phi_{III} < 0$. Proofs of conditions for sets of strategies (e) through (h) to be equilibria follow similarly and are omitted. ■

1.2.2 Special case: Symmetry

When the sets of firms $\mathcal{M}_1$ and $\mathcal{M}_2$ are "symmetric", a common case in models studied in the literature, the general conditions characterising each possible equilibrium can be simplified. What I mean by symmetry here needs to be made precise:

Definition 1.2 *(Symmetry) There is symmetry about the sets of firms* $\mathcal{M}_1$ *and* $\mathcal{M}_2$ *when for all admissible parameter values:*

- $\sum_{i\in\mathcal{M}_1} \Pi_{i,s_1} = \sum_{i\in\mathcal{M}_2} \Pi_{i,s_2}$

- $\Pi_{M_1,s_2} = \Pi_{M_2,s_1}$

- $\Pi_{M_1,s_3} = \Pi_{M_2,s_3}$

- $\sum_{i \in \mathcal{M}_1} \Pi_{i,s_0} = \sum_{i \in \mathcal{M}_2} \Pi_{i,s_0}$

In words, the sum of payoffs of (independent) firms in $\mathcal{M}_1$ under market structure s_1 (where only $\mathcal{M}_2$-firms have merged) equals the sum of payoffs of (independent) firms in $\mathcal{M}_2$ under market structure s_2 (where only $\mathcal{M}_1$-firms have merged). The payoff to (merged) firm M_1 under market structure s_2 equals the payoff to (merged) firm M_2 under market structure s_1. When the two mergers have taken place (market structure s_3), the payoff to (merged) firm M_1 equals the payoff to (merged) firm M_2. When no merger has been carried through (market structure s_0), the sum of payoffs of (independent) firms in $\mathcal{M}_1$ equals the sum of payoffs of (independent) firms in $\mathcal{M}_2$.[13]

Recall the above example, where firms 1, 2 and 3 are located in country A, firms 4 and 5 are located in country B and the sequential merger game specifies firms 1 and 4 first deciding whether to merge (if so forming M_1), followed by firms 2 and 5 deciding whether to merge (if so forming M_2). Symmetry would require that: (i) the sum of payoffs to firms 1 and 4 under market structure $s_1 = \{M_2, 1, 3, 4\}$ equals the sum of payoffs to firms 2 and 5 under market structure $s_2 = \{M_1, 2, 3, 5\}$; (ii) the payoff to M_1 under market structure $s_2 = \{M_1, 2, 3, 5\}$ equals the payoff to M_2 under market structure $s_1 = \{M_2, 1, 3, 4\}$; (iii) under market structure $s_3 = \{M_1, M_2, 3\}$, the payoff to M_1 equals the payoff to M_2; and finally, (iv) under fragmented market structure $s_0 = \{1, 2, 3, 4, 5\}$, the sum of payoffs to firms 1 and 4 equals the sum of payoffs to firms 2 and 5.

By symmetry, it follows from Definition 1.1 that $\Phi_I = \Phi_V$ and $\Phi_{II} = \Phi_{VI}$. Only four merger surplus functions Φ need now be computed. The number of possible sets of strategies by the players that can be supported in equilibrium collapses to

[13]Notice that this notion of symmetry, defined around the *set of firms* $\mathcal{M}_1$ and $\mathcal{M}_2$, is weaker that a concept of symmetry couched directly in terms of the firms themselves, allowing for some asymmetry among firms.

six: sets (b) and (g) are no longer possible[14]. The conditions for each to form an equilibrium are simplified somewhat, as stated in the following proposition (see Figure 1-3):

Proposition 1.2 *(NSC for equilibria in the case of symmetry) In the special case of symmetry, Nash-strategy sets (b) and (g) cannot be supported in equilibrium. Necessary and sufficient conditions for each of the six remaining sets to be supported as the unique equilibrium (for a given combination of parameter values) are:*

Set (a) $\Phi_I \geq 0$ *and* $\Phi_{II} \geq 0$

Set (b) This cannot be supported in equilibrium

Set (c) $\Phi_I < 0$, $\Phi_{II} \geq 0$ *and* $\Phi_{III} \geq 0$

Set (d) $\Phi_I < 0$, $\Phi_{II} \geq 0$ *and* $\Phi_{III} < 0$

Set (e) $\Phi_I \geq 0$, $\Phi_{II} < 0$ *and* $\Phi_{IV} \geq 0$

Set (f) $\Phi_I \geq 0$, $\Phi_{II} < 0$ *and* $\Phi_{IV} < 0$

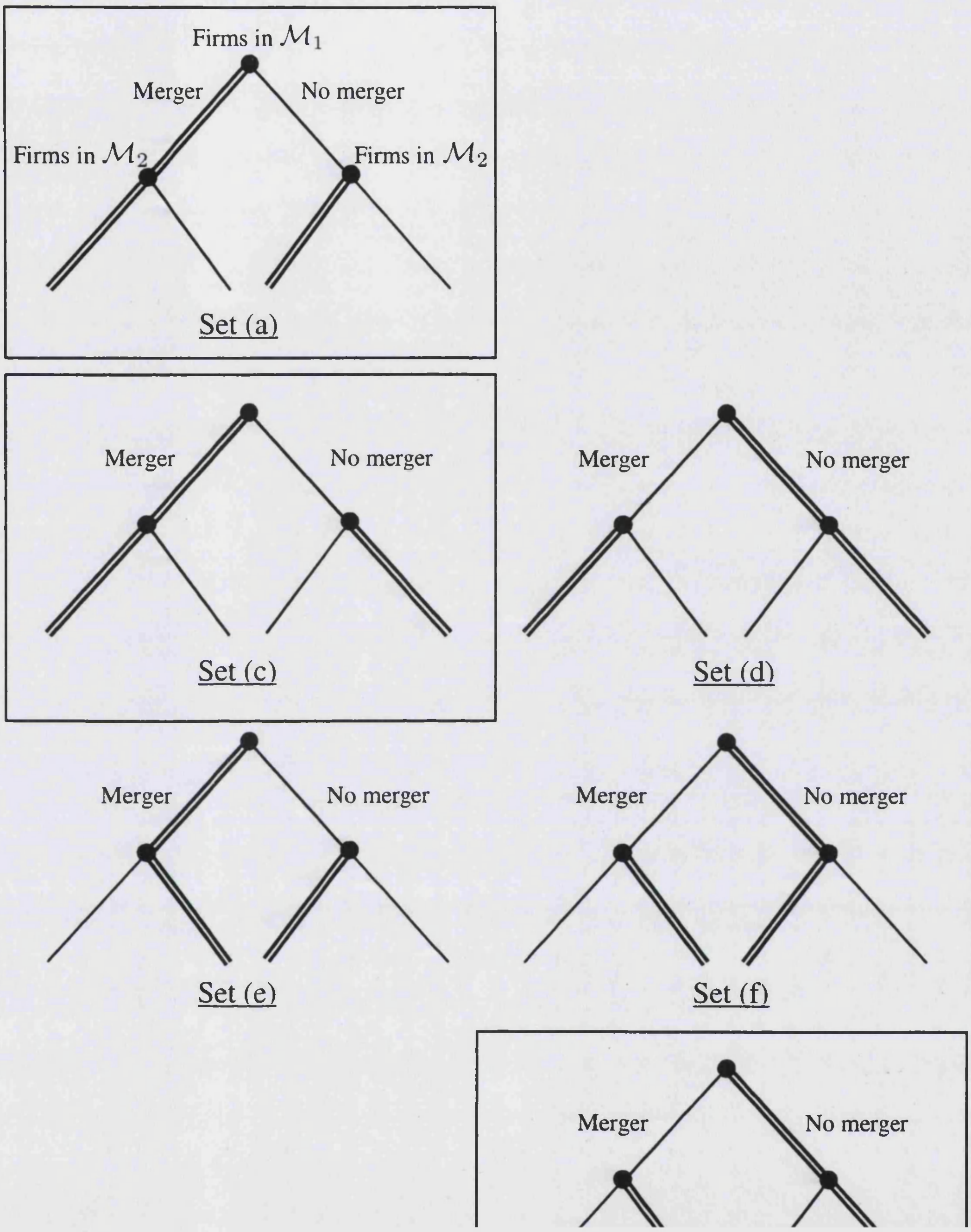
Set (g) This cannot be supported in equilibrium

Set (h) $\Phi_I < 0$ *and* $\Phi_{II} < 0$

Proof. This follows from the definition of symmetry, the definition of the Φ functions and Proposition 1.1. ■

A simple corollary to Proposition 1.2 states a sufficient condition that rules out each of sets (d), (e) and (f) as candidate equilibria. A model that satisfies this sufficient condition for every admissible combination of parameter values can only have sets (a), (c) and (h) among its equilibria. That is, of the eight possible equilibria in Figure 1-2, only the three equilibria highlighted in Figure 1-3 remain in the case of any symmetric model that satisfies the corollary's sufficient condition. This is the case for several models considered in the literature, in addition to the two models considered in this chapter's subsequent application.

[14] This is intuitive. For example, consider set (b). If it is optimal for $\mathcal{M}_2$-firms to merge regardless of $\mathcal{M}_1$-firms' earlier decision, including when $\mathcal{M}_1$-firms decide to merge (this requires $\Phi_{II} \geq 0$), then it cannot be optimal for $\mathcal{M}_1$-firms to decide not to merge, knowing that $\mathcal{M}_2$-firms will subsequently decide to merge regardless of their decision (by symmetry, this would require $\Phi_{VI} = \Phi_{II} < 0$, in contradiction).


Firms in $\mathcal{M}_1$
Merger
No merger
Firms in $\mathcal{M}_2$
Firms in $\mathcal{M}_2$
Set (a)
Merger
No merger
Set (c)
Merger
No merger
Set (d)
Merger
No merger
Set (e)
Merger
No merger
Set (f)
Merger
No merger

Corollary 1.1 *(Sufficient condition to rule out sets (d), (e) and (f) as candidate equilibria, in the case of symmetry) Consider that symmetry holds in the sense defined above. If for all admissible parameter values, it holds that (i) whenever* $\Phi_I \geq 0$ *it happens that* $\Phi_{II} \geq 0$, *and (ii) whenever* $\Phi_{II} \geq 0$ *it happens that* $\Phi_{III} \geq 0$, *then Nash-strategy sets (d), (e) and (f) cannot be supported in equilibrium.*

Formally, this sufficient condition can be stated as

$$\{\theta \mid \Phi_I(\theta) \geq 0\} \subseteq \{\theta \mid \Phi_{II}(\theta) \geq 0\} \subseteq \{\theta \mid \Phi_{III}(\theta) \geq 0\}$$

where θ *is a vector of admissible parameter values (see footnote 8).*

Proof of the corollary follows by inspection of Proposition 1.2[15]. A comment is in order. The corollary states only a *sufficient* condition that rules out sets (d), (e) and (f) as candidate equilibria for symmetric models. As I argue, this condition is satisfied for a fairly wide class of models. (And in the subsequent application I provide intuition as to why this is so.) However, other sufficient conditions can be written that rule out each of sets (d), (e) and (f), as I show in Appendix A.1.

An illustration: Fauli-Oller (2000) I briefly illustrate the sufficient condition of Corollary 1.1 by reference to Fauli-Oller's (2000) (symmetric) model[16]. There are four firms in a homogeneous-good Cournot industry: firms 1 and 2 have zero marginal cost, while firms 3 and 4 have constant marginal cost $c > 0$. Demand is linear, where the price intercept is given by α (and the slope is normalised to -1). As such, the parameters θ of the model are given by α and c, and Fauli-Oller

[15] See Appendix A.1. The sufficient condition of the corollary can be phrased as follows: (i) whenever it is profitable to merge in isolation, it happens that it is profitable to merge conditional on the merger of the rival set of firms; (ii) whenever it is profitable to merge conditional on the merger of the rival set of firms, it happens that each set of firms (i.e. $\mathcal{M}_1$ and $\mathcal{M}_2$) prefers that the two mergers are undertaken over no merger being undertaken.

[16] When there is no geographic dimension to the sequential merger game – as in Nilssen and Sørgard (1998) or in this illustration – the setup of the general analysis just presented requires one minor simplification. Since there is no longer a distinction between firms located in one country or the other, the assumption that at least one firm from each country takes part in each merger decision no longer holds. The two other assumptions remain, that at least two firms take part in each merger decision and that each firm participates in at most one merger decision. The set $\mathcal{A} \cup \mathcal{B}$ can be replaced by $\mathcal{I}$, defined as the initial set of (independent) firms, containing n elements. The necessary and sufficient conditions for equilibria are as in Proposition 1.1 and, in the special case of symmetry, Proposition 1.2.

admits parameters in the range given by $\alpha > 3c$. A slightly simplified version of the sequential merger game considered by the author has firms 1 and 3 first deciding whether to merge (i.e. $\mathcal{M}_1 = \{1,3\}$), followed by firms 2 and 4 who then decide whether to merge (i.e. $\mathcal{M}_2 = \{2,4\}$); a final market competition stage ensues. The merger technology is such that a merged entity produces with zero cost (the cost of its lower-cost constituent). From the reduced-form payoff functions for this standard Cournot model (stated by the author in equation 1), the relevant merger surplus functions of Corollary 1.1 can then be computed: $\Phi_I = \left(\frac{\alpha+c}{4}\right)^2 - \left(\frac{\alpha+2c}{5}\right)^2 - \left(\frac{\alpha-3c}{5}\right)^2$, $\Phi_{II} = \left(\frac{\alpha}{3}\right)^2 - \left(\frac{\alpha+c}{4}\right)^2 - \left(\frac{\alpha-3c}{4}\right)^2$ and $\Phi_{III} = \left(\frac{\alpha}{3}\right)^2 - \left(\frac{\alpha+2c}{5}\right)^2 - \left(\frac{\alpha-3c}{5}\right)^2$. It is easy to show that the sufficient condition of the corollary holds: (i) for all (α, c) such that $\Phi_I \geq 0$ it happens that $\Phi_{II} \geq 0$, and (ii) for all (α, c) such that $\Phi_{II} \geq 0$ it happens that $\Phi_{III} \geq 0$. Consequently, only Nash-strategy sets (a), (c) and (h) *can* obtain in equilibrium as parameter values vary. (It is easy to show from these three merger surplus functions that all three sets indeed *are* supported as unique equilibria at different ranges of parameters.) Two comments are in order. First, notice that the solution to the merger game once the merger surplus functions have been computed, and the sufficient condition has been verified, is immediate. Otherwise the solution can be somewhat more lengthy, as in Fauli-Oller (2000) and Motta and Vasconcelos (2003). This further illustrates the usefulness of the general analysis of this section. Second, notice from Figure 1-3 that the outcomes associated with these equilibria are such that mergers do not occur in isolation but rather bunch together. In other words, whenever a merger happens along the equilibrium path of the game it is undertaken alongside other mergers. I postpone discussion of this result – which hinges on the way non-participating firms react to the merger of rival firms, as is well understood in the literature – to the application.

1.3 Example 1: Cross-border mergers in Sutton's (1991) vertically-differentiated industry

The first example explores the profitability of sequential cross-border mergers in a Cournot oligopoly with vertical product differentiation. I use a partial equilibrium model with differentiated goods due to Sutton (1991) and later applied in a trade context by Motta (1992, 1994).

Before carefully laying out the setup, a brief description is in order. I embed the two countries, a large country and a small country, each with three firms, where firms in the large country offer a product of quality at least as high as that offered by firms in the small country. While product quality varies between the firms located in different countries, it does not (initially) vary across firms located within the same country. The structure of the sequential cross-border merger game in Section 1.2, where two cross-border mergers were allowed to take place in sequence, is extended here to three merger stages: in each stage a large-country high-quality firm and a small-country low-quality firm decide whether to merge or not. If the merger takes place, the level of quality offered by the merged (and multinational) firm is the higher of the qualities offered by its constituent firms (i.e. the quality offered previously by the large-country constituent). Thus cross-border mergers enable firms to transfer technology. Consistent with the type of sequential cross-border merger game spelled out in Section 1.2, in a final stage, firms which have merged or remained independent engage in Cournot competition on the world market.

Demand setup Consider two countries, $l \in \{A, B\}$, with m_l consumers each. Consumers in both countries have identical Cobb-Douglas preferences defined over a quality (differentiated) good and an outside good, indexed by quantities x and y respectively, where u denotes the quality level of the quality good:

$$U = (ux)^{\beta} y^{1-\beta} \qquad 0 < \beta < 1$$

Total expenditure on the quality good in economy l, S_l, is:

$$S_l = \sum_{i=1}^{n_l} p_i x_i = \sum_{k=1}^{m_l} \beta z_k \tag{1.1}$$

where n_l is the number of firms selling in country l ($n_l \geq 2$), x_i and p_i are respectively the quantity and price of the (single) quality good sold by firm i (referred to as variety i), and z_k is the income of consumer k. Assume $S_A \geq S_B$ due either to a larger population or a higher per capita income in (large) country A relative to (small) country B.

Given any vector of qualities and associated prices, the consumer chooses a variety i that maximises the quality/price ratio $\frac{u_i}{p_i}$[17]. All varieties that command positive sales at equilibrium must therefore have prices proportional to qualities:

$$\frac{p_i}{u_i} = \lambda \quad \forall i$$

where λ is a constant. From equation (1.1), we can then write $S = \sum_{i=1}^{n} p_i x_i = \lambda \sum_{i=1}^{n} u_i x_i$ (momentarily dropping the country subscript l for simplicity), expressing the price-to-quality ratio λ as:

$$\lambda = \frac{S}{\sum_{i=1}^{n} u_i x_i} \tag{1.2}$$

The inverse demand function for variety i is also obtained:

$$p_i = \frac{S}{\sum_{j=1}^{n} \frac{p_j}{p_i} x_j} = \frac{S}{\sum_{j=1}^{n} \frac{u_j}{u_i} x_j} \tag{1.3}$$

By assumption each variety i is sold to the nth part of the population.

Supply and historical motivation for entry and quality Firms are assumed to have the same constant marginal cost of production $c > 0$. Goods produced in one country but sold in another (exports) incur an additional unit trade cost $t \geq 0$.

[17] Consumer k chooses to consume quantity x_{ik} of variety i such that $u_i x_{ik}$ is maximised subject to his budget constraint $p_i x_{ik} = \beta z_k$; i.e. he solves $\max_i \beta z_k \frac{u_i}{p_i}$ by selecting a variety i such that $\frac{u_i}{p_i}$ is maximised across i.

As mentioned above, each country is initially embedded with three (independent) firms, where a firm from country l offers a good of quality $u_l \geq 1$. By construction, $u_A \geq u_B$ and define the quality ratio (or gap) v as the ratio of the quality offered by large-country firms to that offered by their small-country counterparts, $v := \frac{u_A}{u_B}(\geq 1)$.

The number of firms in each country and their respective (asymmetric) qualities can be motivated by considering the following long-term entry and investment game (Motta 1992). Countries are initially closed to foreign trade and investment and in this autarkic setting, without forseeing any changes, firms make entry and investment decisions. In each country, the following two-stage game is played. In a first stage, firms simultaneously decide whether to enter and, if so, with which quality. In a second stage, they engage in Cournot competition. In order to offer a good of quality u, firms must incur a fixed and sunk cost $F(u) = u^{\gamma}$, $u \geq 1$.

Consistent with models of this type, in equilibrium the same number of firms enters in each country, with the large-country firms making larger investments in quality than the small-country firms, $u_A \geq u_B$. The convexity of the fixed cost function is chosen so that three firms enter in each country[18].

By this historical motivation, market integration then unexpectedly occurs. Firms are "locked in" with their previous (autarky-based) quality levels, reflecting the long-term nature of their investment decisions (capability-building) as opposed to the short-run process of market integration and market competition. Firms located in each country are now allowed to export to the other market (due to trade integration) and/or can merge with a (now) rival firm located in the other market (due to investment integration), as set out below.

[18] See Appendix A.2 for a derivation. The number of firms (in each country), $n = 3$, depends on the choice of parameter, $\gamma = 3$, of the fixed cost function, $F(u) = u^{\gamma}$. By making the cost function more convex, i.e. making (R&D, say) spending less effective in raising quality, the larger is the number of firms entering in equilibrium. For example, for $\gamma = 5$, $n = 4$ firms enter in equilibrium. We may write $n = n(\gamma)$, where $n' > 0$. Notice that n does not depend on market size: this "non-convergence" result is consistent with the finiteness property of many vertical product differentiation models (Shaked and Sutton 1983, Sutton 1991). For $\gamma = 3$, the three firms entering in country l offer a common quality given by $u_l = \frac{2}{3}\sqrt[3]{\frac{S_l}{3}}$. It follows from the assumption that $S_A \geq S_B$ that the quality offering of a large-country firm exceeds that of a small-country firm: $u_A \geq u_B$.

The experience of some Latin American and Southeast Asian countries, which during the 1990s underwent a relatively unplanned process of trade and investment liberalisation, with the ensuing high volume of cross-border mergers and acquisitions, comes to mind. It is in this historical setting that the cross-border merger game is set.

Sequential cross-border merger game In sequence, each of the three firms located in country A – producing with quality u_A – is allowed to merge with each of the three firms located in country B – producing with quality u_B. With no loss of generality, the first A-country firm and the first B-country firm which decide whether to merge (stage 1) are labelled 1 and 4 respectively, the second A-country firm and the second B-country firm which decide whether to merge (stage 2) are labelled 2 and 5 respectively, and the third A-country firm and the third B-country firm which decide whether to merge (stage 3) are labelled 3 and 6 respectively. Keeping with the notation of Section 1.2, $\mathcal{M}_1 = \{1, 4\}$, $\mathcal{M}_2 = \{2, 5\}$ and $\mathcal{M}_3 = \{3, 6\}$ and the merged (multinational) firms which may come to form are labelled M_1, M_2 and M_3. As mentioned above, the merger-technology assumption is that full transfer of technology takes place, the merged firm being able to produce at the high level of quality u_A not only at its A-country facilities but also at its B-country facilities[19]. In the fourth and final stage, merged or independent firms compete à la Cournot on the world market.

The sequence of merger decisions is depicted in Figure 1-4, as are the possible market structures coming out of the merger stages. The market structure where no cross-border merger occurs, $\{1, 2, 3, 4, 5, 6\}$, is denoted r_0, while that where three cross-border mergers are carried through, $\{M_1, M_2, M_3\}$, is labelled r_3. Given that symmetry holds (in the sense defined in Section 1.2), market structures where one

[19]This full-transfer-of-technology assumption is also made in Fauli-Oller (2000), for example. It is one extreme case of a more general assumption where, by merging with a high-quality foreign firm, a B-country firm can produce at an "average" level of quality given by, for example, the convex combination $\delta u_A + (1 - \delta) u_B$, where $0 \leq \delta \leq 1$ is the technology transfer coefficient. While a low-quality firm located in a developing country will not necessarily achieve "world class" quality by merging with a high-quality developed-country counterpart (indeed one can conceive of world class being easier to achieve through the setting up of a "greenfield" operation, rather than acquiring existing assets), this assumption allows us to examine sequential mergers in the context of international trade and quality asymmetries in a simple way.

cross-border merger is undertaken are denoted r_1 and those where two cross-border mergers occur are denoted r_2. For example, r_2 comprises structures $\{M_1, M_2, 3, 6\}$, $\{M_1, M_3, 2, 5\}$ and $\{M_2, M_3, 1, 4\}$, where there are two multinational firms and one independent firm producing in each country[20].

1.3.1 Equilibria as a function of the quality gap and the trade cost

The reduced-form profits for each firm, as a function of the parameters in the model, for every possible market structure are derived in Appendix A.2. The notation introduced in Section 1.2 needs to be complemented: let subscript a denote an independent A-country firm, subscript b denote an independent B-country firm and subscript m denote a merged (multinational) firm. Thus Π_{a,r_2}, for instance, denotes the payoff to an independent firm located in country A under market structure r_2. Reduced-form profits are a function of the quality gap v, the trade cost normalised by the marginal cost of production $\tilde{t} := \frac{t}{c}$ and the market sizes S_A and S_B.

I assume that $(1 \leq) v \leq \frac{3}{2}$ (the quality ratio is low enough) and $(0 \leq) \tilde{t} \leq \frac{3-2v}{2v}$ (the trade cost is low enough) so that in equilibrium low-quality firms command positive sales in both countries, i.e. there is two-way (intra-industry) trade between countries under any market structure where at least two firms remain independent (clearly trade ceases if all firms merge across borders, in situation r_3). This space of parameter values is labelled $\mathcal{P}$, as before. (Such an assumption is made for simplicity of exposition; extending the space of parameter values outside $\mathcal{P}$, i.e. to include $\{(v, \tilde{t}) \in \mathcal{R}^2 \mid v \geq 1 \text{ and } \tilde{t} \geq 0\} \setminus \mathcal{P}$, does not add to the results, as discussed in Appendix A.2.)

Given that there are three merger stages, merger surplus functions are now denoted Ψ (cf Φ in Section 1.2). Coupled with the fact that symmetry holds, eight

[20]For the sake of simplicity, now that there are three merger stages, the r-notation here takes into account that symmetry holds (cf the s-notation in Section 1.2, introduced for the general case, as in Nilssen and Sørgard 1998).

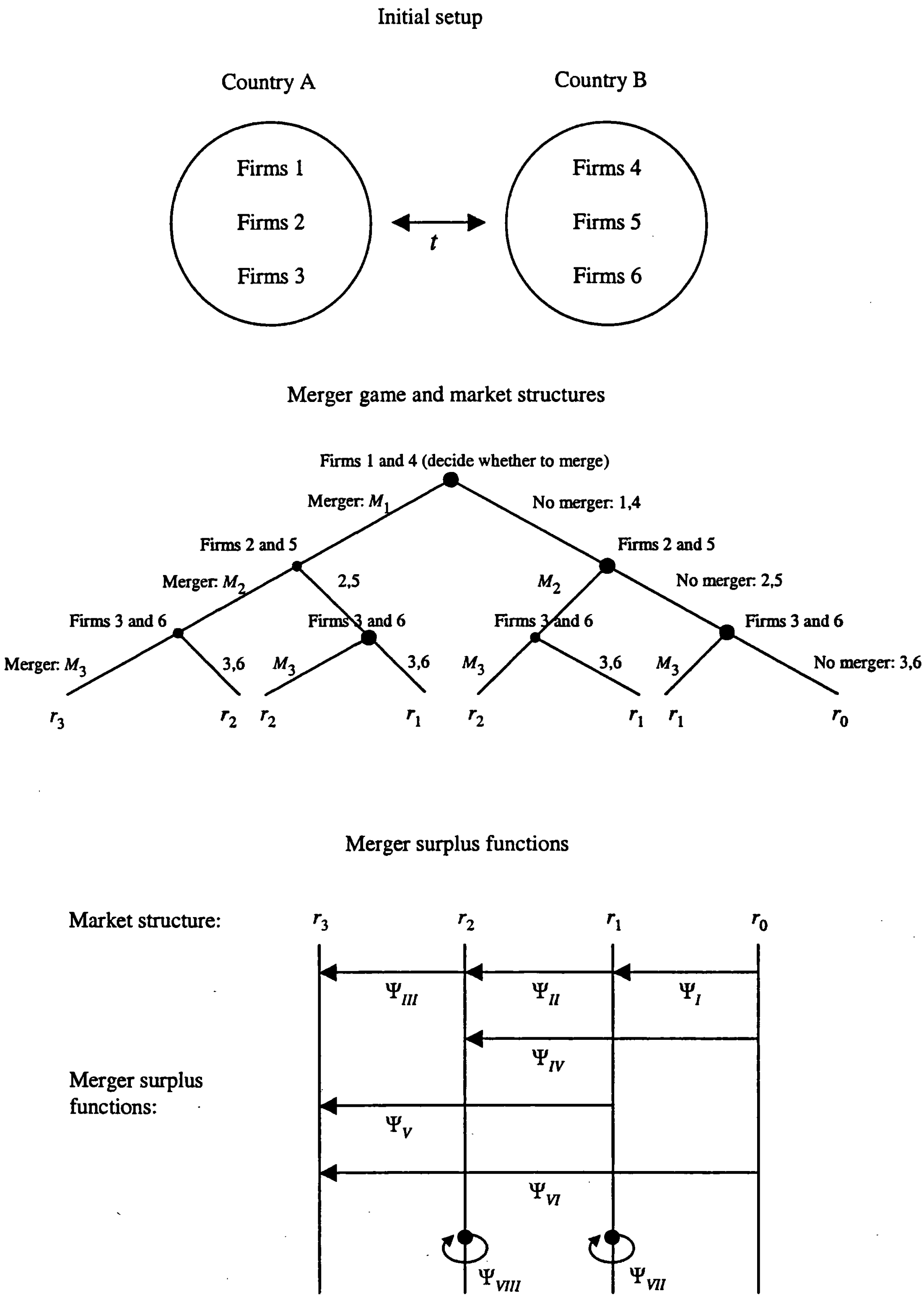


Figure 1-4: (Example 1) Sequential cross-border merger game and merger surplus functions. (The notation is adapted to reflect three merger stages and considers the symmetry of the model.)

merger surplus functions are relevant[21], defined as follows (see Figure 1-4 for an illustration):

Definition 1.3 *(Merger surplus functions) Define eight Ψ functions as follows:*

$\Psi_I := \Pi_{m,r_1} - \Pi_{a,r_0} - \Pi_{b,r_0}$ $\qquad$ $\Psi_V := \Pi_{m,r_3} - \Pi_{a,r_1} - \Pi_{b,r_1}$

$\Psi_{II} := \Pi_{m,r_2} - \Pi_{a,r_1} - \Pi_{b,r_1}$ $\qquad$ $\Psi_{VI} := \Pi_{m,r_3} - \Pi_{a,r_0} - \Pi_{b,r_0}$

$\Psi_{III} := \Pi_{m,r_3} - \Pi_{a,r_2} - \Pi_{b,r_2}$ $\qquad$ $\Psi_{VII} := \Pi_{m,r_1} - \Pi_{a,r_1} - \Pi_{b,r_1}$

$\Psi_{IV} := \Pi_{m,r_2} - \Pi_{a,r_0} - \Pi_{b,r_0}$ $\qquad$ $\Psi_{VIII} := \Pi_{m,r_2} - \Pi_{a,r_2} - \Pi_{b,r_2}$

The sufficient condition of Corollary 1.1 holds Along the lines of Proposition 1.2 and Figure 1-3, how the signs of these merger surplus functions change as parameter values vary pins down the equilibria of the merger game. One can show that in general there are twenty possible equilibria in a symmetric three-stage game, and conditions supporting each can be written. However, the verification that the model satisfies the sufficient condition of Corollary 1.1 simplifies matters considerably: the set of candidate equilibria is reduced from twenty to four elements. This is captured in the following lemma, which is the natural extension of the sufficient condition of Corollary 1.1 to a game with three merger stages.

Lemma 1.1 *For $(v, \tilde{t}) \in \mathcal{P}$:*

(i) $\Psi_I(v,\tilde{t}) \geq 0 \Longrightarrow \Psi_{II}(v,\tilde{t}) \geq 0 \Longrightarrow \Psi_{III}(v,\tilde{t}) \geq 0$

(ii) $\Psi_{II}(v,\tilde{t}) \geq 0 \Longrightarrow \Psi_{IV}(v,\tilde{t}) \geq 0$

$\Psi_{III}(v,\tilde{t}) \geq 0 \Longrightarrow \Psi_V(v,\tilde{t}) \geq 0 \Longrightarrow \Psi_{VI}(v,\tilde{t}) \geq 0$

While proof that the lemma holds is provided in Appendix A.2, the intuition is straightforward. Consider part (i). A merged firm produces less than the pre-merger sum of outputs of its constituents since each constituent now internalises

[21] While in Section 1.2, with two merger stages and symmetry the number of relevant merger surplus functions was four, here it is eight. In general, if there are T merger stages, $T \geq 1$, and there is symmetry, the number of relevant merger surplus functions is given by $T - 1 + \sum_{i=1}^{T} i$. If there are T merger stages but symmetry does not hold, then the number of relevant merger surplus functions increases to $\sum_{i=1}^{2^T-1} i$.

Parameters v, $\tilde{t}$, S_A and S_B are suppressed as arguments of the functions here for simplicity. When I do include arguments, subsequently, I ignore S_A and S_B since Proposition 1.3 holds irrespective of market sizes (these enter the profit functions multiplicatively).

the externality it confers upon the other constituent when making its output decision[22]. Since quantities are strategic substitutes, firms not participating in the merger ("outsiders") respond by increasing output, which may render a merger unprofitable for its participants ("insiders"). (Such "accommodation" was highlighted by Salant, Switzer and Reynolds 1983 in a Cournot model with symmetric firms.) Now, whenever firms in $\mathcal{M}_i$ do not find it profitable to merge when two other cross-border mergers are taking place (i.e. $\Psi_{III} < 0$) and consequently there are less outsiders to free ride on the merger, then they would not find it profitable to merge were only one other merger (i.e. $\Psi_{II} < 0$), let alone none (i.e. $\Psi_I < 0$), taking place, as there would be more free-riders to expand output following $\mathcal{M}_i$-firms' merger[23]. As for part (ii) of the lemma, consider $\Psi_{II} \geq 0 \Longrightarrow \Psi_{IV} \geq 0$, for example. The decision to merge captured by merger surplus function Ψ_{II} entails an opportunity cost given by the sum of the profits of an independent A-country firm and an independent B-country firm under market structure r_1. This opportunity cost is larger than the opportunity cost in the decision to merge captured by Ψ_{IV}, which amounts to the profits of the same two independent firms but under r_0, a more "fragmented" market structure (in the sense that while in r_1 there are five firms selling into each country, in r_0 there are six firms)[24].

The set of remaining candidate equilibria, containing four elements, is illustrated in Figure 1-5, alongside the conditions on the reduced-form functions in the right-hand margin. The equilibria are labelled (a), (c1), (c2) and (h), in a manner consistent with the labelling of equilibria in Section 1.2. In view of Lemma 1.1, only three merger surplus functions need be computed in parameter space $\mathcal{P}$: Ψ_I,

[22]In this model, for $(v, \tilde{t}) \in \mathcal{P}$, this occurs despite the quality jump from u_B to u_A enjoyed by the B-country constituent. More generally, note that if a quality increase (or, isomorphously, a marginal cost reduction) through merger were very large, insiders to the merger could actually *expand* output. A merger, by lowering prices, would then be detrimental to outsiders, who would respond by lowering output (under quantity competition).

[23]Fauli-Oller (2000) interprets a similar result as "previous takeovers stimulate takeover profitability" (pp.197).

[24]In this model, in the space $\mathcal{P}$ of parameter values, it can be shown that despite merging parties having the advantage of producing with high quality in both countries and no longer cross-hauling product between them, non-participants always gain in the event of a merger, be they multinational firms or independent firms. Part (ii) of the lemma follows from noting this. Outside $\mathcal{P}$, for example, for $\tilde{t}$ high enough (and $v > 1$) that not even high-quality A-country firms command positive sales in country B (i.e. trade flows in neither direction), a merger has a detrimental effect on outside (independent B-country) firms.

Ψ_{II} and Ψ_{III}. By inspecting the signs of these functions in $\mathcal{P}$, the game can be solved.

The solution to the game is stated in the next proposition. It states that all four remaining candidate equilibria occur in space $\mathcal{P}$. Figure 1-6 partitions the space into four disjoint zones, each labelled after the equilibrium which can be supported for the combinations of parameters defining the zone. Thus, for low enough values of both the quality gap v and the (normalised) trade cost $\tilde{t}$, in zone (h), the "unprofitability of mergers" result common to models with quantity competition and in the absence of merger "synergies" (Salant, Switzer and Reynolds, 1983) is obtained. Here, along the equilibrium path, as well as in subgames hanging from nodes off the equilibrium path, no cross-border merger occurs. At the other end, for high enough values of v and/or $\tilde{t}$, in zone (a), mergers are always profitable: along the equilibrium path, as well as off it, cross-border mergers always occur. For intermediate values of v and/or $\tilde{t}$, in zones (c1) and (c2), mergers again occur along the equilibrium path. Here, however, the merger decisions in the early stages of the merger game have a bearing on the merger decisions in the later stages.

Proposition 1.3 *Space $\mathcal{P}$ can be partitioned into four disjoint zones, as depicted in Figure 1-6. A straight line segment drawn from $(v, \tilde{t}) = (1, 0)$ to any point on $\tilde{t} = \frac{3-2v}{2v}$ begins in zone (h), crossing into zone (c2), followed by zone (c1) and ending in zone (a). In each zone, the unique equilibrium is given in Figure 1-5, labelled after the zone. The boundaries between zones (h) and (c2), between zones (c2) and (c1), and between zones (c1) and (a), are continuous and downward-sloping, joining a point $(v, 0)$, $1 < v < \frac{3}{2}$, on the v-axis to a point $(1, \tilde{t})$, $0 < \tilde{t} < \frac{1}{2}$, on the $\tilde{t}$-axis. The boundary between zones (h) and (c2) lies strictly below the boundary between zones (c2) and (c1), which in turn lies strictly below the boundary between zones (c1) and (a).*

Intuition and decomposition analysis (A decomposition analysis of the profitability of a merger provides intuition on the results of Proposition 1.3. The reader

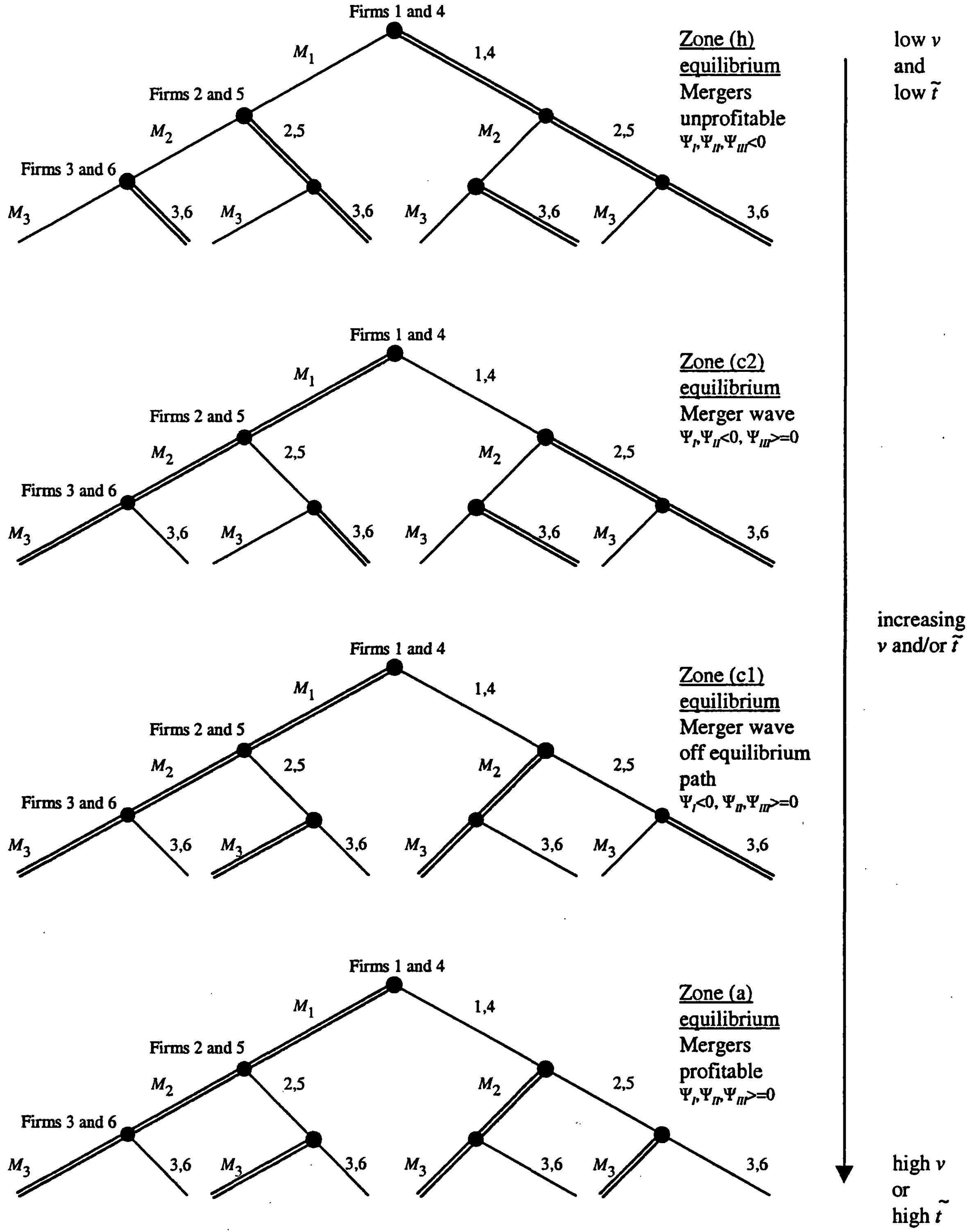


Figure 1-5: (Example 1) Equilibrium of sequential merger game in each zone; v quality gap, $\tilde{t}$ (normalised) trade cost

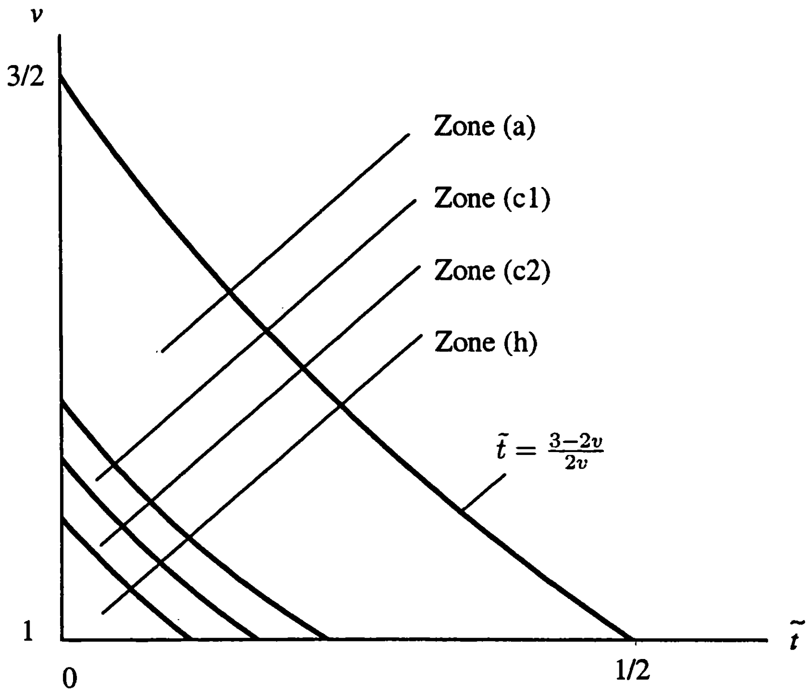


Figure 1-6: (Example 1) Zones in $\mathcal{P}$ space

may wish to skip to Section 1.3.2 without loss of continuity. Proof of the proposition is set out in Appendix A.2.) How the total effect of a single cross-border merger (say, a stand-alone first merger that shifts market structure from r_0 to r_1[25]) on the profit of insiders changes as v and $\tilde{t}$ vary can be broken down in a simple way into three effects. The first effect, labelled the "insider output contraction" effect, is calculated assuming a merger technology where there is *no* quality upgrade in the production facilities located in country B as a result of the merger, and further that outsiders do *not* react to the merger by increasing output, but rather maintain their pre-merger output levels. This profit effect is unambiguously positive[26] for

[25]By "stand-alone" I am considering the profitability of a first merger in isolation, abstracting from whether it is followed by further mergers and therefore from whether r_1 is an equilibrium market structure.

[26]Insider profits rise on sales in both countries. Under this modified merger technology, it can be shown that when $v > 1 + \tilde{t}$ the merged firm discontinues production operations in country B, shipping high-quality product produced in country A to country B. When $v < 1 + \tilde{t}$, the merged

$(v, \tilde{t}) \in \mathcal{P}$. A merger internalises the externality each constituent insider confers upon the other when making an output decision (the insiders' reaction function shifts inward); as a result, the (quality-adjusted[27]) combined output of insiders falls, prices of high and low-quality goods rise in both countries and insider profits rise.

The second profit effect, labelled the "technology transfer" effect, still assumes that outsiders' output remains unchanged upon merger, but now the merger technology assumed in the model, that quality of B-country production facilities jumps from u_B to u_A through merger, is re-established. This effect is also unambiguously positive: insider profits on sales in country B rise as a result of the transfer of technology[28]. By this effect, insiders expand the (quality-adjusted) quantity sold in country B (with prices in country B falling), yet this increase in sales in country B does not offset the output contraction in country B as a result of the first effect (the net effect on country B prices is positive).

The positive "insider output contraction" and "technology transfer" effects on the profit of insiders are countervailed by a negative "outsider output expansion" effect. This third effect captures the impact on the profit of insiders of allowing outsiders to respond to the merger. Since quantities are strategic substitutes, outsiders to a merger free ride on insiders' output contraction (the equilibrium moves out along the outsiders' reaction function as the insiders' reaction function shifts inward). As outsiders' (quality-adjusted) output increases, insiders react by further contracting output, to which outsiders further expand output, and so on, until equilibrium is reached. This third effect on insiders' profit is unambiguously

firm no longer cross-ships, producing low-quality product in country B for domestic consumption.

[27]Quality-adjusted output is defined as output multiplied by the quality gap $v = \frac{u_A}{u_B}$ for product of high quality u_A *when* in competition with product of low quality u_B, and output multiplied by 1 otherwise. This measure of output is natural in view of the inverse demand functions (e.g. see equation (1.3)).

[28]Clearly, there is no effect on sales, prices and profit on sales in country A. As in the model, the merged firm now produces high-quality product in country B for domestic consumption. This effect could alternatively be called the "(insider) trade cost elimination" effect. The upgrade in quality of B-country facilities has an equivalent effect to the elimination of trade costs on cross-hauling by insiders: by eliminating trade costs insiders would produce only high-quality product in country A, which is equivalent to producing high-quality product in both countries for domestic consumption when trade costs are present. Indeed this is an example of the symmetry of the model with respect to v and $\tilde{t}$.

negative.

The overall impact of the cross-border merger on the profit of insiders turns on the relative magnitude of the countervailing effects. Intuitively, the "outsider output expansion" effect is increasingly negative (outsiders' output *expansion* increasingly high) as v and/or $\tilde{t}$ falls: for $(v, \tilde{t})$ sufficiently close to $(1, 0)$, this negative effect on the profit of insiders dominates the two positive effects and the stand-alone merger is unprofitable. To see this, note that at $(v, \tilde{t}) = (1, 0)$, with no quality asymmetries and no trade costs, with six firms competing in each country on an equal footing[29], output expansion by outsiders is at its most severe, rendering a merger at its most unprofitable for insiders.

The different equilibrium outcomes of the sequential game as the parameter take on different values can then be understood by reference to these effects. In zone (h), where v and $\tilde{t}$ are both low, the output expansion by outsiders in response to a merger would be significant, while the advantage enjoyed by the multinational of producing with high quality in both countries and no longer cross-hauling product between them would be limited. Being an outsider in the event of a merger is highly profitable. As a result, no cross-border merger occurs along the equilibrium path.

At the other end, for high enough values of v and/or $\tilde{t}$, in zone (a), mergers are always profitable: along the equilibrium path, as well as off it, cross-border mergers always occur. Under high v and/or $\tilde{t}$, the output expansion by outsiders to a merger is limited. To see why this is so for sufficiently high v even when $\tilde{t}$ is close to zero, note that starting from a market structure with no cross-border merger, there are effectively only three (high-quality, A-country) firms selling in each country. If only one cross-border merger is now allowed to occur, the number of firms effectively selling in each country remains unchanged, hence the limited output expansion by outsiders. To see why there is limited output expansion by outsiders for sufficiently high $\tilde{t}$ even when v is close to 1, note that again a cross-border merger leaves the number of firms effectively selling in each market unchanged (three A-country firms

[29]Recall that I am analysing the profitability of a stand-alone *first* cross-border merger for the purpose of motivating intuition.

selling in country A and three B-country firms selling in country B). When v and $\tilde{t}$ are greater than 1 and 0 respectively, the technology transfer effect is present: there is the added gain from merging arising from upgrading the production facilities located in country B to the higher level of quality.

For intermediate values of v and/or $\tilde{t}$, mergers again occur along the equilibrium path. Here, however, early mergers induce subsequent mergers. In zone (c2), along the equilibrium path, firms 1 and 4's merger is decisive for firms 2 and 5 to merge, and both mergers are in turn decisive for firms 3 and 6 to merge. Stated loosely, more mergers mean less free riding by outsiders in the market competition stage. Notice that if, say, firms 1 and 4 were not to anticipate the effect their merger has on subsequent merger decisions, but viewed their merger in isolation, they would not merge since $\Psi_I < 0$ in zone (c2)[30]. In zone (c1), earlier mergers are still decisive for later mergers. However, due to higher v and/or $\tilde{t}$ compared to zone (c2), and hence lower output expansion by outsiders, only one previous decision to merge suffices for firms 3 and 6 to find it profitable to choose likewise.

1.3.2 Discussion: Foreign investment literature and robustness

In this subsection results are discussed in light of the international trade and investment literature. Considerations are also made on the specification of the extensive form of the merger game and the number of merger stages.

The trade literature has modelled foreign direct investment (FDI) mostly in the form of *greenfield* investment, where firms choose between exporting to a market and setting up production facilities in that market; see for example Motta (1994), who studies greenfield investment in a similar vertically-differentiated global oligopoly. The literature has emphasised the tariff-jumping (more precisely, trade-cost-jumping) rationale behind FDI: unless trade costs are sufficiently low, runs the

[30]In an international oligopoly in the presence of technological innovation, Graham (1985) suggests that a firm's motivation for foreign direct investment may be "to respond to or anticipate the action of a rival... creating a clustering of DFI (foreign direct investment) in the industry" (pp. 69; parenthesis added).

argument, it may be rational for a firm to set up foreign production operations. On the one hand, greenfield investment requires a fixed setup cost but, on the other hand, it enables firms to reduce variable (trade) cost.

Yet empirically, the major channel by which foreign investment is made is that of mergers and acquisitions (UNCTAD 2000). By analysing investment in this mode, in counterpoint to greenfield investment, my model points out that cross-border mergers allow firms not only to save on trade costs and transfer technology but also, by reducing the number of rivals, to enjoy muted competition. It explores a different mechanism to the tariff-jumping rationale of why foreign investment may not occur when trade costs (and the quality gap) are low: one where the profitability of investment turns on the response of non-participating firms in a Cournot oligopoly. In such situations where there is low product differentiation along both horizontal ($\tilde{t}$) and vertical (v) dimensions, any cross-border merger would be met with a fierce output response by those producers not party to the merger.

An example is in order. Analysing the substitution-of-exports-for-greenfield-investment decision, Motta (1994) finds that a decrease in the value of the exporting cost, ceteris paribus, results "in an enlargement of the regions (in parameter space) where exports, rather than investments, prevail" (pp.191). This is supported in my model of acquisition investment (in Figure 1-6 the boundary between zones (h) and (c2) is downward-sloping), yet owing to a different mechanism. Motta's result derives from tariff-jumping: the lower is the trade cost, ceteris paribus, the lower are the gains from (greenfield) investment. In my model, the lower is the trade cost, ceteris paribus, the greater is the intensity of competition and the *higher* are the gains to the *industry* from (acquisition) investment (see below). Yet it is precisely when the trade cost is low that free riding on other mergers is so profitable that in equilibrium no (acquisition) investment occurs.

Clearly, for parameter values in all zones including zone (h), the "grand cross-border coalition" – where three cross-border mergers occur and there are no independent firms – is the structure under which industry profits are maximal[31].

[31]When $(v, \tilde{t}) = (\frac{3}{2}, 0)$ or $(1, \frac{1}{2})$, the grand cross-border coalition is weakly dominant. Otherwise

In fact, it is in zone (h), where both v and $\tilde{t}$ are low enough, that the industry gains from merging[32] are greatest: intuitively, the lower is v or the lower is $\tilde{t}$, the greater is the intensity of competition resulting in lower industry profits under r_0, and therefore the higher are the gains to be reaped from eliminating competitors through merger. Despite this, in zone (h) free riding on a merger is "far too" profitable and intense. Due to this prisoner's dilemma character of the game in this zone, the three-merger outcome cannot be supported in equilibrium.

Modelling foreign investment in the form of cross-border mergers suggests just how widespread such activity may be. While it is difficult to rationalise greenfield investment when there are no trade costs ($\tilde{t} = 0$), since location of production does not matter, cross-border mergers will occur when the quality gap is *high enough:* there exist $(v, \tilde{t}) \in \mathcal{P}$ in zones (a), (c1) and (c2) when $\tilde{t} = 0$ and v exceeds certain thresholds. Thus the existence of a (large enough) quality gap is a sufficient condition for mergers to occur, unlike the case of a simple greenfield investment model, where $\tilde{t} > 0$ is also necessary. As with mergers in general, via cross-border mergers firms can dampen competition. My model further predicts the phenomenon of investment-bunching. When cross-border mergers occur they do not happen in isolation: (Merger, Merger, Merger)[33] is the unique equilibrium outcome for most parameter values with the exception of low values for both the quality gap and the trade cost.

It is interesting to compare the equilibrium outcomes of the sequential game with those of a different game, where the three merger decisions are made *simultaneously* in a first stage followed by a market competition stage, as before. As seen above, in the sequential-move game, (Merger, Merger, Merger) is the unique subgame perfect Nash equilibrium (SPNE) outcome for parameter values in zones (a), (c1) and (c2), while (No Merger, No Merger, No Merger) is the unique SPNE outcome in zone (h). Consider now the simultaneous-move game. In each zone

it is strictly dominant for $(v, \tilde{t}) \in \mathcal{P}$.

[32]Note that the industry gains from merging is the difference between profits under market structure r_3 and profits under r_0, i.e. $3\Psi_{VI}(v, \tilde{t})$. That these are monotonically decreasing in $\tilde{t}$, for a given v, and monotonically decreasing in v, for a given $\tilde{t}$, is intuitive and can be shown.

[33]By this representation I mean that all three cross-border mergers are carried out: firms 1 and 4 merge, firms 2 and 5 merge and firms 3 and 6 merge.

the same outcome as that in the sequential game can be supported in equilibrium. Yet in zones (c1) and (c2), this equilibrium outcome is no longer unique: in addition to (Merger, Merger, Merger), (No Merger, No Merger, No Merger) is also a pure-strategy Nash equilibrium outcome.

The set of equilibrium outcomes in the simultaneous-move game is thus a (strict, in zones (c1) and (c2)) superset of the set of equilibrium outcomes in the sequential game. By modelling the merger decisions sequentially, the no-merger outcome which could be supported as a second equilibrium outcome in zones (c1) and (c2) were the merger decisions to be modelled simultaneously, no longer survives. The sequential nature of moves acts as a coordination device, and may help explain the empirical observation of widespread, and bunched together in time, cross-border mergers[34].

Some final remarks pertaining to the number of merger stages. One may ask how robust is the result of widespread cross-border mergers in equilibrium to the number of merger stages (for a game with T merger stages, modifying the setup by, say, embedding each country initially with T independent firms). Consider first the game with two merger stages (and two firms in each country), i.e. $T = 2$. Similar results to those of Figure 1-5 (where $T = 3$) obtain, with one exception. A zone like (h) in Figure 1-5, where mergers are unprofitable and do not occur along the equilibrium path, is no longer obtained. For v and $\tilde{t}$ as low as 1 and 0 respectively, mergers already occur along the equilibrium path; specifically, for $(v, \tilde{t}) = (1, 0)$, a merger wave (in the sense of set (c) in Figure 1-3) obtains[35]. Intuitively, there are at most two non-participating firms to each merger (as opposed to four when

[34] In the specification of the sequential game, I could have explicitly modelled each merger stage as two separate stages, a bidding stage by one of the two constituent firms, the "bidder", followed by an acceptance stage by the other constituent firm, the "target". (This is done by Fauli-Oller (2000), for example, who equips low-cost firms with the ability to place "take-it-or-leave-it" bids to merge with high-cost firms.) Had I done this, the same merger outcomes would be obtained in equilibrium, with the observation that the surplus from merger would be captured by the bidder. In this case, one can show that the sequential nature of the merger game, for parameter values in zone (c2), lowers the average bid accepted in equilibrium by the target firms vis-à-vis the game where all three merger decisions are made simultaneously (see also Kamien and Zang 1990, 1993). Merging sequentially can thus lower the cost of setting up multinational firms and reducing competition.

[35] This contrasts with the result in Fauli-Oller (2000) that a "takeover wave" occurs only for a sufficiently large (cost) difference among firms.

$T = 3$) and thus the "outsider output expansion" effect is never negative enough to "undo" the profitability of merging for participating parties.

As for solving the game when the number of merger stages $T \geq 4$, Appendix A.2 provides expressions for the reduced-form profit functions of the different firms, under the different possible market structures, in a game with T merger stages: from these the merger surplus functions are computed. As T increases, the space $\mathcal{P}$ of parameters $(v, \tilde{t})$ shrinks: the downward-sloping boundary at which two-way trade is just feasible shifts inward towards the "origin" $(1, 0)$. One can show that the unprofitability-of-mergers result for parameter values close enough to, and including, $(1, 0)$ (similar to equilibrium (h) in Figure 1-5), and the profitability-of-mergers result for parameter values close enough to, and including, the two-way-trade boundary (similar to equilibrium (a) in Figure 1-5) carry through for this general case. As with the case where $T = 3$, it can further be shown that for all parameter combinations outside $\mathcal{P}$ (where trade may flow only from country A to country B, or not at all), mergers are profitable; a similar equilibrium to (a) in Figure 1-5 obtains. In sum, the main results of the particular $T = 3$ game analysed carry through more generally, namely that (i) cross-border mergers may be unprofitable only when *both* the quality gap and the trade cost are low, and (ii) the profitability of cross-border mergers is widespread, with such mergers bunching together.

1.4 Example 2: Cross-border mergers in the Perry and Porter (1985) model

I now analyse cross-border merger profitability in another standard model that satisfies the sufficient condition of Corollary 1.1, yet where the motive for merger is different in kind. Perry and Porter (1985) introduce a notion of "firm size" by considering a fixed factor of production (say capital) whose total supply is fixed to the industry; what distinguishes one firm from another is the stock of capital it owns. Thus, a "larger" firm arising from the merger of two firms owning the same amount of capital has a lower marginal cost than either of its constituents at a given level of output (to be made precise below).

Perry and Porter (1985) setup Let $k > 0$ denote the amount of the fixed factor of production whose total supply is fixed to an industry producing a homogeneous good. Firm i's cost function is given by $C(x_i, k_i)$, where x_i denotes its output of this good and k_i is the amount of the fixed factor it owns. The cost function is taken to be homogeneous of degree one in output and capital, implying constant returns to scale and that the marginal cost function, $C_1(x, k) := \frac{\partial C(x,k)}{\partial x}$, is homogeneous of degree zero in x and k. Because of the presence of a fixed factor of production, it is assumed that marginal costs are decreasing in k, $C_{12}(x, k) := \frac{\partial^2 C(x,k)}{\partial k \partial x} < 0$ and hence, by Euler's theorem, marginal costs are increasing in output, $C_{11}(x, k) := \frac{\partial^2 C(x,k)}{\partial x^2} > 0$. Firms compete à la Cournot.

Perry and Porter (1985) specify particular functional forms for demand and cost. Both price and marginal cost are assumed to be linear functions of output. The industry inverse demand function is given by $P(X) = a - X$, where P is price and $X := \sum_i x_i$ is industry output. (I adapt this to the two-country setting by taking $P^l(X^l) = a - X^l$, where P^l and X^l are respectively the price and sales in country $l \in \{A, B\}$.) A firm's cost function is quadratic (and convex) in output, $C(x, k) = gk + dx + \frac{e}{2k}x^2$, where industry fixed costs $g \sum_i k_i$ are distributed in proportion to capital ownership. Coefficients d, e and g are (weakly) positive and $a > d$. (In this two-country setting, I again assume a linear trade technology: exports incur a unit trade cost $t \geq 0$.)

Sequential cross-border merger game I take the simplest structure of the sequential merger game considered in Section 1.2. The (global) industry, with capital stock $4k$, consists of four symmetric firms, each owning one-quarter of this stock, $k_i = k$, each country hosting two firms, $n^A = n^B = 2$. With no loss of generality, firms in country A are labelled 1 and 2, their B-country rivals are labelled 3 and 4, the first set of firms that decide whether to merge is $\mathcal{M}_1 = \{1, 3\}$ and the set of firms that subsequently decide whether to merge is $\mathcal{M}_2 = \{2, 4\}$. The merger technology is such that the capital stock of a merged (multinational) firm is the sum of the capital stock of its constituent firms. Notice that the assumptions rule out economies of scale as a motive for merger.

1.4.1 Equilibria as a function of the rate of change of marginal cost and the trade cost

In view of the symmetry (in the sense of Definition 1.2), I resort to Proposition 1.2 in order to solve the game. The reduced-form profit functions for an independent firm and merged firm under every possible market structure, necessary to compute the merger surplus functions, are derived in Appendix A.3. The merger surplus functions Φ are a function of (i) the demand intercept less the marginal cost when output is zero, $(a - d)$, (ii) the trade cost t, and (iii) the rate of change of the marginal cost of a firm with capital stock k, $C_{11}(x, k) = \frac{e}{k}$. Nevertheless, the Φ-functions can be unambiguously signed, pinning down the equilibria of the game, by $\tilde{e} := \frac{e}{k}$ and $\tilde{t} := \frac{t}{a-d}$, which I will refer to, respectively, as the rate of change of the marginal cost and the (normalised) trade cost.

Clearly, the space of parameter values of interest is that where trade between countries is feasible. (Otherwise there is zero surplus from merger.) Again I label this space $\mathcal{P}$, defined by $\tilde{e} \geq 0$ and $(0 \leq)\tilde{t} \leq \frac{1}{3+\tilde{e}}$.

The sufficient condition of Corollary 1.1 holds It is easy to show that the model satisfies the sufficient condition of Corollary 1.1, implying that of the six possible equilibria depicted in Figure 1-3 for a symmetric game with two merger stages, only three equilibria can obtain: sets of strategies (a), (c) and (h). Inspection of (the sign of) the merger surplus functions Φ_I and Φ_{II} in parameter space establishes that all three equilibria do obtain. Further, the way in which the equilibrium changes as parameters vary is similar to Example 1 (Section 1.3). When the rate of change of the marginal cost $\tilde{e}$ and the trade cost $\tilde{t}$ are *both* low, no cross-border merger occurs along the equilibrium path of the game, as well as off it: set (h) is the equilibrium. At the other end, for high enough values of $\tilde{e}$ and/or $\tilde{t}$, mergers are profitable, along the equilibrium path as well as off it: set (a) is the equilibrium. For intermediate values of $\tilde{e}$ and/or $\tilde{t}$, mergers occur along the path yet the early merger by $\mathcal{M}_1$-firms induces the subsequent merger by $\mathcal{M}_2$-firms: set (c) is the equilibrium. The zones in which the different strategies are supported as equilibria are labelled accordingly as zones (a), (c) and (h). The complementary sets of strategies in Figure 1-3, namely (d), (e) and (f), cannot be supported as equilibria for any combination of parameters. The following proposition summarises the result.

Proposition 1.4 *Space $\mathcal{P}$ can be partitioned into three disjoint zones, as depicted in Figure 1-7. A straight line segment drawn from $(\tilde{e}, \tilde{t}) = (0, 0)$ to any point on $\tilde{t} = \frac{1}{3+\tilde{e}}$ begins in zone (h), crossing into zone (c) and ending in zone (a). In each zone, the unique equilibrium is given in Figure 1-3, labelled after the zone. The boundaries between zones (h) and (c), and between zones (c) and (a), are continuous and downward-sloping, joining a point $(\tilde{e}, 0)$, $0 < \tilde{e} < \infty$, on the $\tilde{e}$-axis to a point $(0, \tilde{t})$, $0 < \tilde{t} < \frac{1}{3}$, on the $\tilde{t}$-axis. The boundary between zones (h) and (c) lies strictly below the boundary between zones (c) and (a).*

The parallel to Proposition 1.3 is clear, both results being driven by a similar underlying mechanism – the reaction of non-participating firms – despite the motive for merger being different in kind: "growth in size" presently as opposed to "technology transfer" in Section 1.3, in addition to dampening competition and jumping tariffs, which both models share. That the response of non-participating firms again takes centre-stage in explaining the profitability of cross-border mergers is intuitive. When *both* the rate of increase of marginal cost and the trade cost are low, non-participating firms respond to a merger by significantly expanding output, rendering the merger unprofitable to its participants. As either the rate of increase of marginal cost or the trade cost increase from these low levels, the "outsider output expansion" effect (see the decomposition analysis in Section 1.3) falls in magnitude and merger profitability increases. As in Example 1, cross-border mergers are widespread and do not occur in isolation, but rather bunch together.

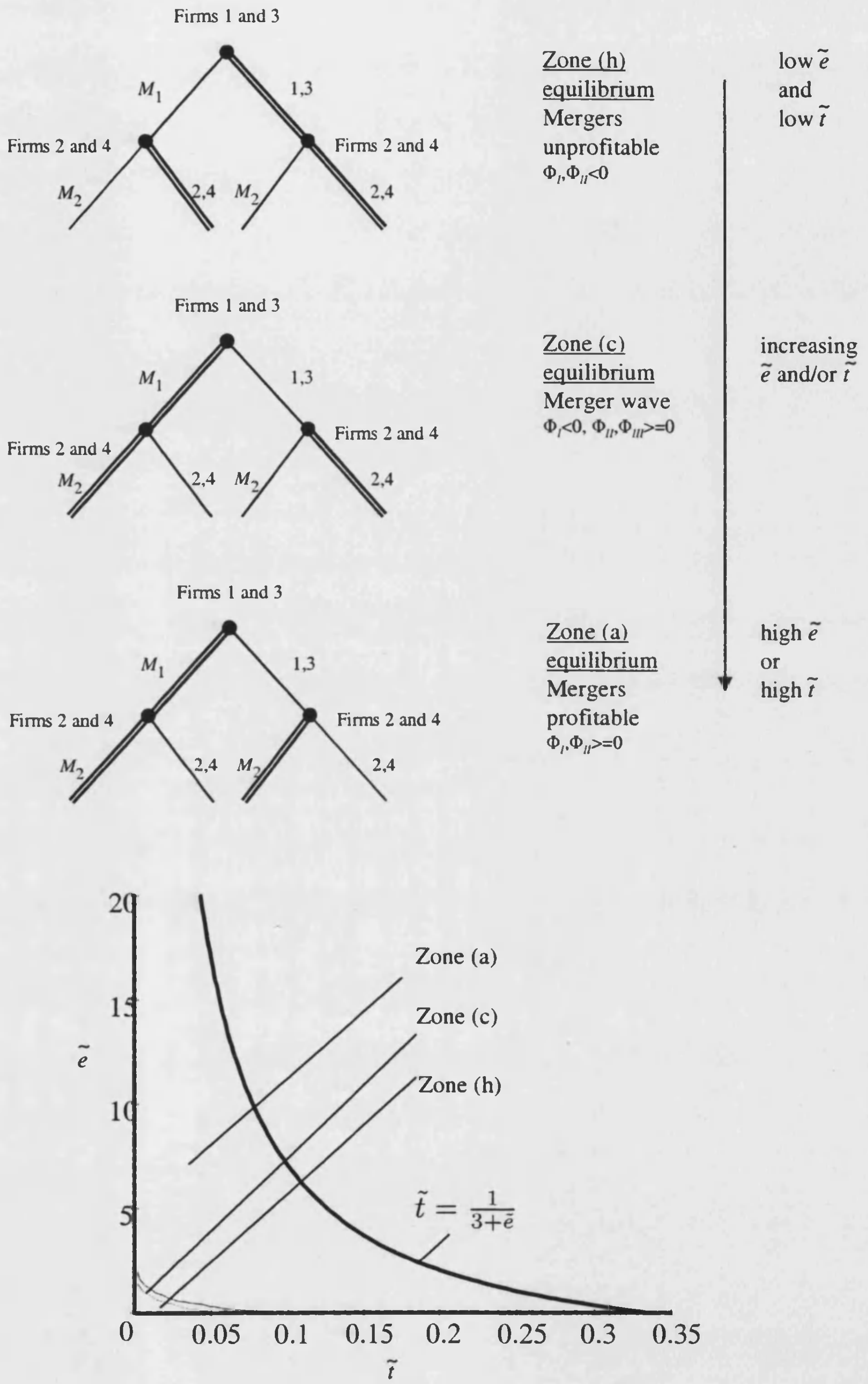


Figure 1-7: (Example 2) Equilibrium of sequential merger game in each zone and zones in $\mathcal{P}$ space; $\tilde{e}$ rate of change of marginal cost, $\tilde{t}$ trade cost

1.5 Revisiting the vertically-differentiated-industry model: Investment integration and fixed costs of merger

The solutions to the two standard models above share the common phenomenon that cross-border mergers are undertaken and bunch together for a "large" region of parameter space. These results, however, beg the question: if the all-merger outcome indeed occurs *along* the equilibrium path for these parameter values, why is it that cross-border mergers are often empirically observed to occur from a sudden point in time onwards (and usually concentrated over a short period)? In other words, why were they not undertaken earlier?[36]

I revisit the vertically-differentiated-industry model in Section 1.3 to introduce the notion of a fixed cost G associated with implementing a cross-border merger. Borrowing from the historical motivation for the quality asymmetry in the model, I ask why do cross-border mergers not take place even before countries undergo unexpected trade and investment integration[37]?

I argue that initially, before countries open up their industries to foreign investment, or lift curbs on the right of foreign firms to acquire shares in domestic firms, this fixed cost G of implementing a cross-border merger is very high, to the extent that cross-border mergers are not profitable and no merger occurs along the equilibrium path. Then, as countries open up to foreign investment, with the relaxation of these curbs on the right of foreign firms to merge with (or acquire) domestic firms, and the cost of doing business (such as acquiring information, or communicating) in a new environment declining as barriers are pulled down, G falls gradually[38]. In the historical context, the point in time at which G begins to fall corresponds to the moment at which unexpected integration takes place.

[36] In particular, one setting in which this flurry of investment activity has been seen to happen is that of countries undergoing integration, such as the European Union (EEC) from the 1980s, or several newly-industrialised countries undergoing trade and investment reform during the 1990s.

[37] By the results of that section, mergers are profitable in autarky (outside $\mathcal{P}$ for $\tilde{t}$ high enough) in the presence of a quality gap ($v > 1$).

[38] According to UNCTAD (2000), "over the period 1991-1999, 94% of the 1,035 changes worldwide in the laws governing FDI created a more favourable framework for FDI. Complementing the more welcoming national FDI regimes, the number of bilateral investment treaties – concluded

I investigate how the solution to the model, for the different zones in $(v, \tilde{t}) \in \mathcal{P}$, changes as the fixed cost G associated with implementing a cross-border merger declines from an initial high level G_0 (at which mergers do not occur). Figure 1-8 illustrates the solution for $(v, \tilde{t})$ in each of zones (a), (c1) and (c2) as G falls[39]. When G reaches a threshold given by $\Psi_{III}(v, \tilde{t})$, the equilibrium outcome (Merger, Merger, Merger) is reestablished for $(v, \tilde{t})$ in each of zones (a), (c1) and (c2). Interestingly, when this threshold $G = \Psi_{III}(v, \tilde{t})$ is reached, the equilibrium for $(v, \tilde{t})$ in *each* of zones (a), (c1) and (c2) (the second game tree from top to bottom in Figure 1-8) replicates the equilibrium for zone (c2) in the absence of fixed costs, depicted in Figure 1-5, which was referred to as a "merger wave"[40 41].

1.6 Concluding remarks

This chapter has sought to explain why the pattern of equilibria in sequential (horizontal) merger games of a certain type studied in the literature is similar across a fairly wide class of models. In this class of models typically only a similar subset of the possible set of equilibria is obtained. By developing general conditions characterising each possible equilibrium, I have shown that the solution to models that exhibit symmetry (in the sense defined earlier) and satisfy a certain sufficient condition will be restricted to this subset of equilibria. These models capture two (commonly) alternative types of situation: (i) one where firms are better off under

increasingly also between developing countries – has risen from 181 at the end of 1980 to 1,856 at the end of 1999. Double taxation treaties have also increased, from 719 in 1980 to 1,982 at the end of 1999. At the regional and interregional levels, an increasing number of agreements (most recently between the EC and Mexico) are helping create an investment environment more conducive to international investment flows" (pp.xv). See also Caves (1991).

[39]The proof is in Appendix A.2. Clearly, in zone (h), no cross-border mergers occur along the equilibrium path for all values of $G \geq 0$: if this was the case in the absence of fixed costs of implementation ($G = 0$), it remains the case when these are introduced ($G > 0$).

[40]A relevant question at this point is: if firms forsee that G will fall further, why not wait for this to occur and thus guarantee a larger surplus from merger? Though this can be ruled out by making firms (sufficiently) impatient in a dynamic setting, or unable to forsee the future trajectory of G, the relevance of the question remains.

[41]As G falls further, assuming no mergers have occurred, for $\Psi_I(v, \tilde{t}) < G \leq \Psi_{II}(v, \tilde{t})$, the equilibria in zones (a) and (c1) (see the third tree in Figure 1-8) replicate the equilibrium for zone (c1) in the absence of fixed costs (the equilibrium in zone (c2) still replicates the equilibrium for zone (c2) in the absence of fixed costs). For $0 \leq G \leq \Psi_I(v, \tilde{t})$, assuming no mergers have occurred, the equilibrium in zone (a) (see the fourth tree in Figure 1-8) replicates the equilibrium for zone (a) in the absence of fixed costs (the equilibria in zones (c1) and (c2) still replicate the equilibria for zones (c1) and (c2) respectively in the absence of fixed costs).

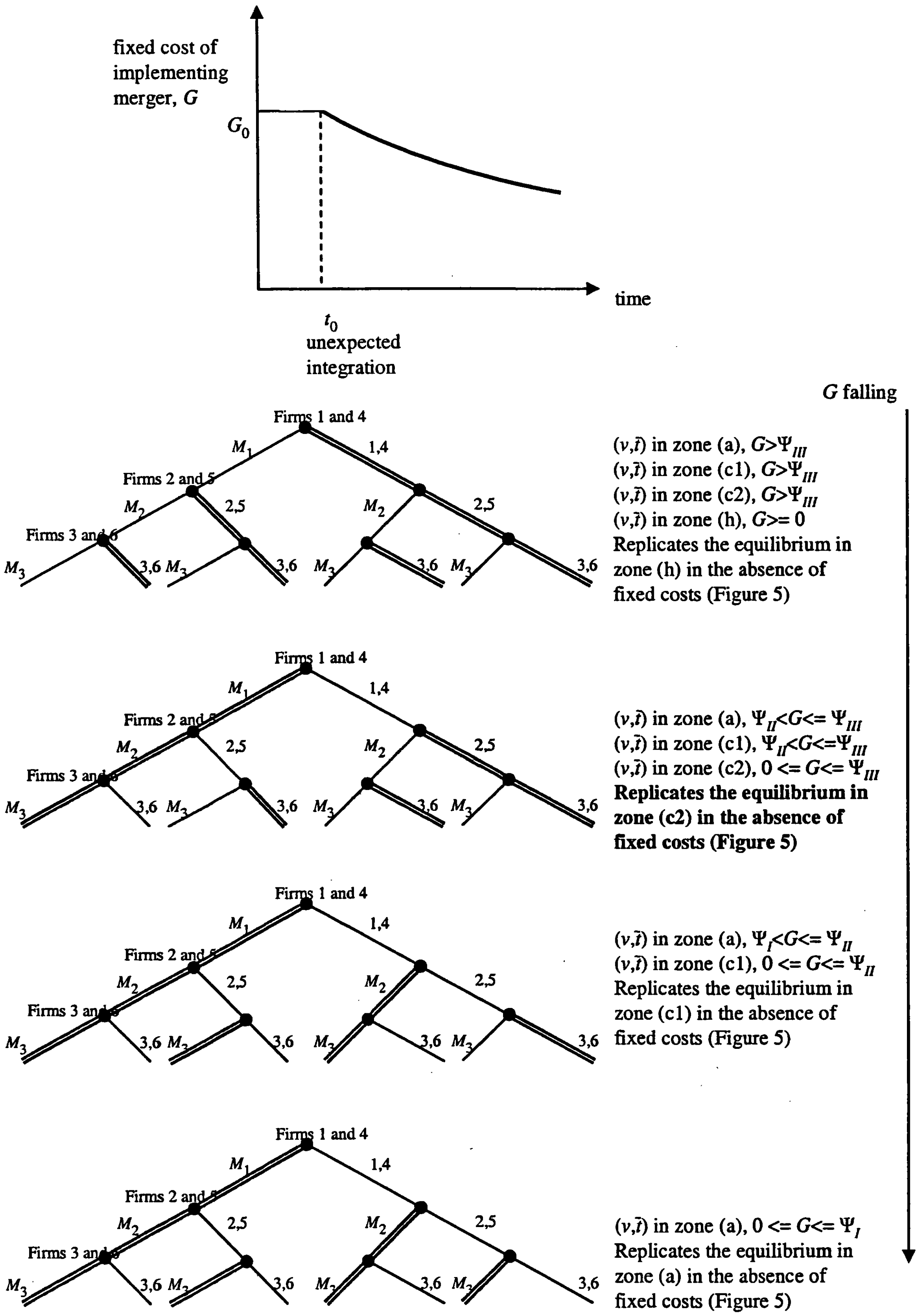


Figure 1-8: (Example 1) Investment integration and sequential cross-border mergers

market structures where their rivals have merged as against market structures where rivals remain independent, or (ii) one where by merging firms may gain a competitive edge over their rivals. This result is of empirical relevance in that the pattern of equilibria found for this class of models is associated with mergers happening, not in isolation, but rather in bunches.

I have then employed the general analysis to solve two other standard models where the sufficient condition holds, illustrating that despite the motive for merger being different in kind – one where mergers are a means of technology transfer, the other where mergers enable firms to grow their holdings of capital – the pattern of equilibria is again similar. In so doing, I have utilised a two-country setting to show that the results can easily be extended to cross-border mergers. I argued that sequential games can be used to explore the profitability of cross-border mergers and can cast light on the phenomenon of bunching. Cross-border mergers are widespread and bunching occurs for "most" combinations of parameter values. The occurrence of mergers hinges critically on the magnitude of the reaction by non-participating firms being limited (e.g. when there is sufficient differentiation along either vertical or horizontal dimensions, as in the first model).

Much research to date on foreign investment has studied the export versus greenfield investment decision by firms as countries undergo market integration. Such models assume that as borders open up and there is increased competition, firms decide whether to enter new markets via exports or by setting up production operations, concentrate on existing markets or exit altogether. In reality, oligopolistic firms have a larger action set which includes merging with one another. Modelling foreign investment in the form of cross-border mergers – in counterpoint to greenfield investment – suggests just how widespread such activity may be. By merging or buying their way into foreign markets firms can dampen competition, a feature which is absent in simple exports-versus-greenfield investment studies. This helps explain the empirical predominance of mergers and acquisitions as a channel of investment by firms in foreign markets.

Finally, modelling mergers in the form of a sequential game as opposed to a game with simultaneous moves, captures the interdependence result – where earlier mergers trigger subsequent mergers – holding for certain regions in parameter space. In these regions the sequential nature of moves acts as a coordination device ensuring that the all-merger equilibrium outcome is unique. This may further help explain the empirical predominance and bunching of cross-border mergers.

Chapter 2

Inferring Conduct under the Threat of Entry: The Case of the Brazilian Cement Industry

2.1 Introduction

Empirical industrial organisation has long been concerned with attempting to measure the degree of market power enjoyed by firms in an industry. Where marginal cost is observed, or can be constructed from known technological parameters, market power can be inferred from the distance between price and cost, informing the researcher of the pattern of firm conduct in the industry. More often than not, marginal cost is not observed. In such cases, a well-established approach in the literature attempts to ascertain firm conduct, along with cost, from the comparative statics of equilibrium. By this approach, a static structural model is specified, typically consisting of a parametric system of demand and firm pricing equations (first-order conditions). One then proceeds to jointly estimate demand, cost and a conduct parameter – indexing the degree of market power – from price and quantity data, as these are moved around by observed exogenous shocks to supply and demand. Such a methodology for estimating cost and conduct, developed by Bresnahan (1982) and Lau (1982), turns on the identifying assumption of orthogonality between the errors of the firm's pricing equation and the excluded exogenous variables which move the demand curve. Intuitively, shocks to demand allow one to observationally distinguish between the hypothesis of a high-cost competitive industry and the hypothesis of a low-cost collusive industry because the response of prices to these shocks is different according to the kind of firm behaviour in the industry: while price-taking firms set output at the point where price equals marginal cost as demand moves exogenously, firms with market power change prices to ensure that marginal revenue is equated to marginal cost.

In this chapter I examine the identification of supply (cost and conduct) in a more general dynamic setting where an industry faces potential entry and this threat of entry constrains pre-entry prices. I develop a structural model of a domestic oligopoly which faces a competitive fringe of elastically-supplied high-cost imports. In equilibrium no imports are observed yet the threat of imports sets an upper limit on prices, equal to the marginal cost of imports. I show that when this price limit binds at the industry equilibrium, the identifying assumption of orthogonality between the error term of the conventional pricing equation and the excluded exogenous demand variables does not hold. Because the constraint posed by the threat of imports is unaccounted for, the standard pricing equation specification does not identify cost and conduct from the comparative statics of demand. In this setting, the standard methodology does not allow one to empirically distinguish the hypothesis of a high-cost competitive industry from the hypothesis of a low-cost cartel where imports restrain prices at the equilibrium. Intuitively, the response of prices to fluctuations in demand is no longer distinct because the threat of imports acts to constrain the ability of the cartel to set marginal revenue equal to marginal cost. Viewed from a different angle, equilibrium market price elasticities of demand are no longer informative since, irrespective of the hypothesis of conduct, the equilibrium remains at the kink of the residual demand curve facing the domestic oligopoly as demand fluctuates.

It is natural that upon not observing imports in equilibrium, or observing no more than a minimal amount, a researcher may come to overlook the restraining effect of imports, thereby misspecifying the structural model for the industry and estimating a static pricing equation imposing the regular moment conditions. It follows that if the industry enjoys market power, in the sense that equilibrium price-cost margins are positive, and yet the threat of high-cost imports constrains equilibrium prices, the lack of responsiveness of prices to fluctuations in demand will lead to the underestimation of the true degree of market power.

To illustrate, I turn to an industry in a country where potential imports restrain market outcomes: the Brazilian cement industry. Unlike its US counterpart, where the penetration of imports has ranged between 10 and 20% of domestic consumption in the past decades, the Brazilian cement industry has historically managed to keep imports at bay[1]. This has been achieved thanks to a combination of domestic price controls, trade barriers, poor infrastructure and a depreciated local currency. At the turn of the 1990s, as price controls were lifted and Brazil began opening up to trade, the threat of imports began to develop "bite" on the industry, reinforced by an appreciating local currency. I consistently estimate demand in each local market (state of the Brazilian federation) and find very low market price elasticities of demand in equilibrium, of the order of -0.5. Two main possibilities arise to rationalise why an industry facing such inelastic demand does not cut output to raise prices to a point where demand is more elastic: (i) there is weak pricing power (e.g. competition or low concentration), or (ii) some dynamic story is appropriate, such as the threat of entry (imports) restraining pre-entry prices[2]. By way of an illustration, I begin by ignoring the restraining effect of imports and follow the standard methodology, estimating a pricing equation and instrumenting with exogenous demand. I obtain cost estimates that are close to prices, suggesting that outcomes in the Brazilian cement industry are competitive (and the standard conduct parameter is estimated to be close to zero). To check these estimates, I then construct actual marginal cost from observed factor prices, the simple fixed-coefficient nature of cement production technology, and the observed flow of cement from plants to markets. In contrast to the estimated price-cost margins that are

[1]The rise of international trade in cement, though puzzling to some, is a fairly recent phenomenon and has been documented carefully by Dumez and Jeunemaître (2000). Despite high inland transportation costs, cement can travel – and does travel – quite cheaply from afar by sea via specialised equipment.

[2]A third possibility hinges on a very special class of models of spatial competition, where a firm is restricted to set only a "mill" price, with delivered prices to consumers who are distributed over space being equal to the sum of this mill price and the transportation cost.

centred around zero, actual price-cost margins are large, amounting to 40-65% of producer prices (net of sales tax). Producers enjoy considerable market power despite the binding high-cost imports constraint[3]. Thus the standard methodology fails to identify supply, severely underestimating the observed degree of market power[4]. This illustrates the theoretical result that when the threat of entry constrains prices, joint inference of cost and conduct will not be consistent because of the lack of responsiveness of prices to fluctuations in demand.

Given that cost and conduct cannot be jointly identified from the comparative statics of equilibrium in this constrained setting, the immediate question is: but what if the researcher *observes marginal cost*? Clearly, a direct comparison of marginal cost to price will provide a test of perfectly competitive behaviour against less competitive models of firm behaviour (and where imports restrain prices). But, other than perfect competition, how may one distinguish empirically between two alternative models of behaviour when *aggregate* (market) outcomes in both of these models are constrained by the threat of imports, and are thus equal? In the Brazilian cement industry, for example, where high-cost imports restrain equilibrium prices, one may wish to identify the model of firm behaviour supporting the constrained outcomes. I propose a test of conduct against a standard benchmark, the Cournot oligopoly solution. The measure I develop uses firm-level quantity data to test the hypothesis of Cournot conduct against the alternative of "more collusive" firm behaviour. It is predicated on the notion that no Cournot firm can perceive that marginal revenue (taking rivals' output as given on the margin) exceeds marginal cost, otherwise the firm would optimally expand output, and this notion holds regardless of whether the imports constraint binds or not in equilibrium[5]. The requirements on the data are large, but the value of the test lies

[3]That the high price ceiling set by the high-cost imports binds is then a consequence of the steepness of the demand curve and this strong price discipline in the industry.

[4]The estimated coefficients on factor prices and other supply-shifters are mostly of the expected sign and significant, which could again mislead a researcher into thinking that the econometric model is appropriately specified. But this owes only to the fact that the estimated coefficients are picking up the expected correlation between cement prices and factor prices.

[5]The reverse notion, that a Cournot firm optimally cuts output when its perceived marginal revenue (were imports not to exist) falls short of marginal cost, no longer holds when the price ceiling imposed by imports binds in equilibrium: cutting output in an attempt to raise price above the price ceiling simply opens the door to imports.

in uncovering firm-level behaviour when market outcomes are constrained by the threat of entry and thus the comparative statics of equilibrium are not informative.

I illustrate the proposed test of conduct by reference to the Brazilian cement industry. I find that conduct across local markets is considerably more collusive than the Cournot benchmark. Market outcomes are characteristic of (tacit) market division, *and this can be identified despite the threat of imports restraining prices.* A story where firm 1 tacitly agrees to give firm 2 the upper hand in market B in exchange for the latter staying away from market A – with typically firm 1 (firm 2) being located slightly closer to market A (market B) than the rival firm[6] – helps to explain the observed shipments. Plants ship to local markets located at their doorstep, while restricting supply to adjacent markets, despite supplying to these latter markets being highly profitable under the static Cournot conjecture[7].

This chapter thus makes three contributions. *First,* it demonstrates that the standard pricing equation specification does not identify cost and conduct in industries where potential entry restrains pre-entry market outcomes, such as domestic oligopolies facing (underlying) competition from abroad[8]. The conventional identifying assumption is not satisfied; in particular the estimated degree of market power will be biased downwards. The implication of this latent effect of imports for antitrust authorities attempting to measure the competitiveness of conduct is increasingly relevant in a world where trade barriers are being pulled down[9]. *Second,* I develop a test of conduct in such settings where potential entry may constrain equilibrium prices. By reference to this test, I show that market outcomes in the

[6] "Slightly closer" is employed in the sense that the freight cost, while important to the cost structure of the industry, clearly does not explain the observed pattern of shipments.

[7] Chapter 3 considers, in light of the different local market structures observed in the Brazilian cement industry, simple dynamic multimarket games which give rise to a pattern of market division in equilibrium..

[8] In addition to potential entry, other "invisible" constraints may be conceived, such as pressure from antitrust authorities. Antitrust authorities are typically fond of "barking" at industries perceived to have market power in the hope of exerting downward pressure on prices. To the extent that some pressure is indeed exerted – i.e. some of the *bark having bite* – such a restraint provides another channel by which the standard methodology yields a downward-biased estimate of market power.

[9] The constraint of imports on market outcomes is not new, however. As I later comment, in a study of the US sugar industry at the turn of the 20th century, Genesove and Mullin (1998) state that "industry pricing was constrained by threats of (domestic) entry or of foreign imports" (p. 367).

Brazilian cement industry are indicative of (tacit) market division. The *third* contribution has policy implications in relation to the cement industry, particularly in developing countries. In a developing country such as Brazil, with its huge housing deficit and infrastructure needs, the importance of the cement industry cannot be overstated. Cement is an essential input to construction and building activity for which there are few substitutes. The industry regularly attracts attention from antitrust authorities, consumer associations and the financial media for its alleged pricing power. Yet to date no study has been undertaken to empirically ascertain the degree of competition in the industry by estimating a structural model using a rich original dataset. This study attempts to fill this void. A clear policy recommendation is that fostering imports can play an important role in curbing the ability of domestic producers to raise prices above marginal cost. In Brazil, recent policy experience has been the opposite; the government has succumbed to the industry's "anti-dumping" lobby and raised the barriers to entry of imports, to the detriment of consumer welfare.

The plan of the chapter is as follows. In Section 2.2 I develop the theoretical framework and address identification. I then turn to institutional aspects of the cement industry, and present the data. Section 2.4 presents the application. (Appendix B provides a discussion of the sources and treatment of the data used, and robustness checks regarding the construction of marginal cost.) Finally, I conclude, reflecting on the policy implications of this chapter.

2.2 Theoretical framework: Towards a test of conduct

In this section I develop a structural model of an oligopoly facing potential entry, where this threat of entry may limit pre-entry prices. Potential entry is modelled as a competitive fringe of foreign suppliers (imports) to the domestic oligopoly market. I then extend the analysis of identification of cost and conduct underlying the standard methodology from the static setting considered by Bresnahan (1982, 1989) to the present constrained setting, where the threat of imports may be constraining market outcomes. I show that the identifying assumption of orthogonality between the error term of the conventional pricing equation and the excluded exogenous demand variables does not hold. Estimates of cost and conduct parameters will be inconsistent; in particular, the degree of competition will be overestimated.

I then consider how a researcher may learn about conduct in an oligopoly facing potential competition from imports. I consider a situation where firms in the domestic oligopoly meet in different spatial (or product) markets and two data requirements on the part of the researcher are met: (i) direct measures of marginal cost *are* available, and (ii) firm-level quantity data is available (at the local market level). I specify a test of conduct based on Cournot behaviour adapted to the constrained setting. This test can reveal details regarding the pattern of conduct prevailing in the industry, despite aggregate outcomes being constrained in equilibrium by the presence of imports.

2.2.1 Domestic monopoly facing competition from imports

Consider a monopolist M producing a homogeneous good at flat marginal cost c_M. The monopolist faces a competitive fringe of foreign suppliers (labelled I for imports), with perfectly-elastic supply at marginal cost $c_I > c_M$. In general, the equilibrium is given by either of two situations. If the marginal cost of imports is lower than the monopoly price in the absence of imports (denoted p^M), the price in equilibrium will be equal to the marginal cost of imports, the monopolist will supply the entire domestic market, yet the foreign fringe exerts downward pressure on price. Alternatively, if the marginal cost of imports exceeds the monopoly price p^M, imports have no "bite" and the equilibrium price will be p^M, with the monopolist again supplying the entire market though in an unconstrained manner. Formally, the equilibrium price p is given by

$$p = \begin{cases} c_I & \text{if} \quad p^M \geq c_I \\ p^M & \text{otherwise} \end{cases}$$

where $p^M = p(q^M)$, $p(q)$ is the inverse demand function and q^M is the quantity that equates the market marginal revenue $MR(q)$ to the monopolist's marginal cost c_M.

Given the assumption that $c_I > c_M$ (imports are high cost), the monopolist always supplies the entire market. Clearly, when $p^M \geq c_I$, the extreme result of imports commanding zero sales rests on the assumption of perfectly-elastic supply from the foreign fringe[10].

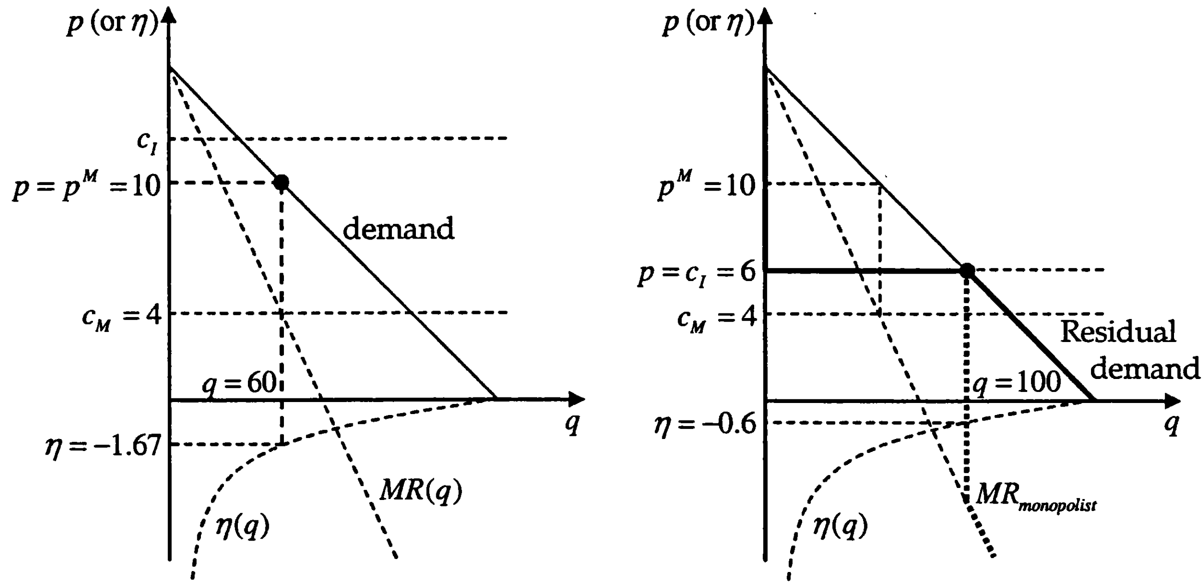


Figure 2-1: Monopolist facing a competitive fringe. Left panel: Imports have no "bite" ($p^M < c_I$). Right panel: Imports constrain price in equilibrium ($p^M \geq c_I$).

An illustration 1 Assume demand is linear, given by $p = 16 - \frac{1}{10}q$ and $c_M = 4$. As shown in Figure 2-1, the monopoly price is $p^M = 10$. If $c_I \geq 10$ (left

[10] As is typically the case with "limit price" models such as the one developed here, one needs to deal with the following question: why does the monopolist need to set a price as low as the limit price, in order to stave off entry (imports), if what is relevant to the entry decision is the *post-entry* price rather than the pre-entry one? In considering such a question, one could ponder why (in the situation of interest where imports do have bite, $p^M \geq c_I$) the monopolist would not set the monopoly price p^M if it is able to cut its price to the importer's cost c_I immediately upon entry. When entry (and exit) is free and there is no entry lag, as is the case in the setting I am considering, where the entrant consists of opportunistic imports in a well-functioning international trading market, the monopolist will set the limit price. In this chapter's application to the Brazilian cement industry, I later consider institutional aspects of the industry that further support the use of a limit-price model. More generally, it may also be argued that the threat of entry does constrain pre-entry prices, even if entry were not free, since pre-entry prices reveal information to the entrant about post-entry prices, such as the cost of the incumbent and/or its disposition to fight if faced with entry (i.e. a predation story of the reputational type). Finally, note that for the purpose of this thesis one need not strictly interpret c_I as the exact cost of imports, but more loosely as a price lower than that set by the monopolist were it to act in an unconstrained manner.

panel), the equilibrium price is $p = 10$ and the monopolist supplies $q = 60$. Notice that the equilibrium market price elasticity of demand $\eta(q) := \frac{\partial \ln q}{\partial \ln p(q)} \simeq -1.67$. Now consider $(4 <) c_I \leq 10$, say $c_I = 6$ as illustrated in the right panel. The market equilibrium now lies at the kink in the residual demand curve faced by the monopolist: $p = 6$ and the monopolist supplies $q = 100$. At the equilibrium, while the price elasticity of demand faced by the monopolist is infinitely high in absolute value, the market price elasticity of demand is only $\eta(100) = -0.6$. Around this latter equilibrium, fluctuations in the marginal cost of imports, say due to fluctuations in the exchange rate, allow one to trace out the demand curve since the kinked equilibrium moves up and down along the market demand curve.

2.2.2 Econometric identification of conduct when costs are not observed

Market demand parameters may be identified from standard cost-shifters excluded from the demand function (left panel of Figure 2-1) or, in the case where imports restrain prices, from fluctuations in the marginal cost of imports, such as movements in the exchange rate (right panel). But, in the absence of information on cost, will *conduct* (and thus cost) be identified from the comparative statics of demand?

In the absence of potential entry, conduct is identified (Bresnahan 1982)
When imports do not constrain prices, as in the left panel of Figure 2-1, conduct is identified from fluctuations in the demand curve. Suppose we wish to distinguish between alternative behavioural hypotheses generating observed price and quantity data: on the one hand a low-cost monopoly or cartel (with cost c_M), and on the other hand a high-cost competitive industry (with cost c_C). When marginal cost is flat in quantity, mere shifts in the demand curve suffice to empirically distinguish the behaviour of a cartel from that of a competitive industry. Rotations of the demand curve will likewise identify conduct. (When marginal cost varies in quantity, only rotations of the demand curve will identify conduct.) Thus demand, cost and conduct parameters are jointly estimated from observed price and quantity

data and observed exogenous demand and cost shifters. The reasoning is captured intuitively in Figure 2-2, where marginal cost is flat. The left panel indicates how a shift in the demand curve has different effects on the initial industry equilibrium E_1 according to the hypothesis of conduct: the equilibrium shifts to E_2^C if pricing is competitive while shifting to E_2^M if there is market power (output expands only to where marginal revenue equals marginal cost). Similarly, the right panel illustrates how demand rotators identify conduct: there is no effect on the industry equilibrium if pricing is competitive (i.e. $E_1 = E_3^C$), yet the equilibrium shifts to E_3^M under a cartel.

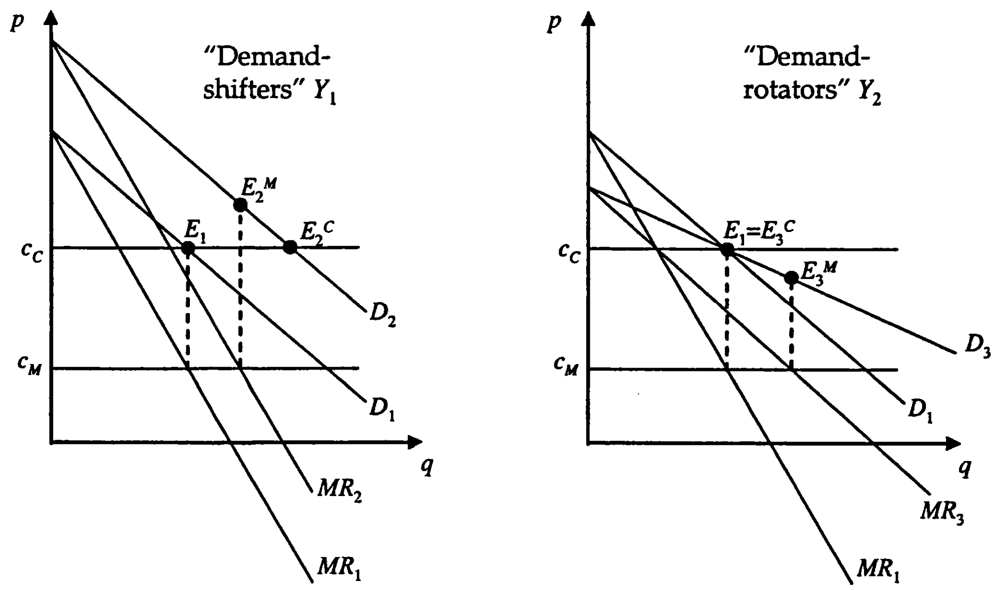


Figure 2-2: Identification in a static model. Left panel: Demand shifts. Right panel: Demand rotates.

When potential entry constrains pre-entry prices, conduct is *not* identified Now modify the cartel hypothesis so that the domestic industry with low cost c_M faces a competitive fringe of imports, with perfectly-elastic supply at high cost c_C that constrains price, as in the model of Section 2.2.1 (with $c_I = c_C$). In other words, under the hypothesis of a cartel, imports restrict the price to be c_C and the equilibrium lies at the kink in the residual demand curve facing the domestic industry (otherwise, for a high enough marginal cost of imports, we are back to the unconstrained situation considered in Bresnahan 1982). We wish to empirically distinguish this *constrained* low-cost cartel hypothesis from the alternative hypothesis of a high-cost competitive (domestic) industry, with cost c_C (and where the presence of imports becomes irrelevant). In Figure 2-3, the equilibrium is initially at E_1. In the left panel, a shift in the demand curve moves the equilibrium to E_2 under both alternative hypotheses. In the right panel, a rotation of the demand curve around the equilibrium point E_1 does not move the equilibrium point under either hypothesis. Thus, in this dynamic setting, unless marginal costs are observed, there is no observable distinction between the hypothesis of a low-cost cartel (with imports restraining prices to be c_C) and the hypothesis of high-cost (domestic) competition[11].

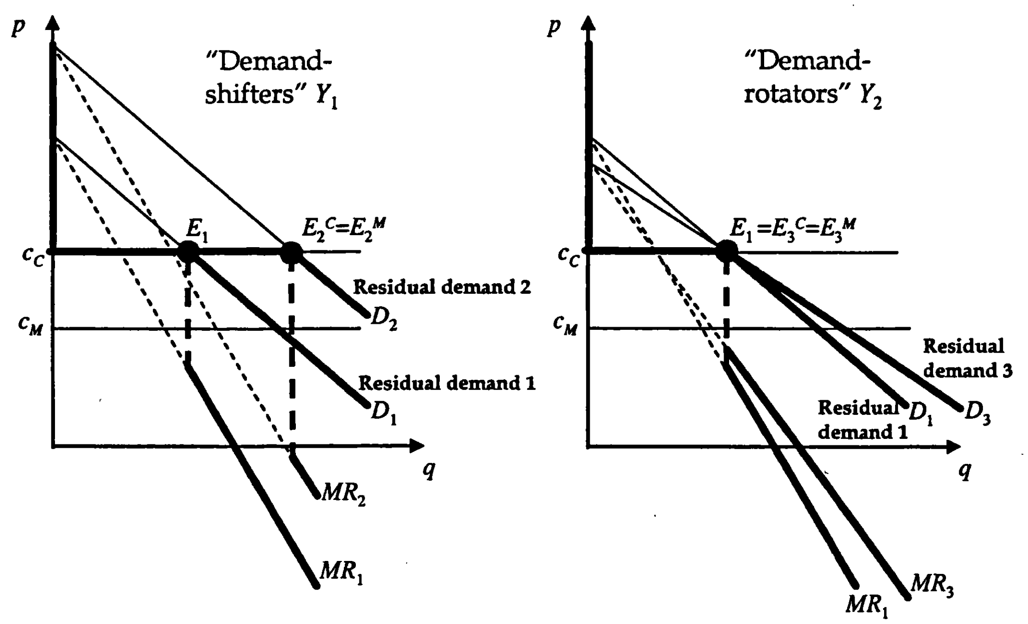


Figure 2-3: Conduct is no longer identified under the threat of entry. Left panel: Demand shifts. Right panel: Demand rotates.

Why is conduct no longer identified? Consider, say, rotations of the demand curve. Intuitively, in the absence of imports, such rotations identify conduct be-

[11]Fluctuations in the demand curve can be broken down into rotations around the price intercept and parallel shifts. When marginal cost is flat, identification is possible only from parallel inward shifts of the demand curve, when these shifts are sufficiently large (we are then back to the situation considered in Bresnahan 1982). Identification is not possible for rotations around the price intercept.

cause firms with market power change prices when demand moves exogenously to ensure that marginal revenue is equated to marginal cost. Now, the threat of imports acts to constrain their ability to set (market) marginal revenue equal to marginal cost and therefore removes the source of price variation which allows conduct to be identified. Viewed from a different angle, equilibrium market price elasticities of demand are no longer informative since the equilibrium lies at the kink of the residual demand curve facing the domestic oligopoly (see below).

Thus, comparing figures 2-2 and 2-3, notice that while fluctuations in the demand curve in the former figure lead to changes in prices under monopoly but not under competition, these same fluctuations in the latter figure leave prices unchanged under both monopoly (facing imports) and competition. In the right panel of Figure 2-3, for example, a rotation of the demand curve around the initial equilibrium point E_1 does not change the equilibrium price (and quantity) under both hypotheses of competition and collusion. Overlooking the effect of imports and misspecifying the structural model to be that captured in the right panel of Figure 2-2 (Bresnahan 1982), a researcher would interpret the stationarity of equilibrium prices as evidence to reject (low-cost) collusion in favour of (high-cost) competition, regardless of the true behavioural model generating the data. (Similarly, for a shift in the demand curve as depicted in the left panel of Figure 2-3, the equilibrium price would again remain unchanged under both hypotheses of competition and collusion. By misspecifying the structural model to be that captured in the left panel of Figure 2-2, one would mistakenly reject collusion in favour of competition, since no price change is observed following a shift in the demand curve.) The more general point is that by misspecifying the structural model and not accounting for the price-constraining effect of imports (or entry), a researcher could be misled into overestimating the extent of competition in the industry, further (residually) overestimating costs.

Estimation of a static pricing equation In the empirical literature on conduct, the following static pricing equation is typically specified on the supply side[12]:

$$p + \theta q \frac{\partial p(q)}{\partial q} = c \tag{2.1}$$

where p is price, q is industry output, c is marginal cost and θ is a conduct parameter. One reason why specification (2.1) may have become so popular is that it nests first-order conditions corresponding to the oligopoly models of monopoly or perfect collusion (where the firm internalises the aggregate inframarginal revenue change from a marginal change in output, so that $\theta = 1$) and perfect competition (where $\theta = 0$), among other models (e.g. symmetric Cournot, θ being the reciprocal of the number of firms in the industry)[13]. Pricing equation (2.1) can be rearranged to the familiar "elasticity-adjusted Lerner index" (or price-cost mark-up):

$$\theta = -\eta(p)\frac{p-c}{p} \tag{2.2}$$

where $\eta(p)$ is the market price elasticity of demand. Clearly, such a specification captures the supply decisions depicted in Figure 2-2, in the absence of imports, under the alternative hypotheses of conduct[14]. Econometrically, (2.1) may be implemented by including a zero-mean error term ε^s and proceeding to the estimation

[12]Examples include Gollop and Roberts (1979), Roberts (1983), Porter (1983), Suslow (1986), Bresnahan (1987), Brander and Zhang (1990), Nevo (2001) and Slade (2004). Note that here I abstract from a criticism that has been made regarding this approach on the grounds that oligopoly theory to date does not underpin a continuum of values for conduct that would support its free estimation (see, e.g., Reiss and Wolak 2002).

[13]Note that (2.1) can be specified at the industry or at the firm level. In the latter case one may include a subscript f for the conduct and cost parameters, to denote the firm. An industry-level pricing equation can be viewed as the average across firms' individual pricing equations (weighted or not by firms' shares), in which case θ has the interpretation of "the average collusiveness of conduct" (Bresnahan 1989). Note also that a common alternative to (2.1) in the literature consists of replacing the inframarginal revenue term $\theta_f q \frac{\partial p(q)}{\partial q}$ by $\theta_f q_f \frac{\partial p(q)}{\partial q}$, i.e. replacing industry output q by firm output q_f in firm f's pricing equation. From the first-order condition, the conduct parameter then corresponds to dq/dq_f, which some have interpreted as a "conjectural variation": by this view, upon expanding its output by dq_f, firm f would hold a "conjecture" dq with respect to the resulting aggregate output expansion.

[14]Notice that the vertical distance between the (inverse) demand function and the marginal revenue function is equal to $-q\frac{\partial p(q)}{\partial q}$. Under the unconstrained monopoly equilibrium of Figure 2-2 this distance is always equal to $p - c$. In contrast, in the constrained monopoly equilibrium of Figure 2-3 this distance will exceed $p - c$.

of

$$p = -\theta q \frac{\partial p(q)}{\partial q} + c + \varepsilon^s \tag{2.3}$$

where p and q are observed, $\frac{\partial p(q)}{\partial q}$ has previously been consistently estimated, and one wishes to estimate θ and c. Since $q\frac{\partial p(q)}{\partial q}$ is endogenous, one needs to find excluded instruments from (2.3). In the absence of imports, exogenous demand variables Y will serve as instruments, since they are correlated with the endogenous variable but uncorrelated with the error ε^s. This is clear from the exogenous fluctuations in demand pictured in Figure 2-2. Specification (2.3) is then estimated by IV or GMM and the identifying assumption is

$$E(Y'\varepsilon^s) = 0$$

The problem with the standard methodology arises in the presence of potential imports (entry), since to the extent that imports constrain market outcomes, fluctuations in the demand curve will be correlated with the error in the specified pricing equation. Due to the price ceiling set by imports, the true model – the data generating process – is given by[15]

$$p = \min\left(-\theta q \frac{\partial p(q)}{\partial q} + c + \varepsilon^s, c_I\right) \tag{2.4}$$

The standard pricing equation which is taken to the data – the estimated model – is, however:

$$p = -\theta q \frac{\partial p(q)}{\partial q} + c + \xi^s \tag{2.5}$$

where the (mis)specified pricing equation error is denoted ξ^s. The theoretical specification (2.1) that underlies the estimated model (2.5) fails to adequately capture the supply decisions (2.4) of an industry with pricing power facing the threat of high-cost imports. This is summarised in the following proposition.

[15]It is clear from (2.4) that, ceteris paribus, the likelihood that the imports constraint binds and thus $p = c_I$ is higher (i) the more collusive is conduct, i.e. the higher is θ; (ii) the steeper is the demand curve, i.e. the higher is $-q(\partial p(q)/\partial q)$; (iii) the higher is the domestic industry's marginal cost c; and (iv) the lower is the marginal cost of imports c_I.

Proposition 2.1 *(Non-identification of conduct) When the threat of entry constrains prices set by an industry with market power, the residual ξ^s in the standard pricing equation is negatively correlated with the excluded exogenous demand variables Y:*

$$E(Y'\xi^s) < 0$$

Consequently, IV (or GMM) estimation using demand perturbations Y will yield inconsistent estimates of conduct and cost. In particular, the true degree of market power θ will be underestimated.

Proof. Let $X_1 := -q\frac{\partial p(q)}{\partial q}$. From (2.4) and (2.5), the DGP can be rewritten as $p = \min(\theta X_1 + c + \varepsilon^s, c_I)$ and the estimated model is $p = \theta X_1 + c + \xi^s$. We wish to determine $E(Y'\xi^s)$. The error of the estimated model is

$$\begin{aligned} \xi^s &= \varepsilon^s \mathbf{1}[\theta X_1 + c + \varepsilon^s < c_I] + (c_I - \theta X_1 - c)(1 - \mathbf{1}[\theta X_1 + c + \varepsilon^s < c_I]) \quad (2.6) \\ &= \varepsilon^s \chi + (c_I - \theta X_1 - c)(1 - \chi) \end{aligned}$$

where the indicator function $\chi := \mathbf{1}[\varepsilon^s < c_I - \theta X_1 - c] = 1$ when the market equilibrium is unconstrained by the threat of entry (imports have no bite) and $\chi = 0$ when the equilibrium is constrained. (It is clear from (2.6) that the data generating process is a generalisation of the static model considered by Bresnahan (1982); this static model would correspond to a specific situation where $\chi = 1$ for all market outcomes, i.e. market outcomes are never constrained.) Assuming that the unobserved supply shock ε^s is orthogonal to the excluded exogenous demand variables Y, $E(Y'\varepsilon^s) = 0$, one may write

$$\begin{aligned} E(Y'\xi^s) &= E\left(Y'\varepsilon^s \chi + Y'(c_I - \theta X_1 - c)(1 - \chi)\right) \\ &\leq E\left(Y'\varepsilon^s \chi + Y'\varepsilon^s (1 - \chi)\right) = E\left(Y'\varepsilon^s\right) = 0 \end{aligned}$$

where the inequality follows from the fact that $1 - \chi = 1$ when $\varepsilon^s \geq c_I - \theta X_1 - c$ (i.e. when the equilibrium is constrained) and $1 - \chi = 0$ otherwise, along with the assumption that $Y > 0$.

Further, let marginal cost be linear in X_2, where X_2 is an $N \times (K - 1)$ matrix of observed variables, both exogenous (such as factor prices, including a constant) and endogenous (such as quantity): $c = X_2\beta_2$. Group the regressors of the estimated model into an $N \times K$ matrix, $X := (X_1, X_2)$, and the parameters to be estimated into a $K \times 1$ vector $\beta := (\theta, \beta_2)$. The estimated model is then $p = X\beta + \xi^s$. Denote as Z the matrix of instruments, containing the exogenous elements of X_2 and the excluded exogenous demand variables Y, and assume the rank condition for identification holds. The 2SLS estimator is given by

$$\begin{aligned}\hat{\beta} &= (X'Z(Z'Z)^{-1}Z'X)^{-1}X'Z(Z'Z)^{-1}Z'p \\ &= \beta + (\frac{1}{N}X'Z(\frac{1}{N}Z'Z)^{-1}\frac{1}{N}Z'X)^{-1}\frac{1}{N}X'Z(\frac{1}{N}Z'Z)^{-1}\frac{1}{N}Z'\xi^s\end{aligned}$$

Noting that (i) $E(X'Z)$ and $E(Z'Z)$ are positive definite, and (ii) $E(Z'\xi^s)$ contains either 0 or negative elements (since $E(Y'\xi^s) < 0$), the application of the law of large numbers to each term along with Slutsky's theorem yields

$$\text{plim } \hat{\beta} < \beta$$

In particular, plim $\hat{\theta} < \theta$. ■

The failure of the orthogonality condition can readily be seen in the linear demand example of Figure 2-3, as I show next.

Example: Shifts and rotations with linear demand Begin by considering a shift in the demand curve as depicted in the left panel of Figure 2-3. Say the inverse linear demand curve $p = a - bq$ shifts outward to $p = a' - bq$, where $a' - a = da > 0$. Recall that under both alternative hypotheses of conduct – low-cost cartel constrained by imports, and high-cost competitive industry – the equilibrium shifts from E_1 to E_2, where $dp = 0$ and $dq = \frac{da}{b}$. Plugging this into the total derivative of the static pricing equation (2.5)[16] and noting that the demand slope $\frac{\partial p(q)}{\partial q}$ remains unchanged at $-b$, one obtains $0 = -\theta\left((-b)(\frac{da}{b}) + 0\right) + d\xi^s$. Thus

$$d\xi^s = -\theta da$$

[16]This may be written $dp = -\theta\left(\frac{\partial p(q)}{\partial q}dq + d(\frac{\partial p(q)}{\partial q})q\right) + d\xi^s$, considering θ and c are constant.

from which it is clear that shifts in the demand curve are correlated with the error in the pricing equation (unless, of course, there is competition: $\theta = 0$) . Now consider a rotation in the demand curve around E_1 (right panel of Figure 2-3). Say the inverse demand curve $p = a - bq$ rotates anticlockwise around $E_1 = (q_1, p_1)$ to $p = a' - b'q$, where $b' - b = db < 0$ and thus $a' - a = da = db.q_1 < 0$. Under both alternative hypotheses of conduct, the equilibrium remains stationed at E_1, and thus $dp = dq = 0$. Plugging this into the total derivative of the static pricing equation and noting that the change in the demand slope $d(\frac{\partial p(q)}{\partial q}) = -db$, one obtains $0 = -\theta\,(0 + (-db)q_1) + d\xi^s$. Recalling that $da = q_1 db$, this translates into

$$d\xi^s = -\theta q_1 db = -\theta da$$

so that rotations in the demand curve around the equilibrium are correlated with the error in the pricing equation.

The limit-price model considered in this chapter provides an example where joint estimation of conduct and costs from a static pricing equation will perform poorly. Another example is provided by Corts (1999) who considers a dynamic model of collusion in a simple linear-demand, homogeneous-good oligopoly with symmetric and flat marginal costs, in which punishment is characterised by Cournot behaviour forever[17]. He simulates market outcomes according to varying assumptions on the persistence of exogenous demand shocks and then shows that estimation of a static pricing equation will in many instances underestimate the degree of market power, as measured by the elasticity-adjusted Lerner index (2.2). The thrust of his argument is that while the estimated conduct parameter is determined by the *marginal* responsiveness of equilibrium quantity (and thus price and the mark-up) to exogenous perturbations of demand, market power is defined by the *level* of the price-cost margin. The estimated conduct parameter will accurately capture market power "only if the true process underlying the observed equilibrium generates behaviour that is identical *on the margin*, and not just *on average*, to a conjectural variations game" (p. 234; by a "conjectural variations" model the

[17]Corts' (1999) illustration is reminiscent of the Rotemberg and Saloner (1986) supergame model of collusion with stochastic demand.

author means a model nested in (2.1) – see my footnote 13). This is clearly not the case for the imports-constrained oligopoly just outlined. As the above discussion makes clear, another way of putting Corts' argument is by stating that the errors in the pricing equation (2.5) are correlated with the demand shocks Y typically used as instruments.

The implication of Proposition 2.1 for empirical work is clear. Consider an industry where firms have market power ($\theta > 0$) and the threat of high-cost imports constrains prices in equilibrium for at least a subset of the data. (In the notation of the proof of Proposition 2.1, this corresponds to $\Pr(\chi = 0) = \Pr(\varepsilon^s \geq c_I - \theta X_1 - c) > 0$ in the available sample.) Suppose a researcher, observing the negligible penetration of imports in equilibrium, fails to realise the price-restraining effect of imports and runs specification (2.5) on the data, thinking that the data generating process is (2.3), when it is actually (2.4)[18]. Thinking that he is imposing $E(Y'\varepsilon^s) = 0$, when in fact he is incorrectly imposing $E(Y'\xi^s) = 0$, the researcher would obtain inconsistent estimates of conduct and cost. The estimated conduct parameter $\hat{\theta}$ will lie below the "true" value $\theta = \eta(p)\frac{p-c}{p}$, as defined in (2.2), underestimating the degree of market power. Intuitively, since prices do not respond to demand shocks as seen above, the coefficient on $-q\frac{\partial p(q)}{\partial q}$ will be biased toward zero. The extent to which potential imports (and entry in general) constrain market outcomes is an empirical question. I briefly illustrate the relevance of the restraining effect of imports by reference to a seminal study of the US sugar industry.

US sugar industry (Genesove and Mullin 1998): constrained market outcomes? Genesove and Mullin (1998) examine the US sugar industry at the turn of the 20th century to test the estimation of cost and conduct using the standard

[18] Note that were the researcher *aware* of the price-restraining effect of imports on a subset of the data, *and* were able to "separate the wheat from the chaff", he could implement the standard methodology using the unconstrained outcomes only (i.e. where $\chi = 1$), or in principle he could use switching regression techniques. Typically (i) the constraining effect of potential entry may be overlooked since entry is not observed in equilibrium; (ii) the level at which the constraint binds fluctuates and is unobserved (e.g. in the example of footnote 8, the level at which the antitrust *bark* has *bite* is by no means clear or stable over time); and (iii) though this is currently work in progress, it would seem that a necessary condition for identification using switching regression is the observation of independent variation in the exogenous variables, such as the marginal cost of imports and the marginal cost of the domestic industry.

static pricing equation (2.3). Thanks to the simple production technology of the industry, marginal cost is observed and can be used to check the performance of the estimation methodology, which they find "performs reasonably well in estimating θ" (p. 370). However, though the difference is small, the estimated conduct parameter $\hat{\theta}$ is lower than the direct measure of market power θ obtained from (2.2). Interestingly, the authors state that this *direct measure* of market power θ would "suggest a more competitive environment than one would expect from an industry that averaged six firms and whose largest firm had an average market share of 63%" (p. 367), and that the "likely explanation is that industry pricing was constrained by threats of (domestic) entry or of foreign imports" (p. 367), despite "very little" sugar actually being imported into the US[19]. Genesove and Mullin point out that: "Although we acknowledge the influence of these competitive fringes, they are not formally incorporated into our analysis" (p. 359). This section's analysis suggests that to the extent that market outcomes in the sugar industry were constrained by the threat of entry, this would lead to a downward bias in the estimated degree of market power.

In Section 2.4 I develop an example using the Brazilian cement industry where market outcomes are constrained by the threat of imports. I illustrate how poorly the standard methodology may perform in such a setting: the estimated conduct parameter heavily understates the direct measure of market power (based on a direct measure of marginal cost). I then use firm-level data to delve deeper into the pattern of conduct in the industry. To this end, we need to consider other models of conduct in the imports-constrained oligopoly model, in addition to the model of monopoly seen at the beginning of this section.

[19] Further evidence of the price-constraining effect of imports is provided: "Although very little refined sugar was ever imported into the United States, in the early years of the Sugar Trust (the largest firm) the threat of European imports affected U.S. prices. In 1888 and 1894, Havemeyer (the Sugar Trust's president) acknowledged setting the price of refined sugar so that none would be imported from Europe" (p. 358; parentheses added). Note that the authors' low direct measure of market power θ stems from a low observed price-cost mark-up $\frac{p-c}{p}$ and a moderate elasticity $\eta(p)$ (of around -1.05 for most part of the year).

2.2.3 From monopoly to oligopoly

The next question is: what are the equilibrium outcomes when there is a domestic oligopoly comprised of $n \geq 2$ firms, instead of a monopoly? Clearly this will depend on the conduct of the domestic firms facing the competitive fringe of imports. I will consider the non-trivial case of imports restraining prices $p^M \geq c_I$ (right panel of Figure 2-1), since the complementary case where imports do not restrain even a monopolist $p^M < c_I$ is standard (left panel of Figure 2-1). I now consider the benchmark models of collusion, Bertrand and Cournot.

Under the most collusive outcome, the equilibrium price is $p = c_I$ and the oligopoly's joint output is $q = p^{-1}(c_I)$, the same output as that of a monopolist (recall Section 2.2.1). In a non-cooperative framework with heterogeneous firms, where side payments are not allowed, output will need to be shared among the firms according to some rule or historical pattern[20]. In a spatial context, where the oligopoly consists of firms with multiple plants scattered across space, meeting each other in different local markets, one possibility is to have the most efficient firm in a given local market supply a large share of output. The most efficient (lower cost) firm in a local market could be the firm with the plant located closest to that market, thus incurring lower transport costs. To the extent that firms' plant configurations are "sufficiently" symmetric, with different firms being the low-cost producer in different markets, the restriction of no side payments can be circumvented and aggregate industry profits can be increased[21]. I return to collusion under multimarket contact (as it applies to the Brazilian cement industry) in Section 3.2[22].

[20] As before, assume flat marginal costs. With homogeneous firms, sharing output equally among the firms – as well as any alternative allocation – maximises joint profits. With heterogeneous firms, were side payments allowed, the optimal allocation rule is to have the low cost firm supply the entire market.

[21] Bernheim and Whinston (1990) explore the incentive constraints under multimarket contact. By pooling a firm's incentive constraints across markets, its share in those markets where it enjoys a low cost (i.e. "on its own turf") may be increased at the expense of its share in markets where it has a high cost. See Section 3.2.

[22] Notice that the price limit set by the delivered cost of imports c_I may provide a natural focal price for coordination. (I thank Margaret Slade for pointing out that this observation is consistent with the Eastman-Stykolt (1966) hypothesis, where by providing the domestic industry with a collusive focal point equal to the world price plus tariff, protection may facilitate oligopolistic coordination of the protected firms. See Harris (1984).)

Under the other polar model of Bertrand competition, the equilibrium price is equal to the marginal cost of the second most efficient plant. Label firms 1 and 2 as the lowest-cost firm and the next lowest cost firm respectively, i.e. $c_1 \leq c_2 (< c_I)$. The equilibrium price is then $p = c_2$, with firm 1 supplying the entire market with $q_1 = p^{-1}(c_2)$. This situation is similar to that depicted in the right panel of Figure 2-1, where it is the next lowest cost firm 2 rather than imports that restrains prices.

Consider finally the standard case of Cournot behaviour among the n firms in the domestic oligopoly[23]. Consider the output decision of firm f. In the absence of imports, denote firm f's reaction function $q_f = R_f(q_{-f})$, where $q_{-f} := \sum_{j \neq f} q_j$ is the joint output of its (domestic) rivals. (This reaction function is derived from the firm's Cournot first-order condition and is drawn as the steeper line in the left panel of Figure 2-4.) In the presence of imports, imports occur if $p(q_f + q_{-f}) > c_I$, or equivalently if $q_f + q_{-f} < p^{-1}(c_I)$, i.e. if domestic output is restricted to fall short of the quantity level at which the marginal cost of imports crosses the demand curve. In this case, where $q_f + q_{-f} < p^{-1}(c_I)$, the quantity of imports is positive and equal to $q_{imports} = p^{-1}(c_I) - q_f - q_{-f}$, so that total supply is $q_f + q_{-f} + q_{imports} = p^{-1}(c_I)$. Thus $q_f + q_{-f} \geq p^{-1}(c_I)$ defines the "imports constraint": its boundary is drawn as the less steep line, of slope -1, in the left panel of Figure 2-4. Clearly, the perfectly-elastic supply of imports ensures that, given the joint output of its rivals q_{-f}, Cournot firm f will set its output such that price is at most equal to the marginal cost of imports, such that imports do not occur. Hence, in the presence of imports, firm f's best response to the joint output of its rivals q_{-f} will correspond to the outer envelope to its reaction function in the absence of imports, $R_f(q_{-f})$, and the boundary to the imports constraint, $q_f + q_{-f} = p^{-1}(c_I)$; denote this "constrained" reaction function as

$$q_f = \tilde{R}_f(q_{-f}; c_I) := \max(R_f(q_{-f}), p^{-1}(c_I) - q_{-f})$$

$\tilde{R}_f(q_{-f}; c_I)$ is illustrated in the left panel of Figure 2-4 as the thick curve. Notice that when $R_f(0) < p^{-1}(c_I)$, as drawn, $R_f(q_{-f})$ will cross $q_f + q_{-f} = p^{-1}(c_I)$. For

[23]For a model with a similar flavour where a Cournot oligopoly may deter entry by producing the limit output, see Gilbert and Vives (1986) (I thank Xavier Vives for pointing this out to me).

high enough q_{-f} such that $\tilde{R}_f(q_{-f}; c_I) + q_{-f} \geq p^{-1}(c_I)$, firm f's optimal reply in the presence of imports is to set the same quantity that it would set in the absence of imports (and the corresponding market price is lower than c_I). This steeper upper segment of firm f's constrained reaction function $\tilde{R}_f(q_{-f}; c_I)$ is collinear with the reaction function in the absence of imports $R_f(q_{-f})$, and the standard Cournot pricing equation holds:

$$p(q) + \frac{p(q)}{\eta(q)} \frac{q_f}{q} = c_f \tag{2.7}$$

where as before $\eta(q)$ is the market price elasticity of demand and $q = q_f + q_{-f}$. Now, for lower q_{-f} such that $\tilde{R}_f(q_{-f}; c_I) + q_{-f} = p^{-1}(c_I)$ (i.e. $\tilde{R}_f(q_{-f}; c_I) \geq R_f(q_{-f})$), firm f's optimal reply in the presence of imports *exceeds* the quantity that it would set in the absence of imports, and price equals c_I (since otherwise imports would occur and price would still be equal to c_I). Here, along the flatter segment of firm f's constrained reaction function $\tilde{R}_f(q_{-f}; c_I)$, firm f's (perceived) marginal revenue falls short of marginal cost:

$$p(q) + \frac{p(q)}{\eta(q)} \frac{q_f}{q} < c_f \tag{2.8}$$

Conditions (2.7) and (2.8) combine to prove the following proposition[24]:

Proposition 2.2 *("Constrained" Cournot first-order condition) In the presence of imports, if firm f behaves as a Cournot player, it will be the case that*

$$p(q) + \frac{p(q)}{\eta(q)} \frac{q_f}{q} \leq c_f \tag{2.9}$$

This condition holds as a strict inequality when the "imports constraint" $q_f + q_{-f} \geq p^{-1}(c_I)$ binds, in which case price is equal to the marginal cost of imports c_I.

The set of Cournot equilibria is found by similarly deriving the rival firms' constrained joint reaction function $q_{-f} = \tilde{R}_{-f}(q_f; c_I)$, which is again the outer envelope of the joint reaction function in the absence of imports $q_{-f} = R_{-f}(q_f)$ and

[24]Condition (2.9) also holds as an inequality in the case of a corner solution (i.e. $p(q_{-f}) < c_f$ such that $q_f = R_f(q_{-f}) = 0$), but this is standard so is omitted from the proposition.

the boundary to the imports constraint, $q_f + q_{-f} = p^{-1}(c_I)$. The set of equilibria is the intersection of $q_f = \tilde{R}_f(q_{-f}; c_I)$ and $q_{-f} = \tilde{R}_{-f}(q_f; c_I)$. For a low enough cost of imports (i.e. an imports boundary sufficiently far from the origin, as drawn in the right panel of Figure 2-4 for the next illustration), there are multiple equilibria and imports restrain prices at the Cournot equilibrium[25].

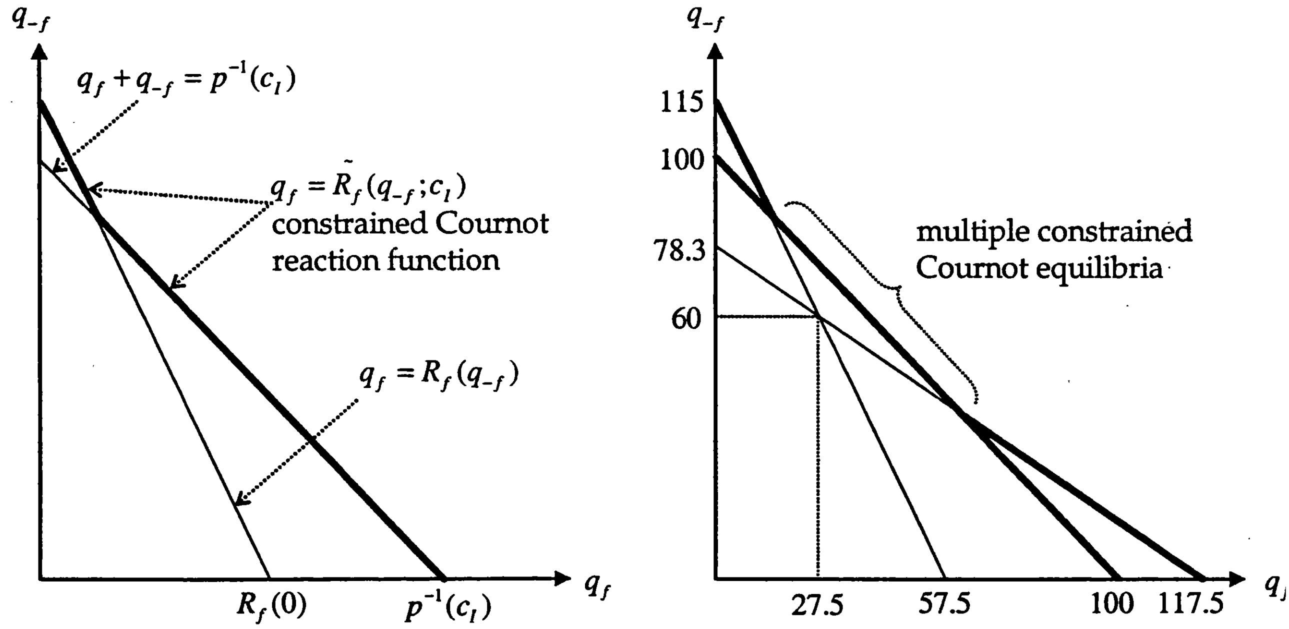


Figure 2-4: Cournot oligopoly facing a competitive fringe of imports. Left panel: Cournot firm f's reaction function, facing domestic rivals and imports. Right panel: Cournot equilibria. Drawn for linear demand and $c_I = 6$ as in the right panel of Figure 2-1.

An illustration 2 This is illustrated in the right panel of Figure 2-4, drawn assuming the linear demand function $p = 16 - \frac{1}{10}q$ and $c_I = 6$ of the earlier illustration to the domestic monopoly case (see the right panel of Figure 2-1), and adding the assumption that there are $n = 3$ firms, that the marginal cost of the firm of interest f is $c_f = 4.5$ and that $\sum_{j \neq f} c_j = 8.5$.[26] In the absence

[25] In the absence of imports, the unique Cournot equilibrium outcome (q_f^C, q_{-f}^C) is defined implicitly by $q_f^C = R_f(R_{-f}(q_f^C))$ and $q_{-f}^C = R_{-f}(R_f(q_{-f}^C))$. Formally, imports have bite under Cournot conduct if $p(q_f^C + q_{-f}^C) > c_I$, or equivalently when $q_f^C + q_{-f}^C < p^{-1}(c_I)$.

[26] For example, if one of firm f's rivals has the same marginal cost as the earlier monopolist, of 4, then the remaining rival has a marginal cost equal to that of firm f, of 4.5.

of imports (or were $c_I > p(87.5) = 7.25$ instead of 6), joint output and price in the Cournot oligopoly would respectively be 87.5 and 7.25 in equilibrium. In the presence of imports, where $c_I = 6$, joint output and price in the constrained Cournot equilibrium are respectively 100 and 6. This joint outcome is the *same* as the most collusive outcome of Illustration 1 in Section 2.2.1.

2.2.4 From theory to application: inference of conduct in the presence of potential imports (when costs are observed)

I argued in Section 2.2.2 that, in the presence of potential imports, estimation of a static pricing equation imposing the regular moment conditions yields inconsistent estimates of conduct and cost parameters. The immediate question is then: But what if the researcher *observes (domestic firms') marginal cost*? In a setting where imports restrain prices at the industry equilibrium, how can one then identify firm conduct?

Clearly, a direct comparison of marginal cost to price will provide a test of competitive behaviour against less competitive models of firm behaviour (and where imports may restrain prices). However, consider the following hypothetical situation. Suppose that prices exceed observed costs and that imports restrain prices were firms to behave à la Cournot, let alone restrain prices were firms to alternatively engage in collusion. One wishes to identify, from the constrained prices and quantities and the observed costs, the underlying model of conduct in the industry. It is not obvious how one can distinguish, say, collusive conduct from Cournot conduct when the equilibrium price under either alternative model of conduct is constrained to be the same and equal to the marginal cost of imports. This situation, for a domestic duopoly, is pictured in the left panel of Figure 2-5. The observed equilibrium outcome is marked with a "+", where clearly the imports constraint binds (since "+" lies on $q_f + q_g = p^{-1}(c_I)$, where the duopolists are labelled f and g). From the observed constrained equilibrium, it is not possible to tell whether firms in the industry behave in Cournot fashion (in which case aggregate equilibrium output in the absence of imports would equal $q_f^C + q_g^C$) or whether firm behaviour is more collusive than the Cournot benchmark (in which case the aggregate equilibrium output in the absence of imports would be lower than $q_f^C + q_g^C$). The example provided by Illustrations 1 (monopoly) and 2 (Cournot) above should clarify. In the example, the cost of imports, in addition to demand and domestic cost conditions, are picked to be such that under either model of conduct the aggregate equilibrium outcome is constrained to be the *same*. Under either the hypothesis of collusion or firms behaving à la Cournot, aggregate industry output is 100 (imports are zero) and price is 6 (equal to the marginal cost of imports). Thus, industry outcomes under full collusion and under Cournot are observationally equivalent in this example in which $p(q^C) > c_I$, where q^C is the "unconstrained" Cournot equilibrium industry output (recall $q^C > q^M$)[27].

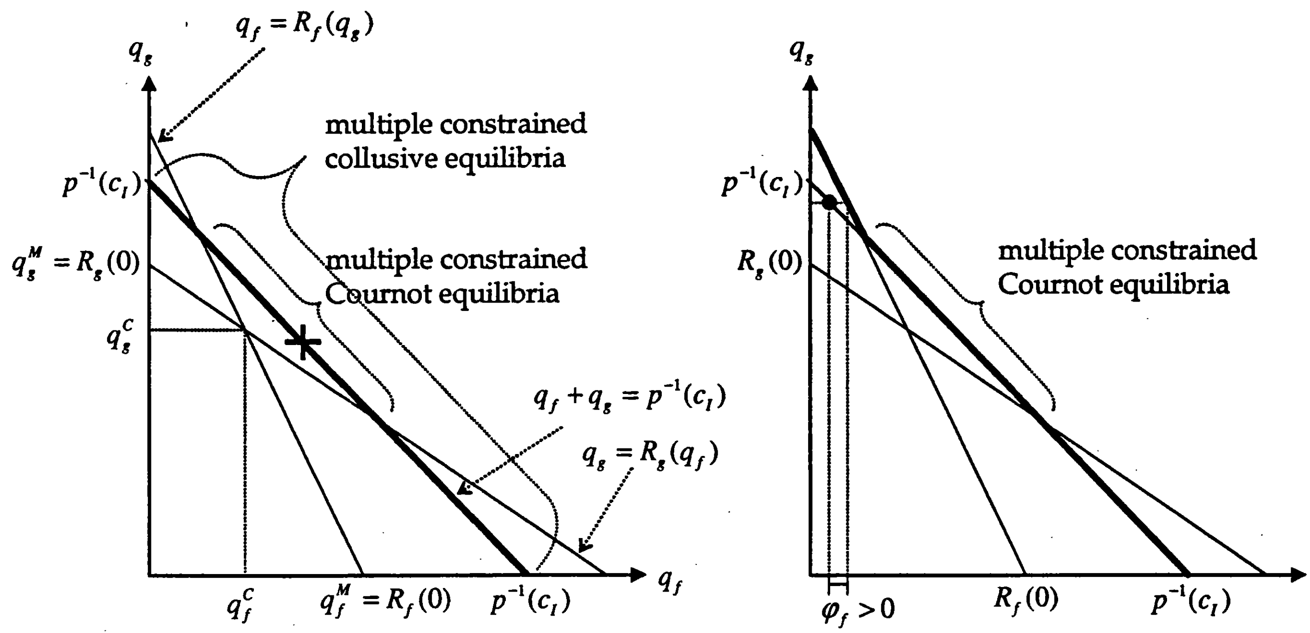


Figure 2-5: Identifying collusion from Cournot when imports constrain equilibrium prices under both models of conduct. Left panel: The imports constraint binds at the industry equilibrium marked "+", which is consistent with either Cournot conduct or more collusive firm conduct. Right panel: Rejection of Cournot behaviour for firm f: $\varphi_f > 0$.

Yet even in the case illustrated – where imports constrain prices not only when

[27]Had I taken $(p^M >)c_I > p(q^C)$ in the example, imports would not have had bite under Cournot conduct and thus outcomes under full collusion and Cournot would be distinct.

behaviour is collusive but also under Cournot – it may be possible to distinguish between competing hypothesis of conduct, despite industry outcomes being equivalent, *if firm-level quantity data is available* (in addition to costs being observed). The point is to recognise that for a Cournot firm, the general (i.e. allowing for the constraining effect of imports) pricing condition (2.9) of Proposition 2.2 has to hold. That is, for no Cournot firm can (perceived) marginal revenue *exceed* marginal cost, otherwise the firm would optimally expand output, and this holds irrespective of whether the imports constraint binds or not (since the latter places a *lower* bound on aggregate domestic output). This pricing condition can be used with observed marginal cost and the estimated market price elasticity of demand to test for Cournot behaviour. Under the hypothesis of Cournot behaviour, one may well observe a firm choosing output to the right of its unconstrained reaction function, given the (joint) output of its rivals, since imports may be restraining price at the constrained Cournot equilibrium: here, the Cournot firm would not cut output in an attempt to raise price above c_I, as an unconstrained Cournot firm would do, since this would only open the door to imports. However, under the hypothesis of Cournot behaviour, one should not observe a firm choosing output to the *left* of its unconstrained reaction function. This is illustrated in the right panel of Figure 2-5. While firm g's behaviour is consistent with Cournot (it does not cut output as a Cournot firm would do in an unconstrained equilibrium, since the imports constraint is binding), firm f's behaviour is *not* consistent with Cournot: firm f is restricting output. This translates into the following test. Rewrite the Cournot pricing condition (2.9) as an equality:

$$p(q) + \frac{p(q)}{\eta(q)}\frac{q_f}{q} = \varphi_f + c_f \tag{2.10}$$

Proposition 2.3 *(Sufficient statistic to reject Cournot behaviour) Under the null of Cournot behaviour, $\varphi_f \leq 0$. When the imports constraint binds, $\varphi_f < 0$ is consistent with Cournot behaviour. The finding that $\varphi_f > 0$ allows one to reject the hypothesis that firm f is behaving in Cournot fashion, in favour of more collusive behaviour, regardless of whether the imports constraint binds or not.*

As I will argue, there is strong evidence to suggest that outcomes across the different local markets in the Brazilian cement industry can be characterised as follows. The market price elasticities of demand are estimated to be very low, of the order of -0.5. Demand across local markets with widely differing market structures (i.e. concentration indices or number of firms) is consistently inelastic: elastically-supplied imports appear to restrain prices. Yet consumer prices considerably exceed the marginal cost of even the least efficient producers serving a given local market. I observe many instances in the data where the hypothesis of Cournot behaviour on the part of firms in local markets can be rejected in favour of what appears to be (tacit) market division. The right panel of Figure 2-5 depicts a typical local market: a firm with a plant located within this market will correspond to firm g. It ships more than what it would ship in any constrained or unconstrained Cournot equilibrium. Other firms with plants located within this given local market, or with plants located nearby, will correspond to firm f in Figure 2-5, whose shipments to this local market fall short of their Cournot best responses. (In Chapter 3 I show how such a collusive arrangement may be sustained in equilibrium in a spatial dynamic model where firms meet in different markets.)

2.3 Industry and data

2.3.1 The cement industry

Cement is a homogeneous good produced largely from limestone and clay in weight proportion of roughly 5 to 1. Described simply, limestone and clay are ground and the mixture is burned at a very high temperature in a rotary kiln producing cement clinker. The clinker pellets – once cooled – are then ground and mixed with a retarding agent (gypsum) and varying types of additives to form different formulations of cement[28]. Despite the relative simplicity of the product, the production of

[28]The different formulations of cement are substitutes in most types of user applications. While clinker comprises around 96% of ordinary cement, this proportion can be considerably reduced in other formulations, such as (blast furnace) slag cement or pozzolanic cement. Usually the supply of these different formulations will depend on the availability of additives (i.e. slag or pozzolane) in the proximity of the cement plant, such as a steelworks in the case of slag cement. Each

cement is capital intensive and is characterised by substantial economies of scale. Labour basically performs a supervisory role (Norman 1979). The process is also energy intensive, not only due to the operation of the kiln but also due to the grinding of raw material and clinker[29].

The process exhibits a fixed factor production function since factor inputs are not substitutable. Yet marginal costs do vary across kilns and plants, according to the technology, capacity, age and fuel employed (Jans and Rosenbaum 1996). The last major innovation to the production process took place in the 1970s in response to the energy price shocks. The "wet" process kiln system was replaced by the "dry" process, which consumes less than half the respective energy (since no heat is needed to evaporate water). With the energy crisis in the foreground, firms invested in bigger, more energy-efficient kilns. Maximum kiln capacity in the four decades leading up to 2000 has increased six-fold to four million tonnes per annum (mtpa) (World Cement 2000)[30]. Although equipment suppliers and cement producers work closely together, most innovations seem to originate from the equipment suppliers, and technology can be purchased off-the-shelf[31].

As cement is a low-value commodity relative to weight, transportation costs may assume a significant proportion of cost, leading to geographically segmented markets. Scherer et al (1975, p. 429) list cement as having the second highest freight cost index for shipments out of 101 US industries. In order to meet dispersed demand, firms may trade in (production) scale economies for lower transport costs by scattering their plants across markets[32].

---

type of cement usually needs to conform to legislation that specifies its (physical and chemical) properties. Thus differentiation based on formulation is limited.

[29] In many regions, such as in the Americas and Europe, the supply of limestone is ubiquitous; the raw material is thus usually extracted from a quarry located within the plant complex. The setup of a modern plant with capacity of 1.5 million tonnes per annum (mtpa), including the prospecting rights over limestone reserves, can require a capital outlay of up to US$300 million (US$200 per tonne of capacity).

[30] See Rosenbaum (1989) and Johnson and Parkman (1983) on process and capacity changes in the US industry.

[31] "Turn-key" plants may be ordered from suppliers. Research and development (R&D) spending by the cement producers themselves is limited: operating at the forefront of cement-production technology, the Japanese producer Taiheiyo (Chichibu Onoda prior to 1998) spends less than 1% of sales revenue on R&D.

[32] See Scherer et al (1975) and Newmark (1998). Pre-empting entry may further reduce initial plant scale (Johnson and Parkman 1983).

Demand for cement is essentially driven by the construction industry and is, similarly, cyclical. In developed markets, shipments are largely made in bulk to ready-mixed concrete firms and construction firms. By contrast, the lion's share of the industry's production in developing countries is dispatched in bags to resellers (retailers) who sell on to individuals ("do-it-yourself buyers"), reinforced by the fact that over the past decade or two many governments in such markets have been scaling down on infrastructure investments. The demand curve for cement is typically steep since cement makes up only a moderate part of most construction projects and there are few substitutes.

World demand, estimated at 1620 mt in 2000, has been growing at around 3% p.a. (International Cement Review 2001). Growth is concentrated in emerging markets while demand in North America and Western Europe has been growing slowly or is stagnant[33]. Over the past 15 years, a significant process of consolidation has been running its course in the global cement industry. While family-run and state-owned firms have been put on sale, a few multinational firms have been on a buying spree, aggressively moving into new markets or expanding in markets where they previously operated. The combined production share (excluding China) of the world's six largest firms (C_6) in 2000 was estimated at 35%, up from 23% in 1995 and 14% in 1985.

2.3.2 The Brazilian cement industry in the 1990s

On the basis of output, Brazil ranks sixth in the league of cement-producing countries, with output of approximately 40 mtpa in the period 1998 to 2000 (SNIC 2002[34]). As shown in Figure 2-6, in 1999 57 active plants were scattered across a geographic area slightly smaller than that of the US[35]. This spatial distribution is not even, however, as consumer markets and thus plants are concentrated along the coastal states, in particular the relatively wealthy and populated states in the

[33] Around 30% of consumption occurs in China, notoriously a producer of low-quality cement in energy-inefficient, environmentally-unfriendly "backyard" mini cement plants.

[34] Unless specified otherwise, facts from this section are drawn from reports of the Brazilian cement industry's trade association (SNIC), backed up by other sources. See Appendix B.

[35] With a population corresponding to two-thirds that of the US, cement consumption per capita in Brazil amounts to 232 kg as compared to 415 kg in the US (SNIC 2002).

Southeast and South regions of the country[36]. States to the northwest of the centre of the country are sparsely populated and are largely covered with jungle.

In 1999, as also depicted in Figure 2-6, these 57 plants were owned by 12 firms. The two largest firms, Votorantim and Grupo João Santos, respectively with nationwide shipment shares of 41% and 12% in 1999, were both domestically-owned, traditional family-run businesses. The subsidiaries of the large multinational firms Holcim and Lafarge followed, with shipment shares of 9% and 8% respectively. As Figure 2-7 indicates, this national picture hides a lot of variation at the local, statewide level.

The 1990s saw two distinct periods in the history of the Brazilian cement industry. Up until mid 1994, a period of very high inflation and low macroeconomic growth, cement consumption was stagnant at around 25 mtpa. With the successful implementation of the *Real* economic stabilisation plan in July 1994 (see below), cement consumption resumed its growth at a rate of 10% p.a., reaching 40 mtpa by 1998-99, pulled by exogenous growth in the construction sector[37]. The post-stabilisation phase of the 1990s also saw a flurry of acquisition activity in the cement industry, with the expansion of incumbents and the entry of foreign firms which did not previously own assets in Brazil. Compared to the 12 firms that ran operations by 1999, the industry had consisted of 19 producers in 1991.

Given the short shelf life of cement, firms produce for immediate consumption. Stocks at producers amount to approximately one week of sales, with roughly another week of sales being stocked down the trade. Around 90% of shipments from producer plants to buyers in consumer markets is carried out by road – as opposed to rail or water. In line with other developing countries, as mentioned above, around 80% of volume is shipped in bags to resellers who then sell on to small-scale consumers; only 20% is shipped in bulk by the industry directly to consumers, usually ready-mixed concrete firms, large construction firms or producers of construction aggregates.

[36]The Federative Republic of Brazil is a federation of 27 states. The coastal states are those running clockwise from the north-most point of the country – the state of Amapá (AP) – to the south-most state of Rio Grande do Sul (RS).

[37]As will be discussed shortly, real cement prices also fell in the early days post stabilisation.

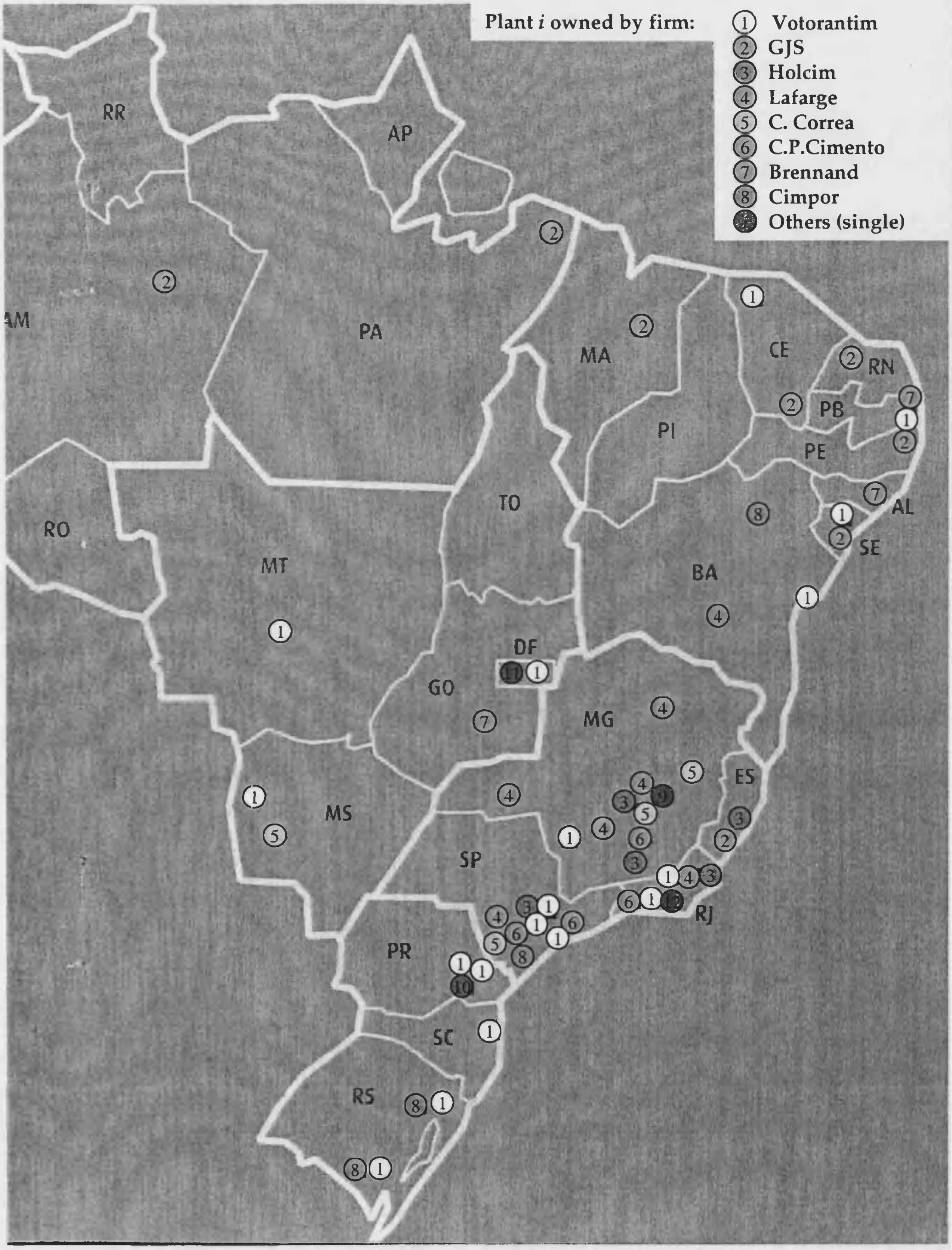


Figure 2-6: Active plants in 1999

	Mean	Standard Deviation	Maximum	Minimum	Memo: Total across 27 states
Cement consumption in state (kt)	1,483	2,324	11,723	55	40,045
Number of (active) cement plants located within state[0]	2.1	2.6	11	0	57
Number of cement firms (producers) shipping to state	5.7	2.8	11	1	12
One-firm concentration index in state[1]	57%	17%	100%	25%	41%
Two-firm concentration index in state[1]	83%	13%	100%	49%	52%
Four-firm concentration index in state[1]	97%	6%	100%	77%	70%
Hirschmann-Herfindahl index in state[1]	4494	1823	10000	1830	2106
% shipments originating from state destined for that state[2]	60%	22%	100%	14%	
% shipments origin. from state destined for that and bordering states[2]	92%	9%	100%	70%	
Value Added (volume decomposition) in Construction Sector[3]	475	726	3,431	9	12,352
Land area (x 1000 square kilometres)[4]	315	370	1,571	6	8,515
Population (m, mid 1999)[4]	6.1	7.3	35.8	0.3	163.9
Population density (/sq km)	56.9	84.1	339.5	1.2	19.3
Per capita cement consumption in state (kg p.c.)	211	67	353	104	244
Per capita Value Added in Construction Sector[3]	61	26	108	16	75

[0] Of the 57 plants, 7 were grinding-only operations (with clinker being shipped from a nearby plant with integrated facilities)
[1] Based on shipments from producers located anywhere to buyers located in a given state
[2] Applies only to states from which shipments originate (i.e. states where plants are located)
[3] In rescaled constant monetary units
[4] Source: Brazilian Institute for Geography and Statistics (IBGE)

Figure 2-7: Variation across 27 states of the Brazilian federation, Summary Statistics (time-varying figures refer to 1999)

I now provide a few comments on the 1994 economic stabilisation plan and on the role of imports in the cement industry, given their relevance to the present study.

The July 1994 stabilisation plan The *Real* economic stabilisation plan, enacted in July 1994, successfully brought (very high) inflation under control. Between 1991 and June 1994, the first period covered in this study, inflation as measured by the change in the General Price Index averaged 26% *per month* (i.e. prices doubling every quarter). With the implementation of the stabilisation plan, inflation fell to 22% per annum in the period July 1994 through December 1995, further falling to 10% per annum in the six years between 1996 and 2001.

One of the outcomes of the stabilisation plan was its large positive effect on the level of economic activity. The sharp slowdown in inflation, through the reduction in "inflationary tax", represented a reduction in the transfers from the private sector to the government. In particular, the large mass of consumers among the lower-income groups who previously had no access to instruments of monetary protection, such as price-indexed savings accounts, saw a significant rise in real incomes. Given their high propensity to consume, this boosted the demand for consumer goods – notably food, clothing and durables – and the demand for housing. Coupled with commercial construction projects resulting from a more favourable investment climate, the demand for housing led to a significant increase in the activity of the construction sector of the Brazilian economy, and thus in the demand for cement.

The four years following stabilisation also saw the appreciation of the local currency, of direct relevance to the competitiveness of cement imports. This period of a strong local currency abruptly came to an end with the devaluation of January 1999.

The role of imports in Brazil Imported cement (including the intermediate product clinker) constitutes a small share of domestic consumption. As shown in Figure 2-8, in the period 1989 to 2003, this share has amounted to at most 2-3% of consumption across Brazil, though the trend appears to be rising since the trade liberalising reforms of the early 1990s (and despite a dip in 1999 and 2000 following the devaluation of the local currency – see below). This low level stands in stark contrast to the penetration of imports in the US. Carlsson (2001) reports that "imports represent a substantial and increasing part of the market in the United States, ranging between 10 and 17 percent of domestic consumption since 1985" (p. 7). The share of imports in some coastal US markets is actually as high as 30%[38]. The presence of imports in Brazil thus pales in comparison to the US, despite most of its markets being located along (or in proximity to) an extensive Atlantic coastline.

[38]Despite the bulkiness of cement relative to its price, the development of specialised seaborne handling and transportation equipment from the 1970s enabled imports to make their presence felt in coastal markets. Dumez and Jeunemaître (2000) provide a historical account of the rise of international trade in cement. On the other hand, in both the US and Brazil, exports account for less than 1% of domestic production (though in Brazil the current trend is upwards).

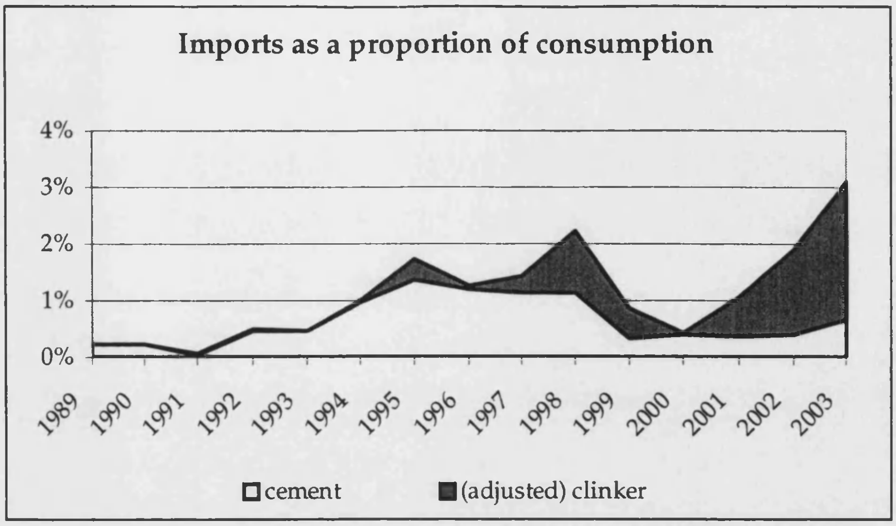


Figure 2-8: (Official) Imports of cement and clinker as a proportion of domestic consumption. Source: SECEX, MDIC. Clinker quantities are adjusted by the author to reflect usage in the production of cement (assumes 80% of clinker imports used in production of slag cement, with a 40% clinker content).

As this study finds, however, the limited penetration of imports hides their welfare-enhancing role in restraining domestic prices, curbing the market power of domestic producers. The trade liberalising reforms of the early 1990s, coupled with the appreciation of the local currency in the four years following stabilisation in mid 1994, opened the door to the threat posed by the entry of imports. To provide a flavour, Figure 2-9 depicts the evolution of cement prices in the state of Rio Grande do Sul – where one-firm and two-firm concentration ratios respectively amounted to 55% and 84% in 1999 – both in current local currency (the *real*, R$) and in a currency of foreign trade, proxied by the US dollar. Domestic cement prices in local currency are highly correlated with the price of the US dollar in local currency (i.e. the exchange rate), to the extent that despite the occurrence of large variations in the exchange rate during the period, and thus in the domestic cement price in local currency, the domestic cement price converted into US dollars is quite steady since 1995[39] [40]. In the cross-section of local markets, one would also expect cement prices to be increasing in the market's distance from the coast. This is verified to be the case. It is important to add, however, that neither of these two observations – regarding (i) the correlation of cement prices in local currency and the exchange rate, and (ii) that prices appear to be increasing in distance – are offered as proof of the claimed role of imports in restraining prices. While consistent with the claim, they are also consistent with alternative stories, such as factor prices being set in hard currency on the world market (fuel oil and diesel?), or with producers incurring higher transport costs to distribute cement in less densely populated areas. The estimation of a very low market price elasticity of demand in equilibrium, in Section 2.4.2, coupled with high price-cost margins and supported by interview evidence, will be the key element in support of my claim.

2.3.3 Data available: Plant-to-market cement flows and the construction of marginal cost

A detailed account of the sources and treatment of the data is provided in Appendix B. Here I offer a short description and briefly discuss how I compute the marginal

[39]Until January 1999 Brazil had an exchange rate fixed by the government. The local currency (the *real*) was floated in January 1999 in the midst of the "Brazil currency crisis", depreciating by 70% against the US dollar in one month, but later partially receding. Other periods of above-average exchange-rate instability took place in 2001 (commonly attributed to the Argentina crisis next door) and in the second half of 2002, with the uncertainty surrounding the outcome of the presidential election late that year. The relatively flat evolution of domestic cement prices in US dollars is consistent with imports setting a price ceiling of between 6-7 US dollars per bag of cement (this would correspond to the US-dollar equivalent of c_I, as defined in Section 2.2). The observation that it seems to take domestic producers between 6-12 months to raise domestic prices back to this ceiling in US dollars upon large unexpected devaluations in the local currency (i.e. in 1999, in 2001 and in 2002) suggests that raising domestic prices in local currency is not friction free (perhaps the industry is wary of attracting negative publicity).

[40]To provide an example, an equity analyst of an investment bank wrote that "(a)lthough imports accounted for only 1.6% of the Brazilian total consumption in 1995, reaching 451.3 thousand tons, it represents a constant threat to domestic producers, pressing down domestic prices and imposing a price ceiling of US$ 70 per ton" (Zaghen 1997; pp. 24). The author refers to the price "at the coast" as the exporter's FOB price plus international insurance and freight, excluding cost upon arrival in Brazil, such as inland freight, sales taxes and resellers' markups. Further evidence suggesting concern by domestic producers as to the threat of imports is provided by their successful lobbying of government in passing antidumping measures – namely a 23% import tariff – against Venezuelan and Mexican cement producers in the late 1990s who were starting to make inroads into local markets particularly in the north and northeast of the country.

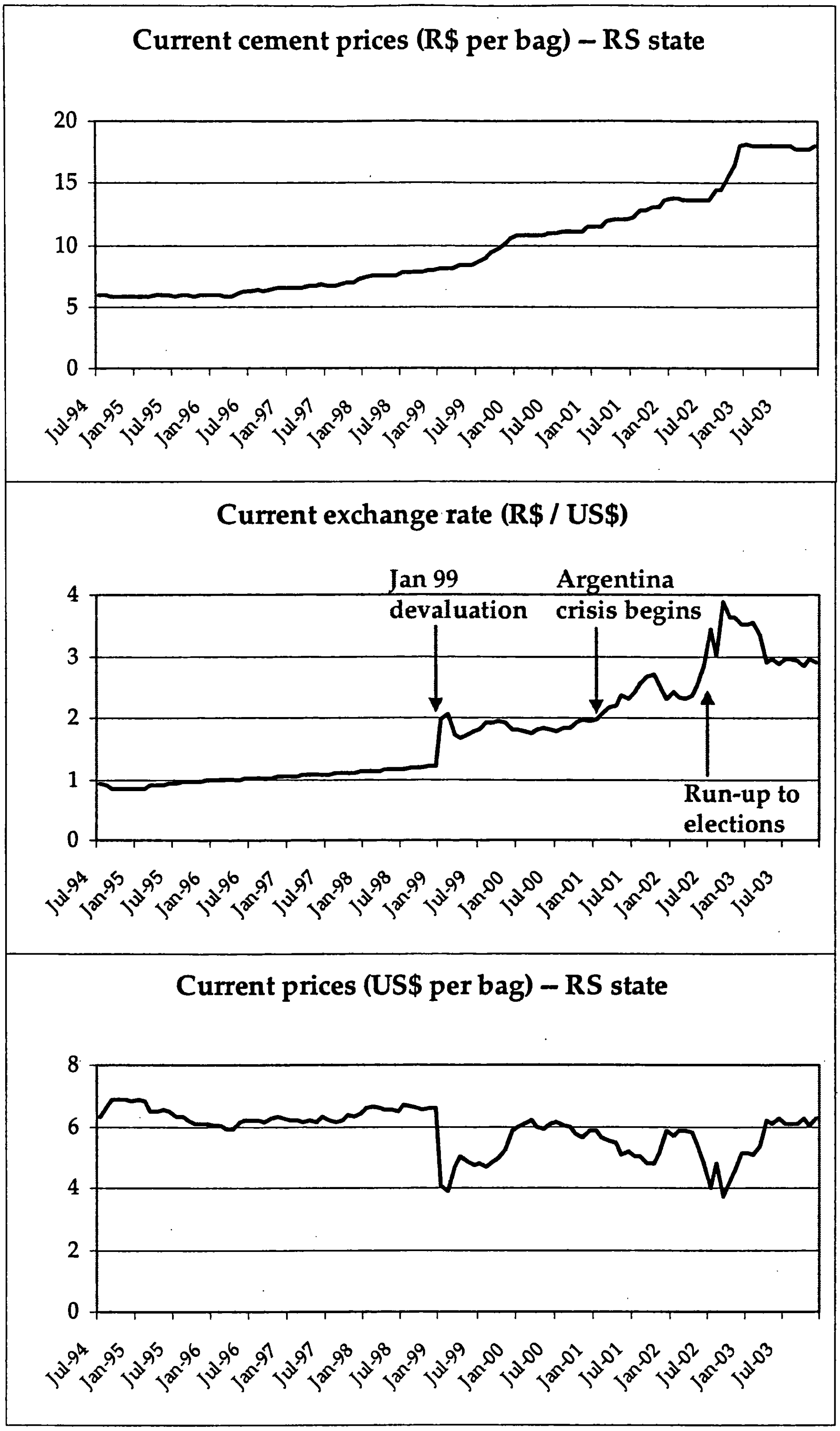


Figure 2-9: Evolution of cement prices in *RS* state since July 1994. In current local currency units (R$) per bag and US$ per bag

cost of each plant in serving each local market.

On the demand side, I observe monthly cement consumption and consumer prices (i.e. prices set by retailers, also referred to as resellers) across the 27 states in the period 1991 to 2003. I take each state to represent a local market. As demand shifters, I observe alternative series of economic activity, either in the construction and building sector or aggregated across sectors of the economy, which I use as proxies for the exogenous demand for cement[41].

The key ingredient on the supply side is the observed breakdown of shipments from each plant to each of the local markets (states), enabling me to map the flow of cement from the plant to the consumer. In addition to plant ownership, I observe plant characteristics – e.g. capacity, number of kilns, type of fuel usage, proportion of shipments in bags as opposed to bulk[42] – and local factor prices, such as fuel oil, coal, electricity and wages. I do not observe freight prices paid by cement producers but I approximate these by using data on freight prices for agricultural goods collected over the period 1997 to 2003 for thousands of different routes across Brazil. The transportation of goods such as soyabean and maize are reportedly close substitutes in the supply of cement freight (Soares and Caixeta Filho 1996).

Considering that the technology of cement production is of the fixed coefficients type, I use engineering estimates, factor prices and plant characteristics to directly calculate the marginal cost of each plant in serving each market. As I argue in Appendix B, these marginal costs are indeed estimated upper bounds to the true marginal costs. (When I turn to the testing of conduct in Section 2.4.4, such a bias, however, reinforces the results.) In view of the fixed-coefficient technology and my understanding of the industry, I model plant marginal cost as flat in quantity up to

[41] This follows from the fact that cement is an input to construction and yet accounts for a small share of construction budgets. Taking such construction activity to exogenously move the demand curve for cement is a typical assumption: see, for example, Syverson (2004) who uses construction sector employment as an exogenous measure of demand in ready-mixed concrete (an industry located downstream to cement).

[42] Aggregating across all plants, between 1997 and 1999 81% of shipments were in bags. In terms of the means of transportation, 91% of shipments were by road. The breakdown of shipments among different buyer channels is also available, with resellers accounting for 76% and ready-mixed concrete firms accounting for 11%, in this same period.

capacity. Notice that I do not observe producer prices, only consumer prices. However, I back out producer prices assuming competition at the retail (reseller) level and taking into account the high proportional sales taxes. The assumption of competition among resellers follows from several field interviews, including interviews with producers' sales representatives and resellers. (I also check the robustness of this assumption by, for example, comparing *observed* producer prices that I was able to obtain from a subset of producers to the backed-out producer prices.) This study thus considers the entire supply chain from the producer of cement (and extractor of the raw material) to the retail consumer, encompassing the reseller: in addition to plant marginal cost, total plant-to-market marginal cost consists of plant-to-market freight, sales taxes and the reseller's mark-up.

Figure 2-10 depicts cement prices (in units of local currency for the standard 50 kg bag, at a constant December 1999 level[43]), cement consumption and exogenous demand (activity in the construction sector) from January 1991 through December 2003 for the largest market, the state of São Paulo (*SP*). The month in which the stabilisation plan was implemented, July 1994, corresponds to observation (month) 43 in the graphs (marked by dotted lines). Following the lifting of price controls in November 1991, prices approximately doubled in the first two years of the pre-stabilisation period I cover, remaining in the high R$ 14 to R$ 16 / bag range until 1994. In the post-stabilisation period they gradually declined back to R$ 7 by late 1996, gradually rising thereafter. The sharp increase in consumption following stabilisation, from a level of 600 mt per month to 1000 mt per month within two years, pulled by a 20% jump in the level of construction activity, is evident from the graphs. Some factor prices are also portrayed. It is interesting to note that in the post-stabilisation phase the correlation between cement prices and the prices of fuel oil and diesel oil (the two major components of cost, used respectively in the kiln and in freight) is high[44]. This is expected in view of (i) my earlier claim

[43]This is done using an economy-wide General Price Index (GPI). Owing to the high levels of inflation prevailing in the first 42 months (out of 156) that I consider, particular attention has been paid to the conversion of current cement prices to constant prices – see Appendix B. Factor prices are similarly converted. In contrast, Figure 2-9 presents current prices (albeit for another state).

[44]From July 1994, correlation coefficients (all highly significant) are as follows: 0.72 between cement prices and the (US dollar) exchange rate; 0.86 between cement prices and the price of fuel

(at this point) that imports set a price ceiling for cement and thus cement prices (in local currency) are highly correlated with the exchange rate, and (ii) oil is a global commodity and policy in the oil sector from the second half of the 1990s has prescribed domestic oil prices varying in line with the world price (and hence with the exchange rate). Though the picture varies across states, if only due to different changes in industry structure and demand conditions, the case for the state of São Paulo is broadly representative for Brazil as a whole, in addition to accounting for around one-third of the nation's cement consumption.

A glance at price-cost margins and the robustness of constructed marginal cost With respect to firm profitability, Figure 2-11 shows the evolution of average consumer prices, marginal cost and price-cost margins on the leading firm Votorantim's actual sales across Brazil, in constant local currency units per bag. (Figure 5-2 in Appendix B breaks this figure down into figures for each of the 25 states where Votorantim is present.) Prices and marginal cost have been increasing since late 1996, the latter owing chiefly to increases in the price of fuel oil and diesel (freight) and the fact that sales taxes are proportional to prices – recall that cost relates to the entire supply chain, including freight, sales taxes and the reseller's cost. The picture is similar across firms. *In sum, the industry wields considerable market power, despite the threat of imports.* Across producers, across states and over time, the price-cost margin as a proportion of the consumer price lies in the region of 25-45% (equivalent to 40-65% as a proportion of the producer price net of sales tax).

I conduct two robustness checks of the calculated marginal costs and the resulting price-cost margins (for further details, see Appendix B). The first check consists of comparing my measures of price-cost margins as a percentage of net producer sales (i.e. net of sales taxes) to reported EBITDA (earnings before income tax and depreciation allowance, also known as operating cash flow) as a percentage of net sales for the firm Cimpor, over the period 1998 to 2003. (This firm, which bought its way into Brazil in 1997, is listed on the Lisbon stock exchange, and

oil; 0.77 between the price of fuel oil and the exchange rate.

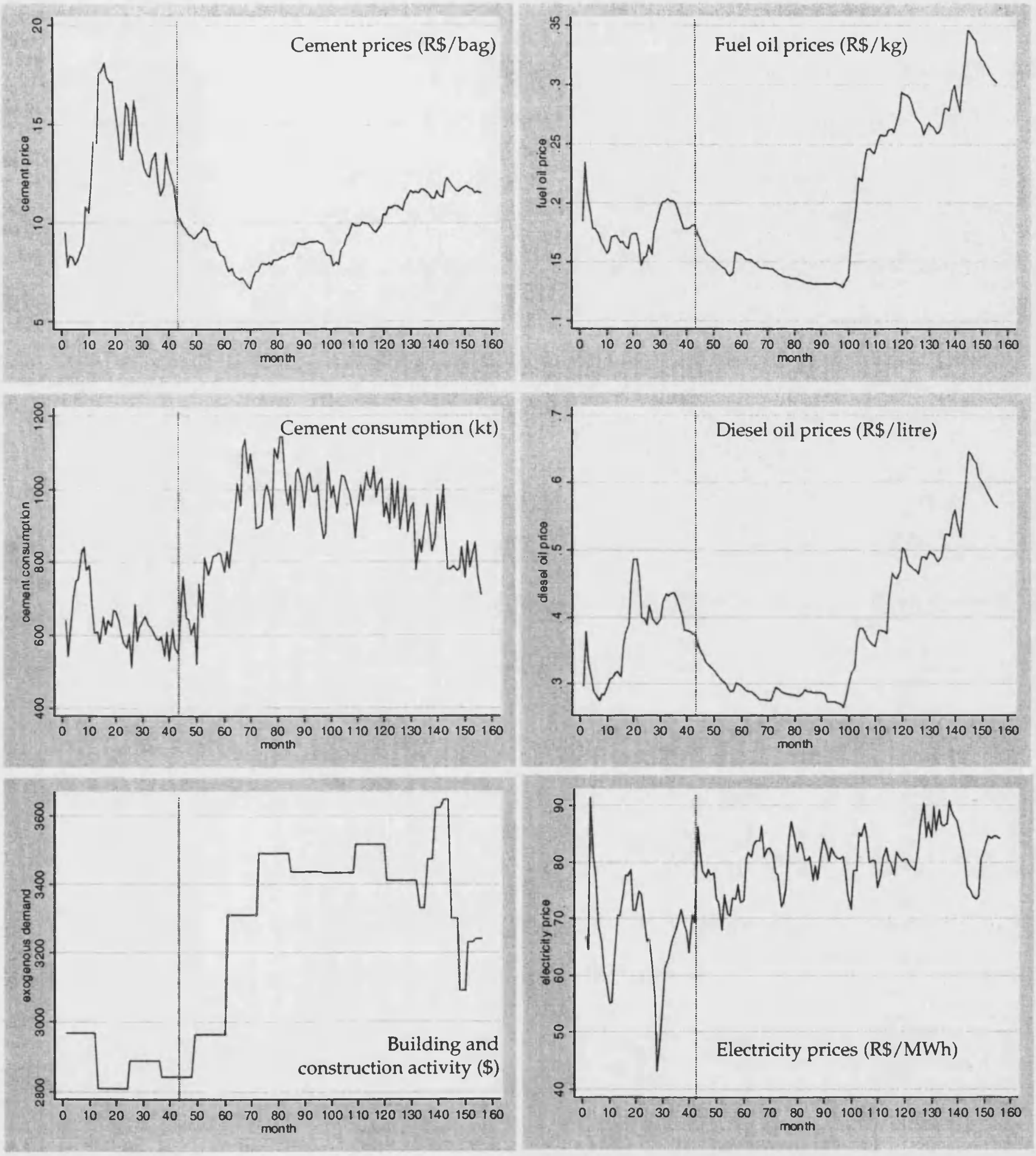


Figure 2-10: Cement prices, consumption, exogenous demand and factor prices for the state of São Paulo. All prices are in constant December 1999 values. Monthly observations, observation 1 corresponding to January 1991. July 1994, the month in which the stabilisation plan was enacted, is marked by the dotted lines.

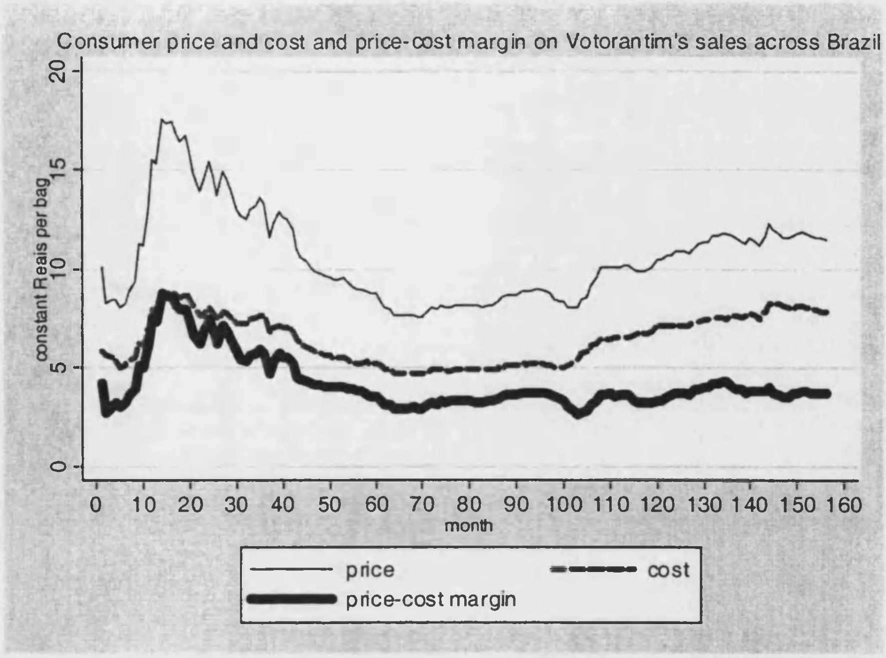


Figure 2-11: Evolution of consumer prices, marginal costs and price-cost margins on Votorantim's sales. Averaged across all states. In constant Reais per bag (at December 1999 values).

fortunately reports its financial results broken out by country of operation and line of business.) The time series fit between constructed and reported figures is good of Goods Sold does not include freight but it includes depreciation.)

A final word on capacity utilisation Throughout the time period, capacity considerably exceeds production, including the three years post stabilisation of steep consumption growth (1995 to 1997), although the slack is lower. Capacity utilisation hovers around an average 65%. The reader is referred to Appendix B for details (and a discussion of a strategic role for capacity is provided in 3.2).

2.4 Inferring demand and conduct in the Brazilian cement industry

2.4.1 A "road map"

Having laid out the theoretical framework and provided an outline of the industry and the data, this Section turns to the empirical application. Section 2.4.2 begins by estimating demand in each local market (state). The market price elasticities of demand are estimated to be very low, of the order of -0.5. Demand across local markets is consistently inelastic at the equilibrium, including local markets where the one-firm concentration ratio is as high as 80%. Recall that observed (constructed) price-cost margins are high in equilibrium. I argue that the reason why the industry, with its considerable pricing power, does not further raise prices (and further restrict output) is that the competitive fringe of imports sets a price ceiling which binds at the industry equilibrium. As such, while the market elasticity is low (-0.5), the price elasticity of the demand that the *domestic* industry faces is much higher: the equilibrium lies at the kink in the residual demand curve facing the domestic oligopoly (recall the right panel of Figure 2-1 in the theoretical framework developed earlier).

I have shown that when the threat of entry constrains prices, joint inference of conduct and cost from the estimation of a static pricing equation will not be consistent. The lack of price variation as demand moves exogenously will lead to the overestimation of competition; econometrically, the exogenous demand variables will be correlated with the error in the misspecified pricing equation. This is what Section 2.4.3 shows with regard to the industry at hand. *I assume costs are not known* and proceed to estimating a market-level pricing equation, instrumenting with exogenous demand. The conduct parameter is estimated to be close to zero and costs are estimated to be close to prices, wrongly suggesting that the outcomes in the Brazilian cement industry are competitive. The negative bias in the estimated price-cost margins is severe in light of the high price-cost margins I measure directly, as presented earlier in Section 2.3.3.

Having rejected competitive conduct based on the known price-cost margins, Section 2.4.4 delves deeper into the pattern of conduct in the industry, taking into account the constraint posed by imports on industry outcomes. I use the test of Proposition 2.3, which is based on direct measures of costs and the observed firm-level quantity data (i.e. the flow of cement from plants to local markets), to show that conduct is considerably more collusive than the Cournot benchmark. Market outcomes are characteristic of (tacit) market division, *and this can be identified despite the threat of imports restraining prices.*[45]

2.4.2 Demand

There are L (geographic) markets (identified with states of the Brazilian federation), indexed by $l = 1, ..., L$. Scattered across these L markets are I plants, indexed by $i = 1, ..., I$.[46] Let $i = 0$ index the aggregate fringe of foreign suppliers. The flow of cement for consumption can be summarised in a set of $(I+1) \times L$ matrices, one matrix for every time period t, where element q_{ilt} denotes the quantity of cement shipped by plant i for consumption in market l in that time period. Let q_{lt} denote total shipments to market l in period t, i.e. consumption; then $q_{lt} = \sum_{i=0}^{I} q_{ilt}$. The

[45]In Chapter 3 I show how such an arrangement may be sustained in equilibrium in a context where firms meet in different markets. Inspired by the Brazilian cement industry, I provide examples of simple dynamic games which give rise to such collusive behaviour in equilibrium.

[46]Not all plants are active (in the sense that cement is shipped from them) in each time period t; in a given time period, some plants may be yet to enter (or reenter), while others may have exited.

demand function in each market l can then be written:

$$q_{lt} = D(p_{lt}, Y_{lt}, \alpha_l, \epsilon_{lt}^d) \tag{2.11}$$

where p_{lt} is the price of cement to the consumer, Y_{lt} are exogenous variables shifting demand (e.g. output in the construction and building sector), α_l are market-specific parameters to be estimated and ϵ_{lt}^d is an econometric error term. (Demand function $D(.)$ is the inverse of the inverse demand function $p(.)$ considered in the theoretical framework, i.e. $D(.) = p^{-1}(.)$.)

Estimation of (2.11) must deal with the (potential) endogeneity of prices. The choice of instruments will depend on whether imports restrain domestic prices at the industry equilibrium (i.e. whether the imports constraint binds, in which case prices are given by the marginal cost of imports c_I, as in Figure 2-5), which in turn depends on the behaviour of domestic firms[47]. There are therefore two situations to consider.

Identification 1: Imports restrain domestic prices at the industry equilibrium In practice, due to the presence of frictions, cement prices will not be exactly equal to c_I. Prices and c_I should be highly correlated however. As mentioned in Section 2.2 (recall the right panel of Figure 2-1), fluctuations in the marginal cost of imports allow one to trace out the demand curve (assuming c_I does not rise to the extent where imports no longer have bite). The marginal cost of imports is a function of factors such as the exchange rate, world fuel prices (used in the production of clinker abroad and in the international transport of cement), tariffs and port handling charges, and domestic freight to the consumer (the latter being highly correlated with the domestic price of diesel oil). Observed factors such as the exchange rate, world fuel prices and domestic diesel oil prices (all in local currency in constant terms) can then be used as instruments for prices in the

[47]Recall that in the case of full collusion and $p^M \geq c_I$, as in the right panel of Figure 2-1, the equilibrium price is $p = c_I$. Likewise, in the case of Cournot, imports will have bite if the aggregate Cournot output in the absence of imports falls short of $p^{-1}(c_I)$ – see footnote 25 and the right panel of Figure 2-4.

estimation of (2.11) (under the identifying assumption that these factors are not correlated with the unobserved market-specific demand shocks ϵ_{lt}^{d}).

To the extent that the "frictions component" of cement prices – i.e. the part of prices not determined by the marginal cost of imports c_I – is orthogonal to the unobserved demand shocks across local markets, prices can be treated as predetermined and (2.11) can be estimated by OLS[48].

Identification 2: Imports do *not* restrain domestic prices at the industry equilibrium When imports do not restrain domestic prices, traditional cost-shifters may be used to instrument for cement prices. These include factor prices (i.e. prices of kiln fuel such as fuel oil and coal, electricity prices which determine the cost of grinding, the price of diesel oil which drives the cost of freight, and wages, the latter also impacting freight in addition to the cost of production) and other supply-shifters such as plant capacity, to the extent that changes to scale impact (flat) marginal cost[49].

Demand specification

I begin by specifying the market-level demand function (2.11) in loglinear form as:

$$\log q_{lt} = \alpha_l^1 + \alpha_l^2 Y_{lt} + \alpha_l^3 \log p_{lt} + \alpha_l^4 Y_{lt} \log p_{lt} + \epsilon_{lt}^d \qquad (2.12)$$

For each market there are 156 monthly observations, from January 1991 to December 2003. Given the quarterly seasonality of sales, three quarterly dummies – not shown in (2.12) – are included. The inclusion of an interaction term between (log) price and the exogenous demand variable (construction and building activity),

[48]The model I have in mind when imports restrain domestic prices is as follows. Cement prices p are determined by the marginal cost of imports c_I and a frictions component ζ, i.e. $p = c_I + \zeta$. As for c_I, as the econometrician I observe some cost drivers V^1 but not others V^2, where $c_I = V^1\kappa + V^2\phi$, and κ and ϕ are parameters. Under the identifying assumption that $E(V^1\epsilon^d) = 0$, where ϵ^d captures the unobserved demand shocks in (2.11), V^1 (e.g. the exchange rate) can be used to instrument for prices in the estimation of (2.11). In addition, if $E(\zeta\epsilon^d) = 0$ (and of course $E(V^2\epsilon^d) = 0$ as well), demand equation (2.11) can be estimated consistently by OLS.

[49]Other instruments can be used as a robustness check, such as Hausmann-type instruments (prices of cement in other local markets) or lagged prices or first differences.

$Y_{lt} \log p_{lt}$, allows the demand curve in logs to rotate – in addition to shift, through the level term Y_{lt} – as exogenous demand varies.

By the earlier discussion, (2.12) is estimated, for each local market, by (I) OLS, (II) 2SLS using the exchange rate and other prices relevant to the marginal cost of imports (such as world oil prices and local diesel oil prices) as instruments (all in constant local currency), and (III) 2SLS using factor prices as instruments. These three sets of results are depicted in Figure 2-12 for the state of São Paulo (*SP*), denoted respectively as "OLS", "IV imports bite" and "IV imports no bite"[50]. (I illustrate using results for the state of *SP*, as this is the largest market, accounting for 29% of national cement consumption in 1999; as shown below, however, results across states follows a common pattern.) Most estimated coefficients are significantly different from zero, many at the 1% level of significance. The interaction term is found to be negative and highly significant: the demand curve (in logs) rotates anticlockwise as exogenous demand expands[51]. Each fitted equation is evaluated at two different values for the exogenous demand variable: at the mean for the pre-stabilisation (high inflation) phase, $\bar{Y}_{SP,pre} = 2883$, from January 1991 through June 1994 (42 observations), and at the mean for the post-stabilisation (low inflation) phase, $\bar{Y}_{SP,post} = 3338$, from July 1994 through December 2003 (114 observations). The (average) market price elasticity of demand during the pre-stabilisation phase amounts to (an inelastic) -0.17, rising to -0.33 during the post-stabilisation phase (see the respective coefficients on log price in column (II), respectively $\hat{\alpha}^3_{SP}+\hat{\alpha}^4_{SP}\bar{Y}_{SP,pre}$ and $\hat{\alpha}^3_{SP}+\hat{\alpha}^4_{SP}\bar{Y}_{SP,post}$). Thus, as prices in the economy stabilise and an average 16% exogenous increase in the demand for cement occurs, the price elasticity seems to double from around -0.2 to around -0.4. Clearly, a formal test that the price elasticity has increased is equivalent to verifying that the coefficient on the interaction term is significantly negative. This is so: the p-value for this (one-tailed) test is 1.5%. Importantly, to check the robustness of the low elasticity I repeat regressions (II) and (III) using only the 114 observations from

[50]Standard errors are heteroskedasticity and autocorrelation-robust (1-lag Newey-West errors). Estimates for the three quarterly dummies are not shown, but are usually significantly negative in the first quarter, significantly negative or insignificant in the second quarter, and insignificantly negative, insignificant or significantly positive in the third quarter.

[51]Estimation of specification (2.11) *excluding* the interaction term renders a significantly positive coefficient on the level of exogenous demand, as expected (results are not shown).

the post-stabilisation subsample (July 1994 on). This confirms a low elasticity of -0.4 (results are not shown). Figure 2-13 plots the fitted demand curve for the pre- and post-stabilisation phases (i.e. evaluated at the respective means $\bar{Y}_{SP,pre}$ and $\bar{Y}_{SP,post}$), indicating that as stabilisation took place and exogenous demand grew, the demand curve shifted out and rotated anticlockwise. In addition to the state of São Paulo (SP), similar plots are drawn for the three next largest markets, the states of Minas Gerais (MG), Rio de Janeiro (RJ) and Bahia (BA). These suggest that this pattern may be typical across states, as I argue after considering some specification tests.

Specification tests I test for the presence of heteroskedasticity and serial correlation using diagnostic tests such as Pagan and Hall (1983). While under OLS I can clearly reject homoskedasticity, I can no longer reject homoskedasticity under 2SLS. In any case, I choose to allow for heteroskedasticity and serial correlation by calculating heteroskedasticity and autocorrelation-robust standard errors (1-lag Newey-West errors). As for the choice of instruments, I verify the overidentifying restrictions in specifications (II) and (III) using different tests such as those based on Hansen's J statistic or variations of the Sargan statistic (see Wooldridge 2002, or Baum, Schaffer and Stillman 2003). In both specifications (II) and (III), I can reject the null that the set of instruments is orthogonal to the error term, casting doubt on the validity of the set of instruments, despite the finite sample properties of such statistics suggesting that the test results should be interpreted with caution. (Further, from the structural model one would not expect that market-specific demand shocks be correlated with the instruments, such as the exchange rate and the price of diesel oil.) To check the extent to which overidentification may be driving efficiency in the estimation at the expense of consistency, I reestimate specifications (II) and (III) in each case using only a subset of the initial set of instruments such that the validity of instruments can no longer be rejected. This is shown in Figure 2-12 as regressions (II B) and (III B) respectively. Comparing estimates for (II B) against those for (II), for example, the estimated elasticities are similar. I also test for the endogeneity of prices (and the interaction term) using endogeneity tests à la Durbin-Wu-Hausman (Wooldridge 2002). Under specifications (II) and (III),

		(I) OLS	(II) IV imports bite	(III) IV imports no bite	(II B) IV subset imports bite	(III B) IV subset imports no bite
No. obs.		156	156	156	156	156
R^2		0.840				
Intercept	coef	2.241 *	2.828 **	2.439 *	0.212	0.729
	s.e.	(1.202)	(1.210)	(1.236)	(1.357)	(1.333)
Exog. demand	coef	0.00159 ***	0.00141 ***	0.00152 ***	0.00225 ***	0.00203 ***
	s.e.	(0.00038)	(0.00039)	(0.00039)	(0.00043)	(0.00042)
Log Price	coef	1.093 **	0.852 *	1.003 *	1.954 ***	1.702 ***
	s.e.	(0.498)	(0.504)	(0.514)	(0.564)	(0.554)
Interaction	coef	-0.000428 ***	-0.000355 **	-0.000396 **	-0.000709 ***	-0.000607 ***
	s.e.	(0.000160)	(0.000163)	(0.000166)	(0.000181)	(0.000176)
Quarterly dummies		Included	Included	Included	Included	Included
Evaluating at the mean of exogenous demand pre-stabilisation: 2,883						
Intercept	coef	6.825 ***	6.898 ***	6.815 ***	6.699 ***	6.594 ***
	s.e.	(0.143)	(0.144)	(0.145)	(0.155)	(0.167)
Log Price	coef	-0.142 **	-0.171 ***	-0.138 **	-0.091	-0.048
	s.e.	(0.055)	(0.056)	(0.056)	(0.060)	(0.065)
Evaluating at the mean of exogenous demand post-stabilisation: 3,338	16% growth versus pre-stabilisation phase					
Intercept	coef	7.549 ***	7.541 ***	7.507 ***	7.724 ***	7.521 ***
	s.e.	(0.129)	(0.136)	(0.135)	(0.142)	(0.141)
Log Price	coef	-0.337 ***	-0.333 ***	-0.318 ***	-0.414 ***	-0.325 ***
	s.e.	(0.058)	(0.060)	(0.060)	(0.063)	(0.062)
Test of overidentifying restrictions			Fail	Fail	Pass	Pass

Note: Heteroskedasticity and autocorrelation-robust standard errors (Newey-West 1 lag)
*** Significant(ly different from zero) at the 1% level; ** Significant at the 5% level; * Significant at the 10% level
Dependent variable is Log Consumption
Quarterly dummy variables for quarters 1, 2 and 3 are included but estimates are not shown

Figure 2-12: Demand estimates for the state of *SP*

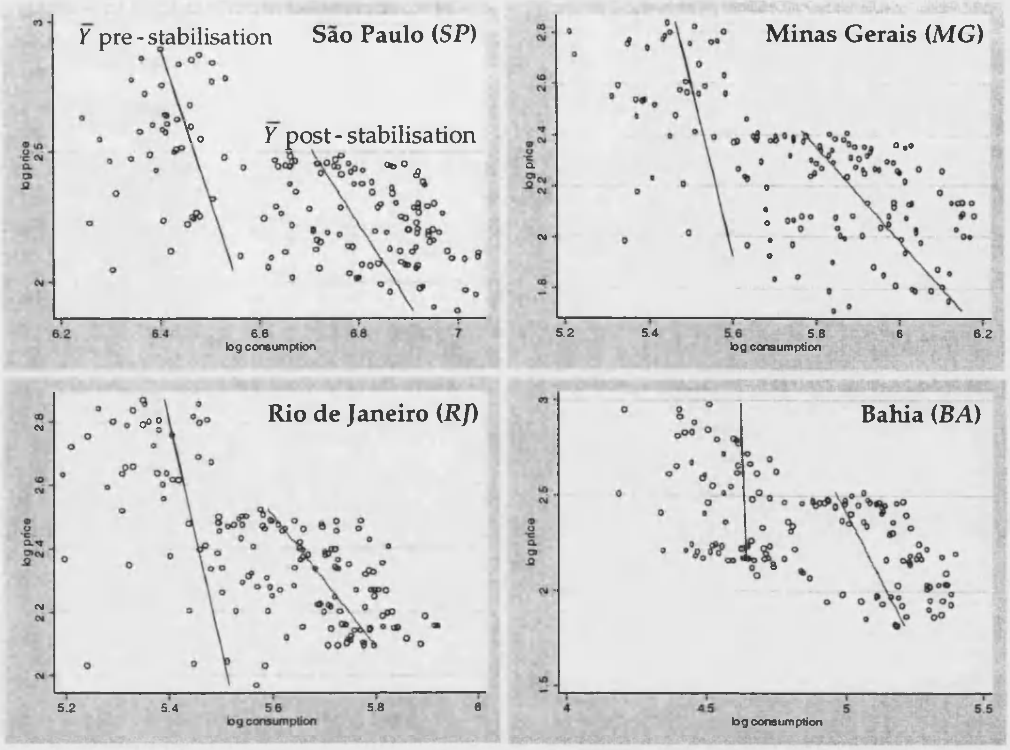


Figure 2-13: Fitted demand curves for the four largest markets. (Log) Price against (Log) Consumption. Evaluated at the respective means of exogenous demand Y for the pre- and post-stabilisation phases.

using the initial set of instruments for each specification, I cannot reject the n hypothesis that prices are orthogonal to the error, lending some credibility to t OLS estimates, as explained earlier. However, repeating these tests under spec cations (II B) and (III B), using only a subset of the instruments, I can reject t null of orthogonality in some instances, but not all. Finally, given the efficiency GMM under heteroskedasticity, I reestimate specifications (II) and (III) (and th counterparts which use a subset of the initial number of moment conditions) usi GMM. Results – not shown – are similar.

Results by state Figure 2-14 summarises results across states, from regressi (II) using the full sample (pre- and post-stabilisation phases). At first glance, t picture is mixed when compared to the results just described for the state of *S* The estimated price elasticity for 3 out of the 27 states is positive (evaluating significantly so, suggesting that for these states the demand curve slopes upward! The estimated coefficient for the interaction term is positive for 9 out of the 27 states (4 of which significantly so), suggesting that as exogenous demand expands in the post-stabilisation phase – and indeed it does expand in every one of the 27 states – the demand curve for these states rotates *clockwise*. The estimates for these states stand in contrast to the results obtained for the state of *SP*, which I claimed represented a typical pattern across states. Upon closer inspection, however, one notices that these "outlier" states are mostly located to the northwest of the centre of the country, an area which is sparsely populated and largely covered in jungle. Together they account for 60% of Brazil's land area but only 13% of its population and only 11% of its cement consumption in 1999 (see Figure 2-6). It is thus reasonable to believe that the measurement error associated with data collected for these states is large[52]. I thus choose to drop these 10 states from the analysis, segregating their estimates at the bottom of Figure 2-14.

For the remaining 17 states, the pattern is similar and consistent with the results reported for the state of *SP*. Evaluating exogenous demand at its mean annual value in the post-stabilisation phase, the price elasticity of demand is negative for all 17 states, and significant at the 1% level in 15 states. Price elasticities in the post-stabilisation phase vary from a minimum (in absolute) of −0.14 to a maximum of −0.72, with a mean of −0.41 and a standard deviation of 0.14[53]. It is worth

[52] Indeed the leading global market research and data collection firm, ACNielsen, well-known to marketing professionals in consumer goods industries, and with over 30 years of experience in Brazil, does not audit any of these "outlier" states, to the northwest of the centre, except for the Federal District (*DF*) and the state of Goiás (*GO*). I also choose to drop *DF* because it consists essentially of a city, the federal capital Brasília, with a large population of 2m and two cement plants, embedded at the corner of the state of *GO*, near to the borders of the states of *MG* and *BA*, over which product must flow which I do not observe. For this reason the measurement error associated with consumption figures for *DF* may be large. I also drop *GO* on the basis of probable measurement error: in addition to its low population density, until 1988 the states of *GO* and *TO* comprised one single state, previously known as Goiás, when in 1988 the northern half of the state broke away to form the state of Tocantins (*TO*).

[53] Other studies of cement have found low market price elasticities in equilibrium. For example, Röller and Steen (2002) find an average −0.46 for Norway (treating it as a single market) using yearly observations between 1955 and 1968. Also using yearly observations, over 25 years up to 1989, Jans and Rosenbaum (1996) report an average fitted elasticity of −0.81 across 25 regional US markets. It is conceivable that in these markets imports have also been restraining the prices set by the domestic oligopolies. The explanation commonly advanced behind such inelastic demand is that cement accounts for a low share of construction budgets and it has few substitutes (except in highway construction, where asphalt is a substitute). Yet while helping to explain the steepness of the inverse demand curve, this does not explain the steepness *at the equilibrium*. One must still

		(II) IV-imports bite					
	Cement consumption	Interaction		Log Price: *Y* evaluated at mean pre-stabilisation		Log Price: *Y* evaluated at mean post-stabilisation	
State	in 1999 (kt)	coef	s.e.	coef	s.e.	coef	s.e.
20 *SP*	11,723	-0.000355	(0.000163) **	-0.171	(0.056) ***	-0.333	(0.060) ***
17 *MG*	5,090	-0.001067	(0.000235) ***	-0.147	(0.063) **	-0.549	(0.059) ***
19 *RJ*	3,809	-0.002660	(0.000575) ***	-0.137	(0.059) **	-0.481	(0.057) ***
16 *BA*	2,461	-0.003048	(0.000815) ***	-0.027	(0.065)	-0.361	(0.079) ***
21 *PR*	2,321	-0.001015	(0.000647)	-0.137	(0.087)	-0.278	(0.088) ***
23 *RS*	2,221	-0.001057	(0.000762)	-0.228	(0.037) ***	-0.379	(0.097) ***
22 *SC*	1,648	-0.003488	(0.002647)	0.020	(0.091)	-0.180	(0.095) *
13 *PE*	1,225	-0.003389	(0.001675) **	-0.285	(0.093) ***	-0.469	(0.061) ***
10 *CE*	1,139	-0.005347	(0.001662) ***	-0.142	(0.125)	-0.562	(0.113) ***
18 *ES*	837	-0.003029	(0.002317)	-0.370	(0.078) ***	-0.480	(0.068) ***
8 *MA*	765	-0.020114	(0.007056) ***	-0.097	(0.187)	-0.564	(0.126) ***
12 *PB*	565	-0.036712	(0.007397) ***	-0.123	(0.081)	-0.715	(0.111) ***
11 *RN*	531	-0.005411	(0.004692)	-0.145	(0.146)	-0.300	(0.078) ***
25 *MS*	454	0.000899	(0.004419)	-0.431	(0.047) ***	-0.415	(0.071) ***
14 *AL*	384	0.080309	(0.030990) **	-0.475	(0.127) ***	-0.351	(0.112) ***
9 *PI*	379	0.015324	(0.012214)	-0.657	(0.272) **	-0.330	(0.103) ***
15 *SE*	282	0.003937	(0.020794)	-0.145	(0.136)	-0.136	(0.099)
Memo: States to the northwest of the centre of the country, mostly sparsely populated							
26 *GO*	1,152	0.002590	(0.002127)	-0.163	(0.053) ***	-0.040	(0.079)
5 *PA*	802	0.000622	(0.004211)	-0.369	(0.144) **	-0.318	(0.246)
27 *DF*	694	-0.038925	(0.015687) **	0.014	(0.074)	-0.153	(0.078) *
24 *MT*	540	0.012129	(0.004524) ***	-0.300	(0.100) ***	0.210	(0.121) *
3 *AM*	327	-0.023351	(0.007191) ***	0.107	(0.101)	-0.611	(0.166) ***
7 *TO*	282	0.000000	(0.000000)	-1.052	(0.218) ***	-1.052	(0.218) ***
1 *RO*	217	-0.028803	(0.020138)	-0.152	(0.440)	-1.194	(0.344) ***
6 *AP*	78	0.357032	(0.192129) *	0.126	(0.315)	0.744	(0.215) ***
4 *RR*	66	1.217022	(0.340927) ***	-1.366	(0.307) ***	0.190	(0.305)
2 *AC*	55	-0.222964	(0.109062) **	-1.406	(0.437) ***	-2.673	(0.352) ***

Note: Heteroskedasticity and autocorrelation-robust standard errors (Newey-West 1 lag)
*** Significant(ly different from zero) at the 1% level; ** Significant at the 5% level; * Significant at the 10% level

Figure 2-14: Demand estimates by state, Summary

pointing out that elasticities are low even in states where the supply of cement is highly concentrated, such as the state of Santa Catarina (SC), where the one-firm concentration ratio is 78% (in 1999). Evaluating exogenous demand at its mean value in the pre-stabilisation phase, price elasticities are negative in 16 out of 17 states, 9 of which are significant at the 10% level or higher. In the pre-stabilisation

explain why an industry, facing such inelastic demand at the market price, does not cut output in an attempt to raise prices and thus move up along the demand curve to a point where demand is more elastic. To rationalise why the industry seeing such inelastic demand does not restrict output, one would expect either weak pricing power (e.g. competition or low concentration, such that the elasticity facing individual producers is higher than the aggregate elasticity) or some dynamic story, such as the threat of entry restraining prices.

phase the mean price elasticity is lower: -0.22.[54]

Summary of demand and robustness checks (see Chapter 3) Regardless of the type of price instruments employed (or using prices themselves, under OLS), which depends on whether one accepts that imports restrain domestic prices at the industry equilibrium, I estimate very low market price elasticities of demand, of the order of -0.5. In Section 3.4 I consider further robustness checks. The overall conclusion arising from these checks is that this section's finding – that demand is inelastic (of the order of -0.5) at the equilibrium in each local market – is robust.

2.4.3 Estimation of a static pricing equation: inconsistent cost and conduct estimates

I now *ignore the constraint posed by imports* on the prices set by the domestic oligopoly. Based on the estimates of demand of Section 2.4.2, and *assuming costs are not known*, I proceed to estimating a pricing equation such as (2.3), as is standard in the empirical literature on market power[55]. I then compare the estimates for cost and conduct, summarised in the estimated elasticity-adjusted price-cost mark-ups (2.2), to the observed (constructed) values. I show that the estimation exercise considerably overestimates the degree of competition and the marginal cost, incorrectly suggesting that price-cost margins in the Brazilian cement industry are centred around zero.

I begin by specifying the (standard) structural econometric model. Recall the market-level demand function (2.11) defined at the beginning of Section 2.4.2 and

[54]Given that exogenous demand – activity in the construction and building sector – rises on average in *every* state concurrent with stabilisation, this finding that the price elasticity of demand also increases is equivalent to estimating a negative coefficient for the interaction term. Indeed the fitted interaction coefficient is negative in 13 out of the 17 states, in 8 of which at the 5% level of significance. See footnote 33 of Chapter 3 for one possible explanation behind the increase in elasticity upon stabilisation.

[55]Since I ignore that the true model is (2.4) and that the estimated model is thus (2.5) with $E(Y'\xi^s) < 0$, here I refer to the pricing equation error as ε^s and not ξ^s, as a researcher overlooking the price-constraining effect of imports would do, thinking that he is specifying (2.3) with $E(Y'\varepsilon^s) = 0$.

write its inverse $(D^{-1}(.) = p(.))$ as

$$p_{lt} = p(q_{lt}, Y_{lt}, \alpha_l, \epsilon^d_{lt})$$

Define plant i's costs as

$$C_{it} = C(q_{it}, \mathbf{q}_{it}, W_{it}, Z_{it}, \beta, \varepsilon^s_{it})$$

where $q_{it} := \sum_{l=1}^{L} q_{ilt}$ denotes plant i's shipments aggregated across markets l (equal to production), W_{it} are the prices it pays for its factors, and Z_{it} are other exogenous variables that shift supply. Note that costs by plant will not only depend on the plant's total shipments q_{it} but also on the destination of these shipments $\mathbf{q}_{it} := (q_{i1t}, q_{i2t}, ..., q_{iLt})$, owing to market-specific factors such as freight. β is a vector of common parameters to be estimated and ε^s_{it} is a plant-specific error. The I plants are owned by F firms, indexed by $f = 1, ..., F$. Define $\mathcal{O}_{ft}$ as the set of plants owned by firm f in month t. In period t firm f solves

$$\max_{q_{ilt}|i\in\mathcal{O}_{ft},\forall l} \sum_{l=1}^{L} \left[p(q_{lt}, .) \left(\sum_{i\in\mathcal{O}_{ft}} q_{ilt} \right) \right] - \sum_{i\in\mathcal{O}_{ft}} C(q_{it}, \mathbf{q}_{it}, .)$$

In words, firm f sets shipments from each plant it owns to each market to maximise its profits, which correspond to the difference between the sum of revenues across markets and the sum of costs across plants. Denote the derivatives of the (inverse) demand and cost functions with respect to q_{lt} and q_{ilt} respectively as $p_1(.)$ and $c(.)$. Following Bresnahan (1989), the first-order condition for multi-plant firm f with regard to shipments from its plant $i \in \mathcal{O}_{ft}$ to market l, i.e. q_{ilt}, yields a pricing equation for each plant i - market l pair:

$$p_{lt} + p_1(q_{lt}, Y_{lt}, \alpha_l, \epsilon^d_{lt}) q_{lt} \theta_{flt} \leq c(q_{it}, \mathbf{q}_{it}, W_{it}, Z_{it}, \beta, \varepsilon^s_{it})$$

The pricing equation can be written as an equality when $q_{ilt} > 0$ (i.e. an interior solution). Recall from the earlier discussion of Section 2.2.2 that this specification encompasses alternative models of conduct. At the two extremes, $\theta_{flt} = 1$ captures full collusion while $\theta_{flt} = 0$ reflects price-taking behaviour (competition); a θ_{flt}

equal to firm f's market share, i.e. $\theta_{flt} = s_{flt} := \frac{\sum_{i \in \mathcal{O}_{ft}} q_{ilt}}{q_{lt}}$ corresponds to firm f behaving as a Cournot player. For $q_{ilt} > 0$ and specifying an additive econometric pricing error, one may implement this pricing equation as

$$p_{lt} = -\theta_{flt} \frac{p_{lt}}{\eta_{lt}} + c(q_{it}, \mathbf{q}_{it}, W_{it}, Z_{it}, \beta) + \varepsilon^s_{ilt} \tag{2.13}$$

(recall η is the market price elasticity of demand).

In what follows, I present estimation results corresponding to a market-level counterpart to the plant-level pricing equation (2.13). As mentioned in footnote 13, the market-level equation should be viewed as an average across plants' pricing equations[56]. In view of the fixed-coefficient technology of production, I specify average market marginal cost c as being linear in average market factor prices W_{lt} (namely fuel oil, coal, electricity, labour and freight[57]) and flat in quantity (though in other specifications I have also allowed average market marginal cost to vary in quantity). I allow cost to shift according to the average size and age of the plants shipping into the market (weighted by shipments), Z_{lt} (e.g. marginal cost in a market served by high-capacity plants should be lower). (Finally, a dummy is included to account for price controls in the first ten months of 1991: this supply-shifter may be viewed as an additional element of Z_{lt}.) The market-level pricing equation is thus

$$p_{lt} = -\theta_l \frac{p_{lt}}{\hat{\eta}_{lt}} + W_{lt}\beta^1 + Z_{lt}\beta^2 + \nu_l + \varepsilon^s_{lt} \tag{2.14}$$

where ν_l is a market-specific fixed effect and the market-specific conduct parameter θ_l is time-invariant (other specifications I have fitted allow θ_l to vary over time,

[56] Owing to the lack of firm-level data, most empirical IO studies have no choice but to estimate a market-level equation. Though I have the luxury of observing plant-level data, I here choose to follow suit to simplify the exposition. Importantly, I have estimated a plant-level pricing equation and have ensured robustness of the conclusions I derive from what follows.

[57] The average factor price, say that of electricity, for a given market is calculated as the average price of that factor paid by the plants sourcing that market (weighted by the sourcing plants' shipments to that market). The price effect of substitute kiln fuels (fuel oil and coal) are interactions of the average price of the fuel and the average use of that fuel in the production of cement shipped to the market (i.e. given the location of coal mines in the South of the country, coal prices have a larger effect on the cost of cement plants located in the South). A market's average plant-to-market freight price is modelled as the interaction of the average distance to market across the plants sourcing that market (again weighted by shipments) and a transport price index (based heavily on the price of diesel oil).

such as upon stabilisation). Equation (2.14) is fitted using fixed-effects instrumental variables panel data estimation, where the endogenous regressor $\frac{p_{lt}}{\hat{\eta}_{lt}}$ is instrumented using excluded exogenous demand variables Y_{lt}, and thus the orthogonality condition $E(Y'\varepsilon^s) = 0$ *is imposed.* Since the elasticity $\hat{\eta}_{lt}$ is an estimate based on the demand estimates from Section 2.4.2, I compute bootstrapped (heteroskedasticity-robust) standard errors (with 1000 repetitions, reestimating demand in the first stage for each bootstrap sample) for the fitted coefficients $\hat{\theta}_l$ and $\hat{\beta}$ (and $\hat{\nu}_l$). Notice that though knowledge of the nature of technology is used when specifying marginal cost to be linear in factor prices, *marginal cost is assumed to be unknown*: this is estimated from the observed supply-shifters (W_{lt}, Z_{lt}) and the estimates of the fixed coefficients (β, ν) as $W_{lt}\hat{\beta}^1 + Z_{lt}\hat{\beta}^2 + \hat{\nu}_l$.

Figure 2-15 reports estimation results. The comments that follow refer to estimation (I), based on the entire period, though estimates based only on observations from the post-stabilisation period (i.e. from July 1994) are provided to demonstrate robustness of the conclusion that follows. The coefficients on the prices of fuel oil, coal, electricity and freight are all positive and significant. The coefficient on the average size (age) of plants is negative (positive), as expected, though not significant. On the other hand, contrary to intuition, the price of labour is significantly negative[58]. The price-cost margins are estimated to be very low; these are pictured in Figure 2-16, along with 95% confidence intervals, for the state of Rio de Janeiro (RJ), for example. The dual to these cost estimates are the low estimated conduct parameters $\hat{\theta}_l$, not significantly different from 0, suggesting competition[59]. For the state of RJ, a $\hat{\theta}$ of 0.0079 would correspond to the equilibrium price-cost margins of a static symmetric 130-firm Cournot industry (1/0.0079). Dividing $\hat{\theta}_{RJ}$ by the (negative of the) estimated elasticity $\hat{\eta}_{RJ}$ of -0.48 from Figure 2-14, the estimated

[58]Interestingly, this counter-intuitive estimate is also obtained by Jans and Rosenbaum (1996), who estimate a market-level structural model of the US cement industry (it must be pointed out that, despite their penetration into US markets, a price-constraining role for imports is not considered in Jans and Rosenbaum's model). In attempting to explain the negative coefficient on wages, the authors cite Clark (1980), who suggests that plants that pay higher wages may do so because their labour force is more productive (and possibly more unionised, as Clark finds the more productive US cement plants to be).

[59]It is worth mentioning that the estimated confidence intervals for $\hat{\theta}_l$ vary considerably according to the specification (such as the functional form for demand), though low (absolute) values do seem to be a robust result.

(average) price-cost margin as a proportion of price is only $0.0079/0.48 \approx 1.6\%$ (recall expression (2.2)).

	(I) IV Full sample		(II) IV Post-stabilisation subsample	
	coef	bootstrap s.e.	coef	bootstrap s.e.
No. obs.	2652		1938	
Market-specific conduct parameters				
20 *SP*	0.0167	(0.0194)	0.0021	(0.0152)
17 *MG*	0.0194	(0.0127)	-0.0049	(0.0163)
19 *RJ*	0.0079	(0.0206)	-0.0112	(0.0120)
16 *BA*	0.0004	(0.0100)	-0.0268 *	(0.0142)
(Parameters for 13 other markets not shown)				
Factor prices				
Fuel oil (interacted with fuel use)	18.1368 ***	(2.7773)	20.1344 ***	(2.7119)
Coal (interacted with fuel use)	0.0906 ***	(0.0343)	0.0447	(0.0430)
Electricity	0.0343 ***	(0.0125)	0.0494 ***	(0.0169)
Labour	-7.7850 ***	(1.2287)	-3.9898 ***	(1.1363)
Freight (distance interacted with price of diesel oil)	0.0065 **	(0.0028)	0.0066 **	(0.0026)
Other supply-shifters				
Size of sourcing plants	-9.38E-08	(5.84E-07)	4.56E-08	(6.64E-07)
Age of sourcing plants	0.0191	(0.0316)	0.0188	(0.0321)
Price controls (Jan 91 to Oct 91)	-4.4828 ***	(0.8479)		
Intercept (*SP*)	12.1986 ***	(2.6814)	6.9360 **	(2.8436)
(Other market-specific fixed effects included but not shown)				

Note: Heteroskedasticity-robust standard errors with bootstrapping to account for demand estimation in the first stage. 1000 repetitions, clustered by month (e.g. in (I) a bootstrap sample consists of 156 month draws, and for every month in the bootstrap sample there are 17 markets).
Demand estimates from the first stage of (II) also based on post-stabilisation subsample.
*** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level
Dependent variable is the price of cement in units of local currency per bag (at Dec 1999 prices)

Figure 2-15: Estimation of a static pricing equation, assuming cost is not known. Instrumented with exogenous demand variables.

It is clear from our knowledge of marginal cost and price-cost margins in the industry that these estimates are inconsistent. Figure 2-16 also depicts the (much higher) direct measures of (average) price-cost margins on sales to the state of *RJ*. What lies behind the market price elasticities of demand of the order of only -0.5 in equilibrium, *is not the prevalence of competition, as suggested by* $\hat{\theta}$, *but the constraining effect of imports on prices at the industry equilibrium.* Econometrically, as argued in Section 2.2.2, this constrained equilibrium translates into the corre-

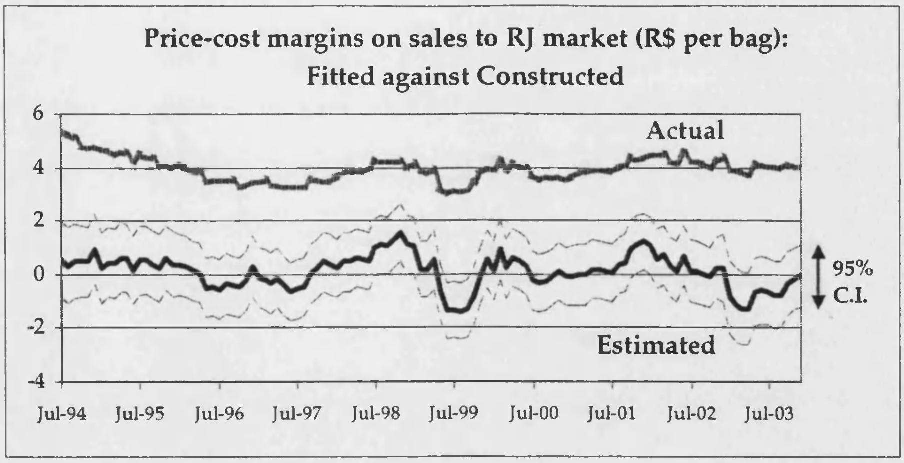


Figure 2-16: Estimated (average) price-cost margin on sales to *RJ* market, as estimated by the static pricing equation, against Actual (constructed) price-cost margin. In constant R$ per bag (December 1999 terms).

lation between the exogenous demand variables being used to instrument pricing equation (2.14) and its residual. The identifying assumption's failure to hold results in the overestimation of the degree of competition. (Indeed, the p-values of overidentification tests à la Sargan and Hansen – where the null is that the set of instruments is valid – is 0.0000 for any overidentifying set.) The finding that the coefficients on factor prices and other supply-shifters are of the expected sign (bar wages) and mostly significant may lead one to misjudge that the econometric model is appropriately specified. But the estimated coefficients are only picking up the expected correlation between cement prices and factor prices. They are not consistent estimates of the structural cost parameters β. The general point is that because the *threat* of entry is not observed, it is only natural that a researcher overlook its role in restraining prices, inadvertently taking the lack of price variation to exogenous movements in demand as evidence in the direction of price-taking behaviour and zero price-cost margins[60].

[60]A comment on this particular industry where conditions (i.e. the steepness of the market demand curve and the marginal cost of imports) are such that demand is so inelastic at the equilibrium. Assume one does not realise that the imports constraint binds at the equilibrium and thus considers the class of behavioural models nested in the static pricing equation $p + \theta_f \frac{p}{\eta} = c_f$ (i.e. the theoretical counterpart to the firm-level pricing equation (2.13), suppressing the error and writing it in the form marginal revenue = marginal cost). It is clear that a η of -0.5 is not consistent with cartel behaviour ($\theta = 1$). Profit maximisation would lead the cartel to cut output until the cartel's marginal revenue were equal (and thus positive) to marginal cost. Nor will such

2.4.4 Inferring conduct in a constrained equilibrium

So far I have argued that what lies behind the very low market price elasticities of demand in equilibrium is the threat of imports restraining prices. As laid out in the theoretical framework of Section 2.2, the competitive supply of high-cost imports sets an upper limit to prices. That market demand is inelastic in equilibrium owes to demand, costs and firm conduct (i.e. the structural parameters of the data generating process) being such that this upper limit to prices binds. The evidence in support of this claim can be summarised as follows:

1. Demand is estimated to be inelastic in all 17 local markets. This is so irrespective of the number and concentration of sellers, ranging from states where the one-firm concentration (C_1) is around 25% to states where C_1 is as high as 80%.

2. In attempting to unravel the low equilibrium elasticities, one must consider price-taking behaviour on the part of cement producers as an alternative explanation to the binding presence of imports. By this alternative explanation – plausible *a priori* – an industry facing such inelastic demand would not be able to cut output to raise prices because competition among producers drives prices down toward marginal cost. However, I reject competition on the basis of the large observed price-cost margins. Producers enjoy considerable market power.

This dynamic story where the threat of entry restrains prices is further supported by a wealth of anecdotal and interview-based evidence. It is also consistent,

a low value of η be consistent with a Cournot industry, unless all firms have small market shares. To see this, notice that if the largest firm has a market share of 50%, then $\max_f\{\theta_f\} = 0.5$ and thus this firm's marginal revenue, to be equated to its marginal cost, is zero. Small market shares across local markets are clearly not the case in the Brazilian cement industry. Any statistical model selection exercise à la Gasmi, Laffont and Vuong (1990, 1992) will thus result in, say, both the cartel model and the Cournot model being rejected in favour of price-taking behaviour by firms ($\theta = 0$). Though misguided in the present case, the high correlation between factor prices and cement prices ensure that the OLS regression of cement prices on factor prices (and other supply-shifters, along with a set of market dummies) – i.e. under the hypothesis of price-taking behaviour – displays good fit: $R^2 = 54\%$. Again, by misspecifying the set of alternative models generating outcomes in the industry, a researcher would mistakenly conclude in favour of price-taking behaviour, not realising that his estimates are simply picking up correlation between cement prices and factor prices.

as argued in Section 2.3.2, with the high correlation observed between cement prices and the exchange rate, the latter having varied considerably over the time period.

I have also shown – and illustrated empirically – that in such a setting the identification of a standard pricing equation from the comparative statics of demand fails. I now turn to the test of conduct spelled out in Proposition 2.3 to cast light on the pattern of behaviour in the Brazilian cement industry. Admittedly, this test provides only a *sufficient* statistic to reject Cournot conduct in favour of more collusive behaviour and the data requirements are large: one must observe both marginal cost and firm-level quantity data. But the value of the test resides in uncovering firm-level behaviour when market outcomes are constrained by the threat of entry. I show that the data rejects Cournot as a benchmark for conduct in the cement industry, in favour of market division. (In the following chapter I illustrate the rationality of such a strategy in an industry where firms meet in different markets.)

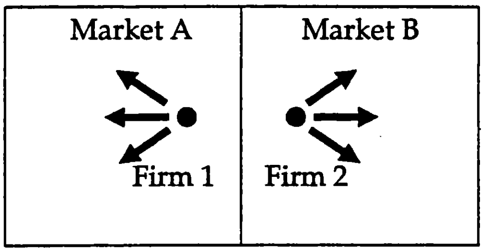
Rejecting Cournot behaviour in favour of collusion

Prior to stating the results of the test of Proposition 2.3 as it applies to the Brazilian cement industry, I provide a flavour of why the data leads to the rejection of Cournot behaviour by considering a specific example extracted from the data. This serves only as an example of a broader trend in the data. As shown subsequently, there are many instances in the data where firms undersupply local markets as compared to the supply decisions of a Cournot firm[61]. In other words, there are many instances where the marginal revenue of a firm in a given market were it adopting the Cournot conjecture – taking its rivals' output as given – significantly exceeds the marginal cost in supplying that market.

A case in point Consider the two adjacent states of Alagoas (AL) and Sergipe (SE), located in the northeast of Brazil (see Figure 2-6). These states are equally small both in terms of market size and geography. Up until 1996 each also had

[61]As in the empirical literature on conduct, the Cournot assumption serves as a benchmark (e.g. Parker and Röller 1997).

only one plant located within its borders: the firm Brennand operated the plant located in AL (respectively firm 1 and market A) and its rival Votorantim operated the plant located in SE (respectively firm 2 and market B)[62]. Consider the year 1996. While firm 1 commands an 83% share in market A, it does not supply to neighbouring market B, right next door to its plant located in market A. Equally striking, firm 2 commands an 89% share in market B, while attaining only a 7% share in the neighbouring market A, next door to its plant in market B. Thus while in market A firm 2's share pales in comparison to firm 1's share, in the neighbouring market B firm 2 dwarfs firm 1 (firm 1 in actual fact does not supply to market B!) Average consumer prices in markets A and B are almost identical, respectively R$ 9.47 (per bag) and R$ 9.46. As explained previously, I calculate firm 1's marginal cost (including sales taxes and the reseller's mark-up) in supplying markets A and B to be respectively R$ 5.20 and R$ 5.47. As for firm 2, I calculate its cost in supplying markets A and B to be respectively R$ 5.30 and R$ 5.17.[63] This is illustrated in the following picture and table, where I take the price elasticity of demand in equilibrium for both markets to be -0.5:[64]



E.g. Market A: *AL* Market B: *SE*
Firm 1: Brennand Firm 2: Votorantim

[62]In late 1996 a third firm, GJS, set up a plant close to Votorantim's plant in SE. However, I abstract away from this in this illustration, by considering the year 1996.

[63]Note that the state-capital cities of AL (market A) and SE (market B) are located less than 300 km apart. Nevertheless, the difference in Votorantim's (say) cost of supplying AL and SE seems low: only R$ 0.13. The reason is that Brazil has an awkward sales tax system which may work against within-state shipments, as happens here, i.e. shipments from Votorantim's plant in SE to resellers in SE are penalised compared to its shipments across the state border to resellers in AL. This mitigates the difference in average freight costs from Votorantim's plant in SE: R$ 0.32 to resellers in SE and R$ 0.77 to resellers in AL.

[64]The subsequent test takes the estimation of η in Section 2.4.2 into account. Also for simplicity, in this illustration I compute an average φ for the year (for each firm-market combination).

(Year 1996)	Price, p	Share, $\frac{q_f}{q}$	MR Cournot, $p+\frac{p}{\eta}\frac{q_f}{q}$	MC, c	Can reject Cournot?
Local market $l = A$ (AL)					
Firm $f = 1$ (Bren)	9.47	0.83	$9.47-\frac{9.47}{0.5}0.83 = -6.25$	5.20	No ($\varphi_{1A} < 0$)
Firm $f = 2$ (Voto)	9.47	0.07	$9.47-\frac{9.47}{0.5}0.07 = 8.14$	5.30	Yes ($\varphi_{2A} > 0$)
Local market $l = B$ (SE)					
Firm $f = 1$ (Bren)	9.46	0	$9.46-\frac{9.46}{0.5}0 = 9.46$	5.47	Yes ($\varphi_{1B} > 0$)
Firm $f = 2$ (Voto)	9.46	0.89	$9.46-\frac{9.46}{0.5}0.89 = -7.38$	5.17	No ($\varphi_{2B} < 0$)

One is thus able to reject the hypothesis of firm 1 behaving in Cournot fashion towards market B in 1996, since (perceived) marginal revenue $p + \frac{p}{\eta}.\frac{q_f}{q} = 9.46$ considerably exceeds marginal cost 5.47, i.e. $\varphi_{1B} = 9.46 - 5.47 = 3.99$, amounting to 42% of consumer price. Likewise, I reject Cournot behaviour for firm 2 towards market B in 1996: marginal revenue 8.14 considerably exceeds marginal cost 5.30, i.e. $\varphi_{2A} = 2.84$, or 30% of consumer price. Firm 1's (firm 2's) supply decision toward market B (market A) corresponds to that of firm f in the right panel of Figure 2-5 in the theoretical framework of Section 2.2.

Thus Cournot behaviour can in these instances be rejected in favour of more collusive conduct. A story where Votorantim tacitly agrees to give Brennand the upper hand in AL in exchange for the latter staying away from SE would help explain the observed shipments. As noted previously, with a view to testing conduct, the marginal costs I construct in this study are conservative, i.e. they err on the high side. (This understates φ, working against the rejection of Cournot conduct.) In spite of this, the φ are not only positive but sizeable: of the order of 30 - 40% of consumer price! Interestingly, note that Brennand (firm 1) ships from its plant in AL (market A) to the states of PB, PE and BA, located at further distances than SE (market B) and where prices are similar to those in SE.

I now compute the test statistic of Proposition 2.3 for each active firm-market-month combination, indexed as before by f, l and t respectively. A firm is active in a given month if it owns a plant which is active in that month; that is, firm f is active iff $\sum_l \sum_{i \in \mathcal{O}_{ft}} q_{ilt} > 0$ (recall that i indexes plant and $\mathcal{O}_{ft}$ is the set of plants owned by firm f in month t). For every month t in which a firm f is active, there are 17 (f, l, t) combinations, one for each of the 17 markets, irrespective of the markets to which firm f *actually* ships in month t. Now at each month t take firm f's marginal cost in serving market l, c_{flt}, as the minimum among the marginal costs in serving market l from the plants that it owns, i.e. $c_{flt} := \min_{i \in \mathcal{O}_{ft}} c_{ilt}$. From the Cournot pricing condition (2.10), compute the test statistic

$$\hat{\varphi}_{flt} = p_{lt} + \frac{p_{lt}}{\hat{\eta}_{lt}} \frac{q_{flt}}{q_{lt}} - c_{flt} \tag{2.15}$$

where $\hat{\eta}_{lt}$ is based on the demand estimates of Section 2.4.2. Recall that $\hat{\varphi}_{flt} > 0$ is sufficient to reject the null hypothesis that firm f is behaving in Cournot fashion towards market l in month t, in favour of more collusive behaviour; a Cournot firm perceiving marginal revenue to exceed marginal cost would expand output beyond the observed output (and, importantly, recall that this statistic allows for the constraining effect of imports). Notice that p, q and c are observed (or constructed: it is clear from (2.15) that the construction of c as an upper bound to the true marginal cost conservatively tilts the test statistic against rejection of Cournot), such that the randomness in $\hat{\varphi}$ stems from the randomness of the estimated price elasticity $\hat{\eta}_{lt}$.

Total number of active firm-market-month combinations, (f, l, t)	37536	
Number of (f, l, t) combinations for which:		
The upper limit to the 95% confidence interval for η_{lt} is negative	24696	100%
$\hat{\varphi}_{flt}$ is greater than zero	16806	
$\hat{\varphi}_{flt}$ is significantly greater than zero at the 5% level	**14849**	60%
$\hat{\varphi}_{flt}$ is positive and exceeds 10% of consumer price	13197	
$\hat{\varphi}_{flt}$ is positive and exceeds 20% of consumer price	**8035**	33%
$\hat{\varphi}_{flt}$ is positive and exceeds 30% of consumer price	3258	
$\hat{\varphi}_{flt}$ is positive and exceeds 40% of consumer price	504	

Total number of active firm-market-month combinations, (f, l, t)	37536	
Number of (f, l, t) combinations for which:		s.t. $\hat{\varphi}_{flt} > 0.2p_{lt}$
$\hat{\varphi}_{flt}$ is positive when $\hat{\eta}_{lt}$ is taken as -0.3	25335	12757
$\hat{\varphi}_{flt}$ is positive when $\hat{\eta}_{lt}$ is taken as -0.5	27270	14237
$\hat{\varphi}_{flt}$ is positive when $\hat{\eta}_{lt}$ is taken as -0.7	28665	15575

The table above summarises the results. There are 37536 active firm-market-month combinations (corresponding, therefore, to an average of $37536/17/156 \approx 14$ active firms across the country in any given month). Since I calculate the 95% confidence interval (C.I.) for φ_{flt} from the 95% C.I. for the price elasticity η_{lt}, I choose to drop 12840 observations for which the upper limit to the C.I. for η_{lt} is positive. In other words, I conservatively consider only the 24696 combinations for which the C.I. for the price elasticity lies in the interval $(-\infty, 0)$. I find that the null hypothesis of Cournot behaviour allowing for the constraining effect of imports, $\varphi_{flt} \leq 0$, can be *rejected at the 5% level of significance in 14849 of these 24696 combinations.* Put differently, under the Cournot conjecture, one would expect firms to expand their supply to local markets in $14849/24696 \simeq 60\%$ of monthly supply decisions vis-à-vis observed outputs – these firms are choosing output to the left of their Cournot reaction functions. As in the earlier illustration, the test statistics $\hat{\varphi}_{flt}$ are not only positive but sizeable: the point estimate for $\hat{\varphi}_{flt}$ exceeds 20% of consumer price in 8035 supply decisions! To check robustness, the table also provides the number of supply decisions for which $\hat{\varphi}_{flt} > 0$ when $\hat{\varphi}$ is calculated using elasticities of -0.3, -0.5 or -0.7. The rejection of Cournot behaviour in favour of more collusive conduct is robust. It is clear from (2.15) and from the table that a market price elasticity of demand greater (in absolute value) than the estimated -0.5 on average reinforces this result. I further comment on the robustness of this conclusion in Chapter 3.

2.5 Concluding remarks

In this chapter, I show that when an industry faces potential entry and this threat of entry constrains pre-entry prices, cost and conduct will not be identified from the comparative statics of equilibrium. The well-established technique of relying on the jumping around (i.e. shifting *and* rotating) of the demand curve to distinguish empirically between the hypothesis of a high-cost competitive industry and the alternative of a low-cost oligopoly with market power fails in the event that the low-cost oligopoly is constrained by potential entry. Because the extent to which the no-entry constraint binds may be unobserved to the researcher, since in equilibrium entry is *not* observed, some studies of market power claiming to consistently estimate structural parameters of the data generating process may actually be underestimating the degree of market power. A typical situation of such limit pricing is that where a domestic oligopoly faces a competitive fringe of foreign suppliers.

This result recommends caution to enthusiasts of the estimation of market power when cost data is lacking, and has important implications for antitrust practitioners, particularly in a world where trade barriers are being pulled down. This chapter provides an additional theoretical setting to a criticism advanced by Corts (1999) regarding the non-robust performance of static structural estimation in the absence of cost when applied to a dynamic model. In an imports-constrained setting, I show one way in which a researcher, equipped with cost and firm-level data, may delve deeper into the pattern of conduct in an industry where firms meet in different markets.

The Brazilian cement industry provides a clear-cut illustration. An elastically-supplied fringe of high-cost imports restrains domestic prices and the market price elasticities of demand are of the order of -0.5 in equilibrium. Despite this binding constraint, the behavioural model generating outcomes in the industry is identified from the observed data. While the estimation of a static pricing equation incorrectly points to competition, I show that conduct is more collusive than the Cournot benchmark, characterised by (tacit) market division. I provide examples of simple dynamic multimarket games which give rise to such equilibrium behaviour. The price limit set by the delivered cost of imports would seem to provide a natural focal price; market division would further enhance collusive pay-offs by limiting cross-hauling.

A clear policy recommendation emanating from the illustration is the finding that producers possess substantial market power and that imports (in the form of cement or the intermediate product clinker) have an important role to play in curbing the ability of domestic producers to further raise prices above marginal cost. To the extent that investments may be made to reduce transaction (entry) costs of imports, one should not necessarily expect to observe an increased share of imports but certainly expect higher consumer welfare in the form of lower prices. Such a recommendation stands in direct contrast to recent policy experience, whereby the cement industry successfully managed to lobby the government into enacting "antidumping" measures against foreign producers who were attempting to make inroads into Brazil's local markets.

Chapter 3

Conduct in the Brazilian Cement Industry: Verifying Robustness Further

3.1 Introduction

In the preceding chapter I showed that the standard methodology for inferring the degree of market power exercised by firms in an industry – in the sense of consistently estimating the true price-cost margins – fails in a general setting where the industry faces the threat of entry and this threat of entry constrains pre-entry prices[1]. Intuitively, the threat of entry dampens the responsiveness of prices to exogenous demand perturbations on which the standard comparative-statics technique rests, removing the source of price variation which allows conduct to be identified. I illustrated this theoretical result by reference to the Brazilian cement industry, where potential (high-cost) imports restrain market outcomes. That is, imports impose a price ceiling which binds at the equilibrium, such that no or few imports are observed. While the standard methodology yielded supply estimates that are indicative of domestic competition – with estimated price-cost margins centred around zero – the true price-cost margins are far from competitive: com-

[1]As long as a non-trivial subset of the observations are *constrained*, this misspecification will lead the analyst to underestimate the degree of market power, finding more competition when there is less.

putation of plant-to-market marginal cost pointed to actual margins amounting to 50% of producer prices. I then went on to propose a test of conduct adapted to such settings where market outcomes are constrained. This test was illustrated again by reference to the Brazilian cement industry. Outcomes in the industry were found to be more collusive than the Cournot benchmark, and can be characterised by tacit market division. In addition to illustrating the proposed test of conduct, such an exercise reiterates the rejection of competitive domestic behaviour, as suggested by the standard methodology when the latent effect of entry is overlooked.

The present chapter's recurring theme is robustness. It explores the robustness and validity of several different theoretical and estimation-related modelling aspects and findings which concern Chapter 2. I begin by considering, in light of the different local market structures observed in the Brazilian cement industry, simple dynamic multimarket games which in equilibrium give rise to the pattern of conduct identified previously. As seen, a story where firm 1 tacitly agrees to give firm 2 the upper hand in market B in exchange for the latter staying away from market A – with typically firm 1 (firm 2) being located slightly closer to market A (market B) than the rival firm – helps to explain the observed shipments. Plants ship to local markets located at their doorstep, while restricting supply to adjacent markets, despite supplying to these latter markets being highly profitable under the static Cournot conjecture. My aim in providing examples of such dynamic games is to indicate the rationality of a strategy of market division in an industry where firms meet in different (geographic or product) markets. Recall that the price limit set by the delivered cost of imports may provide a natural focal price[2]. And by devising strategies which take multimarket contact into account, firms can achieve complete market division, limiting cross-hauling and maximising aggregate profit.

I then turn to the spatial competition literature in the quest to rationalise the very low estimated market price elasticities of demand in equilibrium – of the order of -0.5 across Brazil's local markets. The question posed previously is why does a producer, seeing such inelastic market demand, not cut output in an attempt to

[2] This is consistent with the Eastman-Stykolt (1966) hypothesis – recall footnote 22 of Chapter 2.

raise prices, moving up along the demand curve to a point where demand is more elastic[3]? As explained in Chapter 2, two main possibilities arise to rationalise why a producer does not cut output. The first rationale is that there is competition among incumbents, such that if producer 1 cuts output in an attempt to raise prices, along comes rival producer 2 and undercuts his price. Thus a low *market* price elasticity of demand does not imply a low elasticity of demand faced by *each individual producer.* (Alternatively, one may think of, for example, a Cournot oligopoly with low concentration such that each individual firm internalises only a small proportion of the aggregate benefit of the large price rise that would be brought about by a small reduction in output; thus in equilibrium the price remains at a level consistent with aggregate demand being inelastic.) The second rationale is the one I accept in Chapter 2, that some dynamic story is appropriate, such as the threat of entry (imports) restraining pre-entry prices. One must emphasise that my claim that the domestic industry is constrained by the threat of foreign entry rests on the low market price elasticities of demand in equilibrium[4]. By this rationale, while *market* demand is inelastic at the equilibrium, the demand faced by *incumbents* (domestic producers) in aggregate is elastic – cutting output in an attempt to raise prices above the ceiling c_I simply opens the door to imports[5]. A third possibility, however, arises in the literature on spatial competition, providing a potential substitute (or complement) to the accepted theoretical framework, that a competitive fringe of imports restrains the domestic industry's outcomes. In the cement industry, transportation assumes a significant proportion of cost and firms' offerings are spatially differentiated. Any plausible theory should satisfactorily

[3]One may correctly point out that the demand curve for cement is very steep. (There are essentially two reasons for this. One is that with the exception of highway construction – where highways may be paved either with concrete or with asphalt – cement has no clear substitutes. The second reason is that, in spite of its importance, cement makes up only a moderate proportion of construction budgets.) Now, a steep demand curve still does not explain why the elasticity of demand is so low at the market equilibrium. Why is price (and the elasticity) not higher?

[4]While the finding that demand is inelastic in equilibrium is central to my claim that the domestic industry is constrained by the threat of imports, this claim is backed up by a wealth of interview and anecdotal evidence in addition to being consistent with stylised facts, such as prices being largely uniform along the coastline, increasing in the distance from the coast and highly correlated with the exchange rate. See Section 2.3.2.

[5]As such, in the case of the Brazilian cement industry, market demand is inelastic at the price ceiling set by potential imports. However, in other cases it could well be that the threat of entry binds at a price level where market demand is *elastic.* In such cases it would arguably be more difficult to claim that the threat of entry has a binding effect on industry outcomes.

reconcile the low market price elasticity of demand with the exercise of market power by firms. My motivation stems from the equilibrium in a well-known example of spatial competition, the Hotelling-type circular road model. In this model, inelastic market demand does not imply low price-cost margins (since inelastic market demand does not translate into inelastic demand for the firm). As I show, however, this result of positive price-cost margins even at local markets where spatially-differentiated firms' market boundaries meet owes to a very special class of spatial pricing policy assumed in the model, that a firm is restricted to set only a "mill" price, with delivered prices to consumers who are distributed over space being equal to the sum of this mill price and the transportation cost. This is not the case in the Brazilian cement industry, where firms set prices at the local market level (and these are observed to be quite uniform over space, including those local markets where many firms meet). I conclude that models of spatial competition do not provide further insight over the theoretical framework proposed presently.

It is clear that an empirical researcher must also consider the possibility – given their central importance – that the equilibrium demand elasticity estimates are inconsistent. Their low values, of the order of -0.5, could make one worry that they are indeed underestimates of the "true" elasticities. I thus proceed to performing further robustness checks of the demand estimation of Chapter 2. In addition to the estimation results under different sets of instruments presented in that chapter, here I consider: (i) estimating the state-level demand equations simultaneously, allowing the unobserved demand shocks across states to be correlated; (ii) using fixed effects instrumental-variables panel data estimation, specifying each state as a unit and calculating clustered standard errors (i.e. clustering the observations pertaining to a given state). Notice that, given the vastly larger number of observations (156 months $\times$ 17 states as opposed to 156 months in each state-level estimation of the preceding chapter), this specification increases efficiency at the expense of cross-unit restrictions; (iii) fitting alternative functional forms, such as semi-log-linear and linear, as opposed to the log-linear specification (2.12); (iv) reversing the dependent variable; (v) considering the possibility that instruments are weak; and (vi) considering a dynamic demand function. The overall conclusion arising from these checks is that the preceding chapter's finding – that demand is inelastic (of the order of -0.5) at the equilibrium in each local market – is robust.

Finally, in the spirit of the trade literature, I resort to a *gravity model* in order to statistically analyse the flow of cement between plants, owned by the different cement producers, and local markets (identified with states). The trade literature has extensively used the gravity model to explain the flow of trade and investment between countries. Feenstra (2002) describes the gravity equation as "perform(ing) extremely well empirically" (p.491), while Evenett and Keller (2002) affirm that the equation constitutes "one of the most important results about trade flows" (p.282). Of greater interest than merely explaining the observed flows, or shipments, based on market size, plant location and plant ownership, lies the "detection" of outliers in the data. One example could be the absence of shipments from a given plant (and thus firm) to a given market where one would expect them to occur on account of the plant's location and ownership alone. Barring strong price effects, one could speculatively attribute the presence of such outliers to firm behaviour, as well as institutional detail. It must be emphasised that this exercise is not structural, in that it does not follow from a model of shipments derived from a specification of demand and cost conditions, firm entry and behaviour – this structural approach was pursued in Chapter 2. Here I perform a statistical benchmarking exercise. Rather than arbitrarily establishing the distance over which cement can profitably travel (say 500 km[6]) and from there analysing the pattern of shipments, the merit of this approach is that of letting the data pinpoint the firm-market pairs, if any, which do not conform to an "average" pattern.

To illustrate why this exercise is not structural, consider a hypothetical situation where firms collude by dividing markets such that each firm ships only to customers located in the vicinity of its plants[7]. All the gravity estimation would pick up is a very high coefficient in the "distance" variable, yet one is clearly not identifying distance effects from conduct. And in this hypothetical situation, it may well be

[6]The geographic area relevant to each plant's supply decision (i.e. the plant's "market") is obviously endogenous – it depends on demand and supply conditions. See footnote 3 in the Appendix to Chapter 2.

[7]One would think that the proximity of local markets from a plant (i.e. the plant's shipment radius) may be a natural instrument for coordination, to the extent that each plant's radius may be monitored by other firms.

that there are no outliers to this "average" pattern, if there is "sufficient" symmetry in some sense. But this illustrates a point which, to the best of my knowledge, has not been made in the trade literature's use of gravity models: *that estimated distance effects may also be proxying for firms' strategies of dividing (geographic) markets*, as is the case in the Brazilian cement industry (if not at the supranational level, one could plausibly speculate; see, for example, Ghemawat and Thomas 2004). This is yet another manifestation of the classic identification problem in supply: distinguishing firm costs from firm conduct. Head (2003), for example, lists alternative explanations that have been advanced in the literature to explain why distance is found to matter so much, such as transport costs, perishability, sourcing "synchronisation", communication and transaction costs, and cultural differences. One would clearly wish to add industry behaviour to this list, as the exercise applied to the cement industry makes clear. As shown in Chapter 2, transport costs defined in a broad sense go only a limited way in explaining the observed pattern of shipments.

Indeed, I find that the gravity equation fits Brazilian cement shipments very well. Interestingly, while the estimated elasticities of bilateral trade with respect to distance in the trade literature vary between -0.8 and -1.5% (see, for example, Loungani et al 2002), the distance elasticity in the case of Brazilian cement flows amounts to as much as -3%. Estimated state border effects are also very high, suggesting that state borders may be playing a non-negligible role in the (tacit) coordination of firms' strategies. As for the "detection" of outliers, this statistical exercise flags some odd situations in the data. Such is the case, for example, of the shipments by three of the four leading firms (Votorantim, Holcim and Lafarge – see Figure 2-6) to the three largest local markets (São Paulo, Minas Gerais and Rio de Janeiro – see Figure 2-14). The leading firm Votorantim commands a 50% market share approximately in the largest market *SP*, while commanding less than a 10% share of shipments to the second largest market *MG*, and a 20% share in the third largest market *RJ*[8]. In contrast, both Holcim and Lafarge command shares of 25% each in *MG* and *RJ*, but account for less than 10% of shipments

[8] Of note, Votorantim has plants both in *MG* and *RJ* (recall Figure 2-6).

to *SP*. These anomalies in the data – in the sense that they do not conform to the "average" patterns of shipments, when one controls for distance (which is also proxying for conduct, as per the above discussion), plant location and ownership – are clearly picked up by the gravity model. Actual shipments significantly differ from shipments predicted on the basis of plant location and ownership. There may well be institutional and/or historical reasons behind the existence of these outliers, yet Votorantim's shying away from the large *MG* market is consistent with a market division or multimarket contact story à la Bernheim and Whinston (1990).

The plan of the chapter is as follows. In Section 3.2 I provide simple dynamic multimarket games which in equilibrium give rise to market division. Section 3.3 turns to the spatial competition literature. I then consider the robustness of the demand estimation of Chapter 2. Finally, Section 3.5 estimates a gravity model.

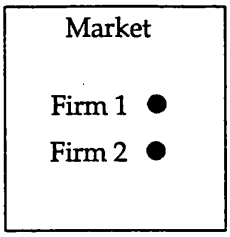
3.2 Collusion and market division

The cement industry is commonly used to illustrate industry characteristics that the literature on tacit collusion has deemed to enhance the likelihood of collusion[9]: see Appendix C for a summary. Unsurprisingly, the cement industry has a long history of anticompetitive behaviour and antitrust litigation across several jurisdictions[10]. Tacit collusion in the Brazilian cement industry, orchestrated for instance via market division, is a concrete possibility in that the characteristics of the industry are consistent with the characteristics which are understood to make tacit collusion more likely. For the sake of illustration, I now turn to some simple collusive arrangements – relevant to the case at hand – which may be supported in equilibrium.

[9] As before, I refer to collusion in prices (or quantities), though I later comment on collusion in capacity investments.

[10] See, for example, Dumez and Jeunemaître (2000). Ghemawat and Thomas (2004) cite the fines imposed by the European Court of First Instance in 1994 on 42 cement-related undertakings across Europe, in what has been one of the EU's largest competition cases to date.

Illustration 1: two firms with plants located in a single local market (*A local market which comes to mind here is Rio Grande do Sul (*RS*), where two firms, Votorantim and Cimpor, operate plants, located very close to one another.*)

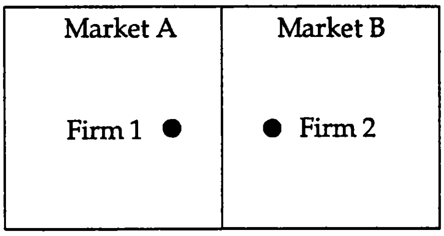


Consider two firms, 1 and 2, with symmetric marginal costs c facing a competitive fringe of imports with marginal cost c_I such that $c < c_I < p^M$, where p^M is the monopoly price as in the theoretical framework of Section 2.2. Recall that in such a setup the most collusive price, which maximises aggregate (domestic) industry profit, is $p = c_I$, where the aggregate industry profit (per period) is $\Pi := (c_I - c)p^{-1}(c_I)$ and $p^{-1}(.)$ denotes the demand function. As is standard in the literature on supergames, the most collusive price may be supported in equilibrium by each firm adopting, say, the following symmetric "grim" strategy in prices: set the collusive price $p = c_I$ in each period unless the rival firm has set a different price in a previous period, in which case set the competitive price $p = c$. Assume for now that, given the symmetry, both firms split the market equally (i.e. firm f's share $s_f = \frac{1}{2}$, $f = 1, 2$). Collusion will then be sustainable if each firm's payoff from sticking to the collusive agreement exceeds its payoff from slightly undercutting its rival and selling to the entire market in a single period; that is, collusion is sustainable if

$$\frac{1}{2}\Pi\frac{1}{1-\delta} \geq \Pi$$

where δ is the per-period discount factor. This incentive constraint rearranges to $\delta \geq \frac{1}{2}$, yielding the standard folk theorem whereby for a sufficiently high discount factor the collusive price $p = c_I$ (or any other price p such that $c < p \leq c_I$) may be supported in equilibrium.

Illustration 2: two firms and two neighbouring markets: firm 1 is located in market A and firm 2 is located in market B next door (*Two local* *markets which come to mind here are the neighbouring states of Sergipe (*SE*) and Alagoas (*AL*), equally small in terms of market size and geography. As stated previously, until 1996 Votorantim operated the only plant in* SE *and Brennand operated the only plant in next-door* AL. *While in 1996 Votorantim commanded an 89% share in* SE, *it attained only a 7% share in* AL*; on the other hand, Brennand commanded an 83% share in* AL *while not supplying to* SE.)



Now assume that there are transport costs $t > 0$ associated with serving the neighbouring market: while the marginal cost of serving a market in which one's plant is located is c (e.g. firm 1 in market A), the marginal cost of serving the market next-door to one's plant rises to $c' = c + t$ (e.g. firm 1 in market B), where $c < c' < c_I$ (c_I as before). Demand is identical in each market, given as before by $p^{-1}(.)$. Denoting $\Pi := (c_I - c)p^{-1}(c_I)$ as before, let $\Pi' := (c_I - c')p^{-1}(c_I)$.. (This corresponds to firm 1's maximal profit in neighbouring market B were it to act as a monopolist in the supply of that market[11]. Clearly $\Pi > \Pi'$.) It is easy to see that the collusive arrangement that maximises aggregate industry profit in each market involves no wasteful cross-hauling and corresponds to complete market division, where each market is supplied only by the low-cost firm and prices in each market are $p = c_I$: firm 1 supplies quantity $p^{-1}(c_I)$ to market A (i.e. shares $s_{1A} = 1$, $s_{2A} = 1 - s_{1A} = 0$) and firm 2 supplies quantity $p^{-1}(c_I)$ to market B (i.e. $s_{2B} = 1$, $s_{1B} = 1 - s_{2B} = 0$). But can this arrangement be supported in equilibrium?

Begin by considering a situation where firms devise strategies *that treat each market separately*, that is, cheating in a market does not trigger retaliation in other markets. Then for the most collusive price $p = c_I$ (or any collusive price p above the competitive price, equal to the high-cost firm's cost c', but lower than c_I, i.e.

[11]Realise that implicit in this statement is the regularity assumption that the monopoly price $p^M(c) := \arg\max_p (p - c)p^{-1}(p)$ is an increasing function of marginal cost c. Thus $p^M(c') > p^M(c) > c_I$, where $p^M(c) > c_I$ is as before, such that firm 1 acting as a domestic monopolist in the supply of neighbouring market B and facing the fringe of imports, would set $p = c_I$.

$c' < p \leq c_I$) to be sustainable in a given market, the collusive arrangement must prescribe a strictly positive share to *both* firms in that market. In this situation, both firms must enjoy a non-trivial share of the collusive pie in each market, as can be seen by each firm's incentive constraint (IC) in, say, market A[12]:

$$s_{1A}\Pi\frac{1}{1-\delta} \geq \Pi + \Pi^{war}\frac{\delta}{1-\delta} \qquad \text{low-cost firm 1's IC in market A}$$

$$(1 - s_{1A})\,\Pi'\frac{1}{1-\delta} \geq \Pi' \qquad \text{high-cost firm 2's IC in market A}$$

where $\Pi^{war} := (c' - c)p^{-1}(c')$ denotes industry profit under retaliation (price war, when $p = c'$), earned by the low-cost firm. (Notice that the minimum share to be prescribed to the high-cost firm is higher the lower is the discount factor: simply rewrite the high-cost firm's IC as $1 - s_{1A} \geq 1 - \delta$.) Hence if firms devise strategies that treat each market separately, complete market division ($s_{1A} = s_{2B} = 1$) cannot be sustained in any collusive equilibrium, regardless of the discount factor.

More naturally, firms will devise strategies that take into account the multimarket nature of their contact, since in each market A or B the same two firms 1 and 2 can supply. By modifying each firm's strategy to ensure retaliation is triggered (i.e. setting price equal to the competitive price c') in *both* markets should any firm undercut the collusive price in *any* market in a previous period, the collusive arrangement that maximises aggregate industry profits across markets – i.e. setting $p = c_I$ with complete market division – can now be supported in equilibrium for a high enough discount factor. To see this, pool each firm's incentive constraints across both markets; firm 1's (say) IC is now

$$s_{1A}\Pi\frac{1}{1-\delta}+(1 - s_{2B})\,\Pi'\frac{1}{1-\delta} \geq \Pi+\Pi'+ \Pi^{war}\frac{\delta}{1-\delta} \qquad \text{firm 1's pooled IC} \quad (3.1)$$

Assuming that the collusive arrangement involves the low-cost firm in one market commanding the same share as the low-cost firm in the other market (i.e. $s_{1A} =$

[12]Note that firm 1 and firm 2's shipments are then respectively $q_{1A} = s_{1A}p^{-1}(c_I)$ and $q_{2A} = (1 - s_{1A})\,p^{-1}(c_I)$.

$s_{2B} = s$) (3.1) can be rearranged to yield firm f's ($f = 1, 2$) incentive constraint:

$$\delta \geq \frac{(1-s)\,\Pi + s\Pi'}{\Pi + \Pi' - \Pi^{war}} \tag{3.2}$$

The collusive arrangement can now involve complete market division, $s = 1$. Indeed, for $p = c_I$, setting $s = 1$ in (3.2) minimises the discount factor threshold above which collusion is sustainable[13]:

$$\delta \geq \frac{\Pi'}{\Pi + \Pi' - \Pi^{war}} \tag{3.3}$$

Thus for a high enough discount factor, the collusive scheme that maximises aggregate industry profit in each market, setting $p = c_I$ in both markets and completely dividing markets, can now be supported in equilibrium. Intuitively, as Bernheim and Whinston (1990) have shown, through multimarket contact "slack enforcement power" may be shifted from the market where a given firm is located, enjoying low cost, high share and high profit, to the neighbouring market where that firm has high cost, low share and low profit[14].

[13]Recall that $\Pi > \Pi'$. Intuitively, when $s = 1$ the short-term gain from cheating (equal to $(1-s)\Pi$ in the own market plus $s\Pi'$ in the neighbouring market) is lowest (and equal to Π'). Two comments are in order. First, one can show that increasing the transportation cost t may increase the discount factor threshold: though the deviant's profits in the period of deviation fall since Π' falls, profits in each later period rise since Π^{war} rises, meaning that the long-term loss of collusive profits from retaliation may become lower. Thus collusion would seem less likely as t rises, given that it makes the incentive-constraint more stringent (for $s = 1$). However, as mentioned in the footnotes to the table on industry characteristics in Appendix C, the *profitability* of collusion may be increasing in t, since the effect of eliminating wasteful cross-hauling, through market division, on profits may now be larger. Hence, a greater payoff from collusion through higher t, conditional on it being sustainable, would suggest that firms would have greater incentive to design and implement a collusive scheme (see Ivaldi et al 2003, footnote 48, on a similar idea). Second, when $t = 0$, then $c' = c$, $\Pi' = \Pi$ and $\Pi^{war} = 0$, implying that incentive constraints (3.2) and thus (3.3) collapse to the familiar $\delta \geq \frac{1}{2}$ of Illustration 1.

[14]Notice that I have assumed Bertrand behaviour should collusion break down, but could just as well have assumed Cournot competition. In this case, the right-hand side of incentive constraint (3.1) would have to be modified as follows: (i) a deviant firm would now earn less than $\Pi + \Pi'$ in the period of deviation (the other firm would still supply its collusive quantities in the period of deviation, and the deviant firm would set its output in each market based on its reaction function, as illustrated in Figure 2-5, thus expanding output in the neighbouring market where it has a low share), and (ii) upon retaliation either firm would also earn positive payoffs in its neighbouring market, where it incurs a higher cost (and, as Figure 2-5 makes clear, the Cournot equilibrium outcome would not necessarily be unique).

Illustration 3: three firms and two neighbouring markets: firms 1 and 2 have plants located in market A and market B next door, while firm 3 has a plant located in market B only (*Two local markets which come to mind here are the neighbouring states of Rio de Janeiro (*RJ*) and Minas Gerais (*MG*). Considering these two markets, four firms have plants located in both* RJ *and* MG*, while two firms have plants located in* MG *only. The extent to which these two latter firms cross-haul cement from their plants in* MG *to the* RJ *market is limited: in 1999 Camargo Correa commanded a 20% share in its home market* MG *but did not supply to the neighbouring* RJ*, while Soeicom had a 9% share in its home market* MG *and a (somewhat) lower 6% share in neighbouring* RJ.)

Market A	Market B
Firm 1 ●	● Firm 1
Firm 2 ●	● Firm 2
	● Firm 3

As before, assume that the cost of supplying a market from a plant located in that market is c but rises to $c' = c + t$ when supplying from a plant located in the neighbouring market. Demand $p^{-1}(.)$ is again identical in each market, recall $p^M(c') > p^M(c) > c_I$ and denote Π' and Π as before. Now consider the following collusive agreement: (i) in market A each of firms 1 and 2 supplies a share s_A of the market, with firm 3 accounting for the remaining $(1 - 2s_A)$ share, such that price is $p = c_I$; and (ii) in market B firm 3 supplies a share s_B of the market, with firms 1 and 2 accounting for the remaining $(1 - s_B)/2$ each, again such that price is $p = c_I$.[15] The (pooled) incentive constraint for each of firms 1 and 2, that operate plants in both markets, becomes:

$$\left(s_A\Pi + \frac{1 - s_B}{2}\Pi\right)\frac{1}{1-\delta} \geq \Pi + \Pi$$

[15] In market A, firms 1 and 2 will then supply $q_{1A} = q_{2A} = s_A p^{-1}(c_I)$ and firm 3 will supply $q_{3A} = (1 - 2s_A)p^{-1}(c_I)$. In market B, firm 3 will supply $q_{3B} = s_B p^{-1}(c_I)$ and firms 1 and 2 will each supply $q_{1B} = q_{2B} = \frac{1}{2}(1 - s_B)p^{-1}(c_I)$.

since a deviation triggers the competitive price $p = c$ in both markets, which can be rearranged to

$$\delta \geq \frac{3 - 2s_A + s_B}{4} \tag{3.4}$$

The incentive constraint for firm 3, with a plant in market B only, is

$$((1 - 2s_A)\,\Pi' + s_B\Pi)\,\frac{1}{1-\delta} \geq \Pi' + \Pi$$

which is equivalent to

$$\delta \geq \frac{(1 - s_B)\Pi + 2s_A\Pi'}{\Pi + \Pi'} \tag{3.5}$$

As in illustration 2, the collusive arrangement that maximises aggregate industry profit across markets involves no wasteful cross-hauling, where each market is supplied only by the low-cost firms (i.e. $2s_A = 1$ such that firm 3 does not supply to market A) and prices in each market are $p = c_I$. Plugging $s_A = 1/2$ and, say, $s_B = 2/3$ such that all three firms produce the same quantity, sustainability constraints (3.4) and (3.5) become $\delta \geq \frac{2}{3}$ and $\delta \geq \frac{\Pi/3+\Pi'}{\Pi+\Pi'}$ respectively.

Capacity and collusion Some remarks about the role of capacity in the Brazilian cement industry. As noted in Section 2.3.3, capacity significantly exceeds production, and this (low) *capacity utilisation* appears to be fairly symmetric across plants and firms, despite the asymmetric capacities across these plants and firms. Thus, for instance, plant 1 with a capacity of 2 mtpa may be running at a 65% capacity utilisation while plant 2, owned by a rival firm, with a capacity of 1 mtpa may be operating at the same 65% capacity utilisation. This observation is consistent with a situation where *all* domestic producers adhere to the collusive arrangement, with no producer "free riding". (This is reinforced by the fact that all producers are long-time members of the cement producers' trade association (SNIC), an active lobbying outfit for the industry.) The corollary to this observation is that there appears to be no relevant fringe to the (tacit) cartel[16]. Further, evenly-distributed

[16]More recently, an entrant, Mizú, has successfully managed to establish a foothold in local markets in and around the state of Espírito Santo (*ES*), where it is based. Set up in 1998 by a large independent ready-mixed concrete firm, Mizú signed a long-term contract with a steel producer

idle capacity across firms would serve the key purpose of disciplining the cartel: the threat of punishment would not be credible were capacity to be tight[17] [18].

3.3 Models of spatial competition

The question which arises is how to reconcile such low market price elasticity of demand with the presence of market power by firms. In Chapter 2 the low elasticities in equilibrium in each of the different local markets – including those where the one-firm concentration ratio is as high as 80% – coupled with the high observed price-cost margins were taken as evidence that the potential entry of imported cement restrains prices set by the domestic industry. While *market* demand in equilibrium is inelastic, the *residual* demand which the domestic industry faces at the price ceiling posed by high-cost imports is elastic. Attempts by the domestic industry, already enjoying a large price-cost margin, to raise prices above this price ceiling would only invite foreign entry.

I now turn to the spatial competition literature to check whether a plausible substitute or complement to this reasoning can be found. Again, any plausible theory should satisfactorily reconcile the low market demand elasticity with the exercise of market power (i.e. positive price-cost margins). My motivation owes to the equilibrium in a well-known example of spatial competition, the Hotelling-type circular road model. In this model, inelastic market demand does not imply low

(Companhia Siderúrgica de Tubarão) to supply it with steel slag, which it grinds and mixes with ground clinker, imported from as far as Japan, producing slag cement. Mizú's (grinding-only) plant is located conveniently next door to the steelworks and to the port of Vitória. In contrast to established cement producers, by 2003 Mizú was selling *up to capacity* (0.7 mtpa). It would appear that the established producers have accommodated Mizú's entry, given its limited capacity and the irreversibility of its investment. (As noted in Appendix B, consumption and shipment figures compiled by the cement producers' trade association, and used in this thesis, do not consider Mizú. The distortion nevertheless is small in view of Mizú's (to date) limited capacity, limited geographic scope and recent entry.)

[17] A further strategic role may be that of helping deter entry, as studied in the literature (e.g. Dixit 1980).

[18] A final comment regarding the possibility that producers collude in capacity investments, in addition to colluding in product market outcomes. Rather than restricting capacity and hence output, producers overinvest in capacity, as just discussed. Other characteristics of the cement industry, such as the lumpiness, infrequency, long life and irreversibility of investment, would further suggest that the scope for collusion in capacities is limited. See Ivaldi et al (2003) for a discussion.

price-cost margins, since inelastic market demand does not translate into inelastic demand for the firm.[19]

3.3.1 Price-setting firms in the circular road model

Consider the circular road model in its simplest form. Consumers are uniformly distributed along a unit circle, transport costs (incurred by the consumer) per unit of distance are given by t, and demand is uniform (each unit mass of consumers demands 1 unit from the firm with the lowest delivered price, given by the sum of the mill price and the transport cost). There are N firms evenly spaced out, producing with zero marginal cost and each firm setting a mill price. Due to the simplifying assumption of uniform demand, it is clear that the market price elasticity of demand is zero, $\eta^{market} = 0$. A firm with (mill) price p facing competition from its two neighbours, each setting a price $\bar{p}$, will face an elasticity of demand given by[20] $\eta^{firm}(p, \bar{p}; t) = \frac{-Np}{t+N(\bar{p}-p)}$, which is increasing (in absolute value) as trans-

[19] To elaborate on the motivation behind this enquiry into the literature on spatial competition, assume away the threat of entry and begin by considering the conventional firm pricing equation (2.13) of Chapter 2, which nests alternative models of conduct. It is conceivable that in an industry where transportation assumes a significant proportion of costs and firms' offerings are spatially differentiated, supply relation (2.13) may correspond to an underparameterised model, where the spatial dimension is not appropriately captured. By this tentative argument, the special ("spaceless") form of pricing equation (2.13) and of the behavioural models nested in it, such as the homogeneous-goods Cournot model, would be feeding a low market price elasticity of demand into too low a price elasticity of demand faced by the firm, and thus inducing the finding of too much competition.

A simple analogy will hopefully make this point clearer. Consider a (brand) differentiated goods industry where firms compete on prices (i.e. Bertrand conduct) and the true substitutability between products is high. A researcher attempts to fit a model of Bertrand competition *imposing* too low a degree of substitution between products (rather than jointly estimating this degree of substitution) and incorrectly rejects this model in favour of price-taking behaviour. The rejection of the true model of behaviour (differentiated Bertrand) in favour of more competition has been brought about by the imposition of too low a degree of substitutition between firms' products (in the form of too low cross-price elasticities). The imposed low cross-price elasticity would be consistent with the firm raising the price of its product above marginal cost; in the absence of it doing so, the researcher incorrectly concludes that the firm is a price-taker.

Note that one can relate this analogy back to the framework of Chapter 2. By overlooking the threat of entry when this has a binding effect on pre-entry prices, a researcher is in effect imposing too low a degree of substitution (between the incumbents and the potential entrants). On assuming away the constraining effect of entry on incumbents, inelastic market demand feeds into inelastic demand facing incumbents. (In so doing, by a similar line of argument to the above, one could be led to reject, say, the hypothesis of Cournot competition in favour of more competition, if Cournot were the true model.)

[20] The distance s from the firm under consideration, setting price p, to its marginal consumer on either side, is given by $s = \frac{1}{2N} + \frac{\bar{p}-p}{2t}$. The firm's demand is thus $q = 2s$; the firm's elasticity of demand is then given by $\eta^{firm} = \frac{dq}{dp}\frac{p}{q}$.

port costs t fall. In the symmetric equilibrium, prices (and price-cost margins) are given by $p^* = \frac{t}{N}$ and the elasticity of demand faced by the firm is $\eta^{firm,*} = -1$ (by construction of equilibrium: recall marginal cost is zero). Notice that inelastic demand at the market level does not translate into inelastic demand for the firm and each firm's price-cost margin is strictly positive for $t > 0$, and increasing in t, despite the market elasticity of demand being zero.

Notice that if the special uniform demand assumption in the example is replaced by a more realistic assumption whereby market demand is decreasing in price, the basic message remains unchanged. For example, assume instead that demand is linear in price: denoting s as the distance from the firm with the lowest delivered price, let demand over the interval $[s, s+ds]$ be given by $f(p,s)ds$, where $f(p,s) = 1 - a(p+ts)$ and $p+ts$ is the lowest delivered price. It is easy to show that the market price elasticity of demand will now depend on the local market's distance from the firm with the lowest delivered price. To illustrate, take $N = 4$, $t = 1$ and $a = 0.1$. In the symmetric equilibrium, the market price elasticity of demand is -0.0253 at the market located at the firm's doorstep ($s = 0$), rising (in absolute value) to -0.0386 for the firm's marginal market (located at $s = \frac{1}{2N} = \frac{1}{8}$). (Analytically, $\eta^{market,*} = \frac{-a(p^*+ts)}{1-a(p^*+ts)}$, where p^* is the equilibrium price). In equilibrium, the elasticity of demand faced by the firm is again $\eta^{firm,*} = -1$ and prices (and price-cost margins) are again non-zero: $p^* = 0.246$.

The circular road model above illustrates that in a context of spatial differentiation a low market price elasticity of demand does not imply low price-cost margins. In the version considered, a strong assumption is that each firm is only allowed to set a (free on board) mill price p_M, with the "delivered price schedule" p_D being given by the sum of this mill price and the travel cost from its plant to the local market located at a distance s, i.e. $p_D = p_M + ts$. One may object to this assumption on the grounds that in many industries firms are able to price discriminate and set a delivered price schedule with a slope[21] other than t. If this

[21]Notice that setting a slope *greater* than t is only possible for the firm if consumer arbitrage is ruled out. The term "price discrimination" is employed as in the spatial literature: a slope strictly less than t is described as price discrimination, in the sense that prices are being discriminated in favour of more distant buyers (also referred to as "freight absorption").

assumption is relaxed, allowing firms to set prices at each local market as opposed to setting only a mill price, competition is now harsher and equilibrium price-cost margins are now lower, but they are still positive. To see this, consider, for simplicity, the case of uniform demand. As indicated in Figure 3-1, a firm's delivered price to a local market (up to its marginal market) will in equilibrium be just short of the transport cost between that local market and the rival firm (recall that marginal cost is zero); its price-cost margin will then be $\frac{t}{N}$ at the market located on its doorstep, falling to 0 at the marginal market (delivered prices are indicated by the thick dashed lines in the figure). While the earlier assumption that each firm set only a mill price ensured a delivered price schedule with slope t (and a constant equilibrium price-cost margin of $\frac{t}{N}$; delivered prices are indicated by the thick smooth lines), allowing firms to price discriminate leads to a delivered price schedule of slope $-t$ (and lower price-cost margins, which are decreasing in the distance from the firm).[22]

Discussion in light of the framework of Chapter 2 The model above in its unrestricted version of spatial pricing policy – where Bertrand firms set prices at each local market as opposed to setting only a mill price – is already nested in the theoretical framework adopted in Chapter 2.[23] (Thus the model is considered as a *candidate* model of behaviour generating outcomes in the industry, yet it is empirically rejected[24].) As for the restricted version adopting the very special class of spatial pricing policy where each firm can set only a mill price, with delivered prices to consumers who are distributed over space being equal to the sum of

[22]See Thisse and Vives (1988). They study a sequential game where spatial duopolists – each firm located at one extremity of a line segment of markets – first choose a spatial price policy (uniform FOB mill pricing or unrestricted price discrimination) and, in a second stage, set prices. Due to the prisoners' dilemma character of the game, price discrimination emerges as the unique equilibrium, despite lower price-cost margins: "(I)t may well be that firms would make more profits by following a uniform price policy. This is so because spatial discriminatory pricing gives more flexibility to a firm to respond to its rival's actions. But then firms may get trapped into a Prisoner's Dilemma-type situation and end up with lower profits due to the intense competition unleashed. Contrary to general belief, uniform (FOB) pricing is therefore *not* evidence of a more competitive environment" (p.124).

[23]Recall, for example, the discussion of Bertrand competition among heterogeneous firms in Section 2.2.3; notice also that pricing equation (2.13) is specified at the local market level, considering the firm's marginal cost of servicing the market, which clearly depends on the location of its plant.

[24]The features of the data listed below recall some elements behind the rejection.

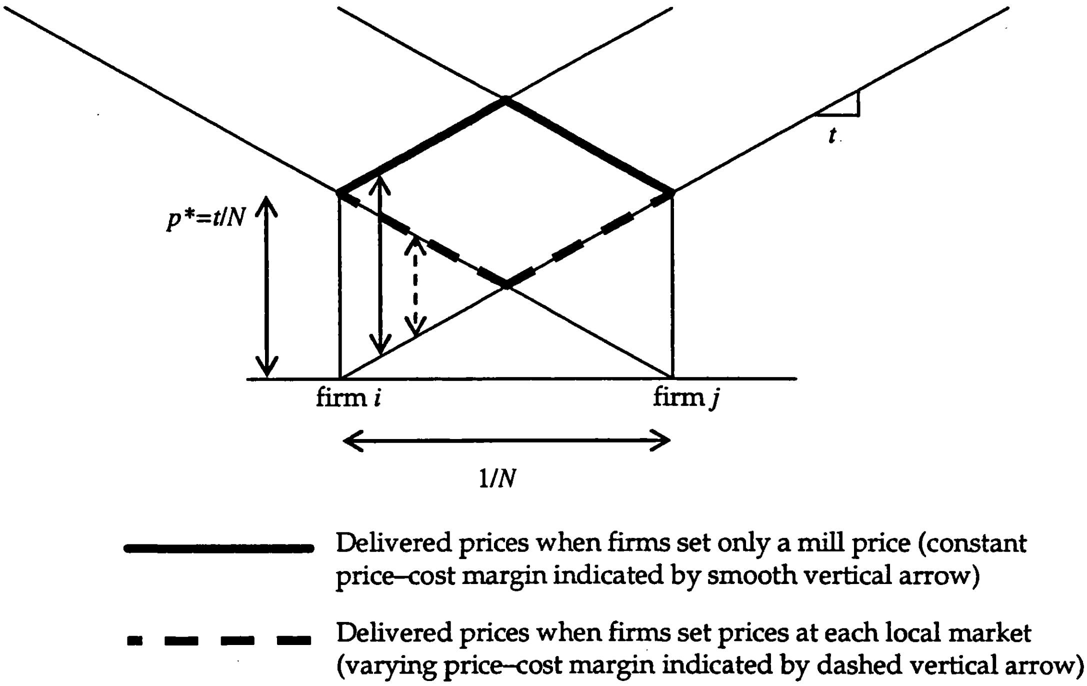


Figure 3-1: Circular road model with uniform demand and price-setting firms: uniform (FOB) mill pricing vs. price discrimination

this mill price and the transportation cost, this is clearly inadequate as a model explaining outcomes in the Brazilian cement industry. Some features of the data are worth recalling briefly: (i) Brazil's cement firms set (delivered) prices at the local market level and are not locked into rigid mill pricing; (ii) prices are quite uniform across Brazil's local (coastal) cement markets, including those markets where many firms meet; (iii) in any given local market, prices considerably exceed the marginal cost of even the least efficient producers serving the market; and (iv) more than one firm typically sells into any given local market[25] (though market shares can be very different), in contrast to the prediction of the spatial competition model above with price-setting firms where at any given local market no more than one firm sells (unless the market is a marginal market). I conclude that no further insight is presently provided over the theoretical framework proposed in Chapter 2. Thus, one is back to requiring some dynamic story as the restraining effect of potential entry to reconcile inelastic market demand and the exercise of market

[25] Often times a given reseller in a given local market will shelve multiple brands (owned by different cement firms).

power.

3.3.2 Quantity-setting firms in the circular road model

I briefly review the circular road model with firms setting quantities, as opposed to prices, in each local market. This is the model of Greenhut and Greenhut (1975)[26]. Similar to before, consumers are distributed along a line segment[27] of length L. Demand at each "point" l is given by the inverse ("spaceless") demand function $p = f_l(q(l))$, where p is the delivered price and $q(l)$ is demand per unit distance (thus revenue in the interval $[l, l + dl]$ is given by $p(q(l))q(l)dl$). Heterogeneous Cournot firms are placed at different points of the line segment. The authors are interested in obtaining the delivered price schedule along the line segment, analysing how it changes (in level and slope) as the number and location of firms, or demand conditions, vary. The equilibrium at each point (local market l) is found by solving a set of pricing equations, one for each firm, as given by equation (2.13), with θ_{fl} replaced by $\frac{q_{il}}{q_l}$ (Cournot behaviour) and $c(q_i, \mathrm{q}_i, .)$ decomposed into a production

[26]McBride (1983) applies this model to the US cement industry in his analysis of vertical integration.

[27]The choice of a line segment, as opposed to a more "realistic" two-dimensional space is again due to analytical tractability, with no loss of generality.

marginal cost MC_i and a transport cost T_{il}:[28] [29]

$$p_l + \frac{p_l}{\eta_l}\frac{q_{il}}{q_l} = MC_i + T_{il} \tag{3.6}$$

The market equilibrium condition can then be found by adding the pricing equation across all firms shipping to a given location (for which there is an interior solution) and dividing by this number of firms (denoted n):

$$p_l + \frac{1}{n}\frac{p_l}{\eta_l} = \frac{1}{n}\sum_{i=1}^{n} MC_i + \frac{1}{n}\sum_{i=1}^{n} T_{il}$$

For the sake of illustration, assume, as Greenhut and Greenhut (1975) do, demand of the exponential form $\frac{p}{\beta} = 1 - \left(\frac{q}{\alpha}\right)^{\gamma}$, defined for all strictly positive values of demand q and parameters α, β and γ. Consider first a situation where there are n_O homogeneous firms (constant marginal cost c) located at the left extremity of the line segment, point O (see the left-hand panel of Figure 3-2). In this case the equilibrium delivered price schedule can easily be shown to be (linear and) increasing in distance from the plant, where an increase in the number of firms n_O leads to lower prices but a steeper schedule. (For simplicity, assume as before a constant unit transportation cost t.) For a finite number of firms, the slope of the delivered

[28]Compared to equation (2.13), the econometric error has also been dropped, time subscripts have been omitted for simplicity, and the equality sign should be replaced by an inequality $\leq$ in the case of a corner solution.

[29]Notice further that while in Chapter 2 pricing equation (2.13), or similarly the Cournot test statistic (2.15), were taken to the data with local markets (on the demand side – note footnote 3 in the Appendix to Chapter 2) defined at the *state* level, the present setup highlights the definition of local markets as separate infinitesimal markets distributed over a continuum in space.

How can one tentatively estimate demand curves at a more disaggregated (local) level (say, the level of *municipality*) when market demand data is observed only at the more aggregated state level? (Note that on average there are 200 municipalities in each state, totalling around 5600 for Brazil as a whole.) Begin by estimating a given state's demand curve from state-specific data. This (inverse) demand curve is the result of horizontally aggregating the inverse demand curves for each municipality, or local market, located in that state. Now assume that the reservation price of the highest-valuation consumer in each municipality is the same across municipalities (i.e. the price intercept is the same). Assume also that as the market size of municipalities (given by their population sizes) varies, the demand curves change multiplicatively (i.e. they swivel around the price intercept). From state S's (inverse) demand curve given by, say, $p = \alpha_S - \beta_S q$, one can back out municipality M's demand curve $p = \alpha_M - \beta_M q$, where $\alpha_M = \alpha_S$ and $\beta_M = \frac{Y_S}{Y_M}\beta_S$, Y_M denoting municipality M's population and $Y_S := \sum_{M \in S} Y_M$ being the total population of state S. It is easy to see that (at a given price) a low elasticity of demand at the state-level implies an equally low elasticity of demand at the municipality level.

The reason why this is not done in Chapter 2 is that on the supply side I observe plants' shipments disaggregated no further than the state level.

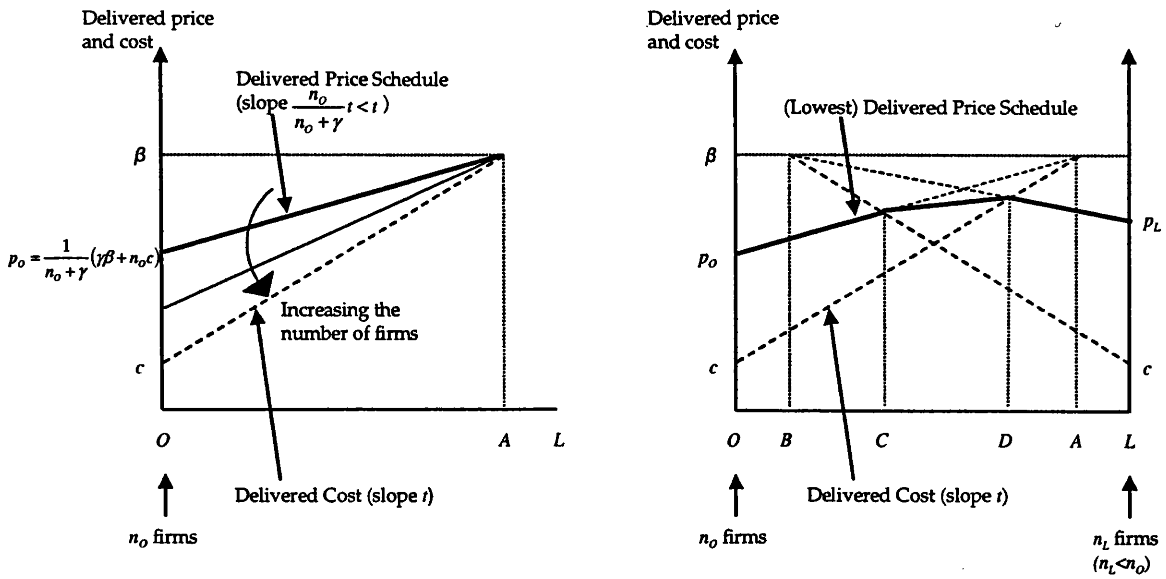


Figure 3-2: Exponential demand along line segment and Cournot firms: price discrimination in equilibrium

price schedule is less than t; in other words, price discrimination (in favour of more distant buyers) takes place in equilibrium. As drawn, no shipments take place to markets beyond point A since the total marginal cost (production and transportation) exceeds the reservation price of the highest-valuation consumer (given by β). Consider a second illustrative situation where to the first situation, n_L firms (exogenously) enter at the right extremity of the line segment, point L (see the right-hand panel; assume $n_L < n_O$ and that the costs of entrants are also c). Markets located between C (where delivered prices equal the total marginal cost of the entering firms) and D (found similarly) are now shared between the incumbents located at point O and the entrants located at point L; delivered prices in the subsegment CD fall compared to the pre-entry situation. Notice also that markets located in the subsegment DA are no longer supplied by firms located at O, and delivered prices also fall[30].

[30]Note that if n_L increases to n_O, p_L will fall to p_O, the subsegment DA supplied only by these firms located at L will increase, and the subsegment CD supplied by all firms will now be shorter and have uniform prices.

Finally, note that if O and L were "sufficiently close" such that the delivered cost were everywhere lower than p_O (equal to p_L, assuming $n_O = n_L$), prices would be uniform across all markets.

Discussion in light of the framework of Chapter 2 Clearly the solution to this model better fits the stylised facts of the Brazilian cement industry that were recalled in the earlier discussion of the circular road model with price-setting firms (e.g. prices can be uniform across local markets – see footnote 30 – and cross-hauling takes places where the spatial markets of firms located at different points overlap, such as in subsegment CD). Yet what is relevant to the discussion here is that, as in the version with price-setting firms, this model with quantity-setting firms is already nested in the theoretical framework adopted in Chapter 2. To see this notice that, as mentioned above, pricing equation (3.6) corresponds to pricing equation (2.13) with Cournot behaviour (or is simply the unconstrained Cournot reaction function (2.7) of Section 2.2.3 defined at each local market). (Again, while the model is considered as a *candidate* for representing behaviour generating outcomes in the industry, it is empirically rejected – see Section 2.4.4.) Once again, I conclude that no further insight is presently provided over the theoretical framework proposed in Chapter 2. Thus, in order to reconcile the low market price elasticities of demand with the exercise of market power in the industry, one is back to requiring some dynamic story such as the restraining effect of potential entry.

3.4 Robustness checks in the demand estimation

The market price elasticities of demand coming out of the demand estimation in Chapter 2 are of the order of -0.5 for each local market. This is very low. Such low values could make one worry that they are actually underestimates of the "true" elasticities. I therefore perform several robustness checks, in addition to the specification tests of the preceding chapter. The overall conclusion arising from these checks is that the demand estimates are robust. It is worth pointing out, however, that the occurrence of any downward bias (toward zero) in the elasticity estimates would only reinforce the test-of-conduct results of Section 2.4.4, that Cournot behaviour can be rejected in favour of more collusive conduct.

Estimating demand across states simultaneously One would expect that unobserved demand shocks across states should be correlated. In this case, the residuals ϵ_{lt}^{d} of the market-level equations would be correlated across markets l. Estimating the 17 state-level demand equations simultaneously should thus yield similar estimated coefficients but somewhat lower standard errors: while one would hope that both 3SLS estimates and (equation-by-equation) 2SLS estimates are consistent, the former procedure should improve efficiency over the latter[31]. Though omitted here, I indeed find 3SLS estimates to be similar to those obtained under 2SLS, and the standard errors are smaller. To take an example, consider again the largest market, São Paulo (*SP*). Evaluating exogenous demand at the mean of the post-stabilisation phase, the price elasticity of demand is −0.366 (standard error 0.038). This compares to an elasticity of −0.333 (standard error 0.060) under 2SLS (Figure 2-12). The average 3SLS-estimated intercept in the post-stabilisation phase is 7.61 (standard error 0.09), compared to 7.54 (standard error 0.14) under 2SLS[32].

Rather than specifying a demand function for each state, an alternative approach is to specify only one demand function for the population of states and to run fixed effects instrumental-variables panel data estimation, treating each state as a unit and calculating clustered standard errors (i.e. clustering the observations pertaining to a given state). Compared to specification (2.12), one would estimate, for example,

$$\log q_{lt} = \alpha^1 + \alpha^2 Y_{lt} + \alpha^3 \log p_{lt} + \alpha_l^4 Y_{lt} \log p_{lt} + \nu_l + \epsilon_{lt}^{d} \tag{3.7}$$

where α_l^4 and ν_l are the market-specific parameters. Notice that, given the vastly larger number of observations (156 months × 17 states as opposed to 156 months for each market-level regression) relative to parameters to be estimated, this specification increases efficiency at the expense of cross-unit restrictions $\alpha_l^j = \alpha^j \; \forall l, \forall j \neq 4$. Estimation results – not shown – again point to very low market price elasticities of demand.

[31] In 3SLS a GLS approach is used to account for the correlation structure in the residuals across the equations.

[32] Across states, the (average) post-stabilisation price elasticity seems to be smaller under 3SLS: the previous mean value across states of −0.41 is now −0.36.

Functional form and reversion of dependent variable It could be that the elasticity estimates are inconsistent and/or biased (downwards) due to functional form misspecification. To investigate this possibility, I fit alternative functional forms, namely semi-log-linear and linear, as opposed to the log-linear specification (2.12), and obtain similarly low elasticity estimates. For example, I estimate the linear demand equation

$$q_{lt} = \alpha_l^1 + \alpha_l^2 Y_{lt} + \alpha_l^3 p_{lt} + \alpha_l^4 Y_{lt} p_{lt} + \epsilon_{lt}^d \tag{3.8}$$

Though again not shown, the fitted (linear) demand curves in 16 out of the 17 states rotate anticlockwise upon stabilisation: the coefficient on the interaction term is negative for these 16 states and significantly so at the 5% level in 13 of them[33]. The average price elasticity across states, computed for each state at the means for the two sub-samples, again almost doubles from -0.20 in the pre-stabilisation phase to -0.39 upon stabilisation, varying in the latter post-stabilisation phase from a minimum (in absolute) of -0.10 to a maximum of -0.67.

Thus when consumption (or its log) is taken as the dependent variable, market price elasticities of demand are estimated in each case at around -0.4 to -0.5. On the other hand, when *prices* are taken as the dependent variable, the elasticity is estimated to be higher, at around -0.7 to -0.8. However, the fit suffers significantly

[33]Recall that $\bar{Y}_{l,post} > \bar{Y}_{l,pre}$ across all states l and realise that with linear demand $\frac{dq_{lt}}{dp_{lt}} = \alpha_l^3 + \alpha_l^4 Y_{lt}$; thus the fitted demand curve rotates anticlockwise upon stabilisation iff $\hat{\alpha}_l^4 < 0$, since the estimated $\frac{dq_{lt}}{dp_{lt}} |_{\bar{Y}_{l,post}} - \frac{dq_{lt}}{dp_{lt}} |_{\bar{Y}_{l,pre}}$ is then $\hat{\alpha}_l^4(\bar{Y}_{l,post} - \bar{Y}_{l,pre}) < 0$. The mean elasticity in the pre-stabilisation period is computed at the means of the variables as $(\hat{\alpha}_l^3 + \hat{\alpha}_l^4 \bar{Y}_{l,pre}) \frac{\bar{P}_{l,pre}}{\bar{Q}_{l,pre}}$ (similarly for the post-stabilisation period). That demand becomes more elastic as the general price level in the economy stabilises from a high rate of inflation makes for interesting reading. How can one interpret such a finding? Intuitively, as the rate of change of the prices of goods and services falls, prices become more meaningful to consumers, carrying greater signal as opposed to noise, making demand more sensitive to variation in prices. Given the potential error associated with the measurement of prices in an inflationary environment, one must cautiously interpret the empirical finding that the price elasticity of demand for cement increased upon stabilisation; however, the estimate that it almost doubled on average across states seems to be a strong result. Much research has been conducted about the economic effects of general price stabilisation, but the hypothesis that stabilisation may lead to increased elasticity of demand has not been tested extensively in the empirical literature. Notice that such an anticlockwise demand rotation would, in the presence of market power on the part of producers, put downward pressure on prices in equilibrium. In the case at hand, cement prices did indeed fall concomitant with stabilisation, but I have argued earlier that, because of the binding imports constraint, this was due to a reduction in the marginal cost of imports (brought about by an appreciation of the local currency and the trade liberalising reforms of the early 1990s), rather than a change in the slope of market demand.

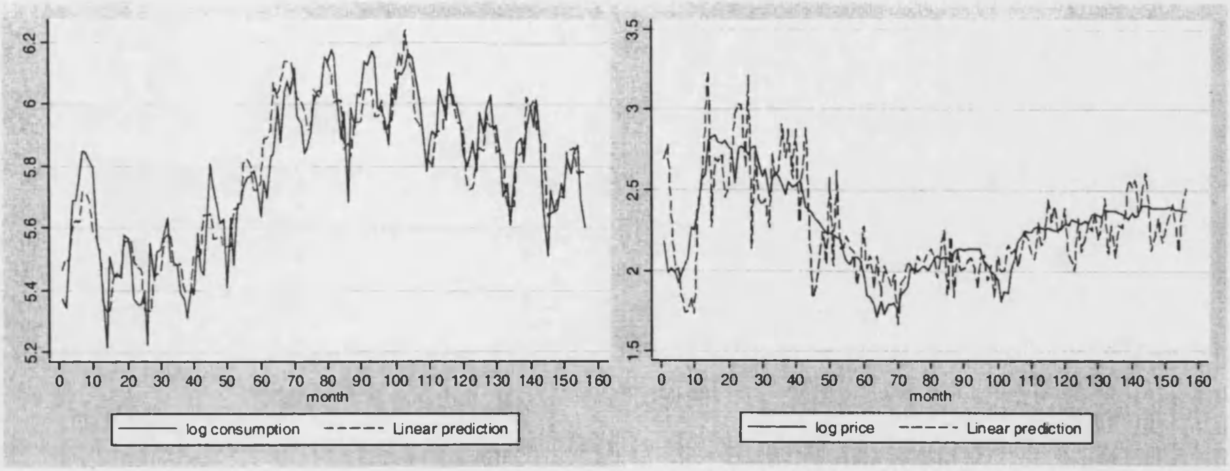


Figure 3-3: Goodness of fit in the estimation of demand for the state of Minas Gerais (*MG*). Left panel: Fit of a 2SLS regression of log consumption on log exogenous demand, log price and an interaction variable, as in (2.12). Right panel: Fit of a 2SLS regression of log price on log exogenous demand, log consumption and an interaction variable.

and the choice of prices as the dependent variable appears to be forcing the estimation of a higher elasticity. Figure 3-3 plots the fit of regression (2.12) – i.e. a

1991, leading to a large and immediate increase in prices, to identify the demand curve (I assume that demand conditions in this period were largely unchanged.) Estimated elasticities are again of the order of -0.2 to -0.3, in line with the estimates in Figure 2-14 for the pre-stabilisation phase.

Dynamics Another possibility is that while in the "short-run" the elasticity amounts to around -0.5, the "long-run" elasticity is higher. Using an autoregressive distributed lag (ADL) demand specification, where prices are regressed on consumption, lagged consumption and lagged prices, Röller and Steen (2002) estimate a long-run elasticity of -1.47, compared to the short-run elasticity of -0.46. The authors argue that "(t)his is in line with intuition, as other materials like wood and metal can be substituted for cement in the long run" (p.10). While intuitive, it is not clear why the difference between the short-run and the long-run elasticities should be this large, particularly in the Brazilian context, where cement is sold primarily to small-scale consumers who in reality do not have substitutes available and, if facing a shortfall, would rather tend to postpone their construction activities. (Note, for example, that the use of cement in highway construction, has to date been minimal, and large-scale buyers, who purchase cement in bulk as opposed to bags, account for only 20% of shipments. One can argue that dynamic effects of demand by Brazil's small-scale consumers, to the extent that they matter at all, may be present in the direction of yielding a *lower* price elasticity of demand in the long-term, the reverse of what Röller and Steen estimate for Norway.) Nor is it clear what is meant by "long-run" and which lags should be included in the ADL demand specification, particularly to the extent that observations in the present study are frequent (monthly rather than annual). Further, even if one could theoretically argue, and empirically demonstrate, that the long-run elasticity were significantly higher, it is not clear either why producers going about making their supply decisions should consider a long-term slope for the demand curve as opposed to a short-term one[35].

[35] Despite these reservations, I have attempted to estimate an ADL specification à la Röller and Steen (2002), including 6 to 12 month lags, with no success.

3.5 Using a gravity model to analyse the flow of shipments

The trade literature has extensively used the *gravity model* to explain the flow of trade and investment between countries. In a typical gravity model, such bilateral flows – indexed by the source country, the host country, and the time interval – are regressed on some measure of the host country's market size, the source country's market size and the physical distance between the countries. Feenstra (2002) describes the gravity equation as "perform(ing) extremely well empirically" (p.491). Other explanatory variables may be added to the equation, such as the flow of communication between countries (Loungani et al 2002) or some proxy of border effects (e.g. Anderson and van Wincoop 2003, Feenstra 2002).

In this section I resort to a gravity-type equation in order to statistically analyse the flow of cement between plants, owned by the different cement producers, and markets (identified with states). Of greater interest than merely explaining the observed flows, or shipments, based on plant location and ownership, lies the "detection" of outliers in the data. One example could be the absence of shipments from a given plant (and thus firm) to a given market where one would expect them to occur on account of the plant's location and ownership alone. It is clear that this exercise is not structural, in that it does not follow from a model of shipments derived from a specification of demand and cost conditions, firm entry and behaviour[36] – this structural approach was pursued in Chapter 2; here I perform a statistical benchmarking exercise, in the spirit of the empirical trade literature[37]. Rather than arbitrarily establishing the distance over which cement can profitably travel (say 500 km) and from there analysing the pattern of shipments, the merit of this approach is that of letting the data pinpoint the firm-market pairs which do not conform to an "average" pattern[38]. By the above example, one may observe

[36]Nevertheless, the use of fixed effects is an (admittedly rough) attempt to control for such local demand and cost conditions, as explained below. The exercise conditions on entry (plant location and ownership). Behaviour, as well as institutional detail, should help explain the residual.

[37]Note, however, that a gravity equation can be derived from a monopolistic competition model with CES preferences (Redding and Venables 2004).

[38]Recall from the introduction to this chapter that this "average" pattern of shipments is consistent with varying degrees of competition (i.e. alternative models of firm conduct). See also

a firm shipping to a market located at a greater distance from its plant(s) than another market of similar size to which it hardly ships. The use of plant, firm and/or market fixed effects attempts to control for price effects, in a similar vein to that of the trade literature (e.g. Redding and Venables 2004). Indeed, such outliers may be a result of historical accident, such as a product quality incident in the past which tarnished the image of a given producer in a given local market. Yet the occurrence of outliers is also consistent with tacit or explicit geographic market division on the part of producers. Two qualifications need to be made at this point. First, such outliers may be picked up to the extent that data on flows is "sufficiently" disaggregated on the spatial dimension; in the situation at hand, the data available is broken out at the state level, i.e. I observe shipments for each plant-state pair as opposed to a (preferred) plant-municipality pair. Second, the location and ownership of plants is the outcome of a complex and long-term sequence of decisions. In this exercise, I condition on the spatial distribution and ownership of plants at each time interval. To the extent that historical accident, geographic market division and other unobserved institutional aspects (e.g. availability of limestone reserves, or constraints on the number of brands resellers can stock) influence the entry of plants and firms, as one expects they should, the observation *and* lack of observation of outliers must be interpreted with caution.

Given that for many firm-market pairs zero shipments are observed[39], I estimate the following tobit model:

$$\begin{aligned} q_{flt} &= q^*_{flt} \quad \text{if} \quad q^*_{flt} = \gamma_1 + \gamma_2 d_{flt} + \gamma_3 Y_{lt} + \gamma_4 I_{flt} + \epsilon_{flt} > 0 \\ &= 0 \qquad \text{otherwise} \end{aligned} \tag{3.9}$$

where q_{flt} denotes cement shipments from firm f's plants to market l in year t, d_{flt} is some measure of the physical distance from firm f's plants to market l in year t, Y_{lt} is some measure (or a vector of measures) of the exogenous demand for cement in market l in year t, and I_{flt} is an indicator variable indicating whether firm f owns a plant located in market l in year t. The inclusion of I_{flt} attempts to

footnote 47.

[39]Recall from Figure 2-7 that the mean number of firms shipping to each state in 1999 is 5.7.

capture within-state border effects, since states are identified with markets. The base measure for Y_{lt} is construction activity, as used in the estimation of demand. I use two alternative specifications for distance. d^1_{flt} is the average distance travelled from firm f's plants to market l, weighted by that firm's observed shipments from its plants to that market each year (or, if zero shipments are observed, then it is the distance from the market of firm f's closest plant). Given the somewhat endogenous nature of the shipment-weighted average distance d^1_{flt}, in that unobserved shocks may impact both d^1_{flt} and q_{flt} via their impact on the flow of cement, I alternatively use d^2_{flt}; this is the physical distance between market l and firm f's closest plant in year t, regardless of how much that plant actually ships[40].

Figure 3-4 presents the results. Regressions depicted in columns (1) through (6) consider the logarithms of q_{flt}, d_{flt} and Y_{lt}. Column (1) includes only distance d^1_{flt} and market size Y_{lt}. (2) adds the firm-located-in-market dummy I_{flt}. All estimated coefficients have the expected sign, and are significantly different from zero. (Heteroskedasticity-robust standard errors, unless otherwise noted, are displayed in parentheses.) Of note, the elasticity of shipments with respect to distance is very high: a 1% increase in a firm's plants' average distance to a market is associated with a 2.7 to 3.7% reduction in shipments. For the sake of comparison, estimates of the elasticity of bilateral trade flows with respect to distance in the international trade literature are typically in the order of -0.8 to -1.5% (Loungani et al 2002). By regression (2), within-state border effects appear very high: controlling for distance, having a plant located within the market (state), raises shipments by 6%. It could be that state borders are playing a role in the (tacit) coordination of firms' strategies. However, the magnitude of this within-state border effect is not robust to the addition of market effects and firm effects. This is what regressions (3) to (5) show. Compared to (2), specification (3) adds market effects, (4) adds firm

[40]Formally, $d^2_{flt} := \min_{i \in \mathcal{O}_{ft}}(d_{il})$, where $\mathcal{O}_{ft}$ is the set of plants owned by firm f in year t.
In both d^1 and d^2, distances to a given market are taken to be the distance to the corresponding state's capital city – each state in Brazil has a state capital, which almost invariably is its most populous city. While this is an approximation to the true distance to market – a plant may certainly ship to municipalities other than the state capital, possibly located closer to it – this avoids the need to arbitrarily pick the distribution of a plant's shipments across municipalities within a state. Unfortunately, I do not observe the breakdown of a plant's shipments across municipalities within a state.

effects, and (5) adds market effects and firm effects. While the distance elasticity is robust to the addition of market effects and firm effects – compare (5) with (2) – the within-state border effect is now lower, though still significantly positive, as expected. However, notice that the addition of market effects considerably raises the market size elasticity. It is somewhat counterintuitive that a 1% increase in a market's construction activity would be associated with a 7% increase in a firm's shipments to that market. Rather, given that increases in market size tend to invite entry – and recall from Chapter 2 that the number of firms shipping to a market is indeed increasing in its size – one would expect the market size elasticity to be closer to 1 (and perhaps less than 1). The probable solution to this puzzling result is that the magnitude of the estimated market elasticity must be analysed jointly with the estimated market fixed effects (not shown): indeed the larger markets tend to have a large negative fixed effect (e.g. São Paulo state, with -8.5) as compared to the smaller states (e.g. Piauí state, with +9.4). The identification of the estimated market size elasticity of 7% in (5) thus appears to come from the variation of market size over time and not across states. It is possible that the observed increase in market size over the time period is correlated with changes to market structure. As discussed in Appendix C, over the time period such changes to market structure have typically consisted of a large firm acquiring a smaller firm with plants in local markets where the acquirer previously hardly shipped to. Upon such an acquisition (merger), two firm-market observations per market become one (merged-)firm-market observation, where the dependent variable shipments is now the sum of shipments from the merging firms' plants. In a period of growing market size, the market size variable may be picking up the effect of such changes in market structure, which are not controlled for. Regression (5) was reestimated dropping the market size variable Y_{lt}, allowing the (cross-sectional) variation in market size to be picked up by the market effects only (results are not shown). Importantly, the estimated distance elasticity, and the within-state border effect, do not vary. This result lends credence to the consistency of these estimates. Another robustness check was conducted by assuming that plants acquired by firms over the time period had been owned by the acquirers throughout the time period, thus in a sense controlling for market structure changes. The estimated coefficient on market size is now lower (4.13 compared to 6.84 in (5)), but the estimates for distance and the within-state border effect remain largely unchanged.

Regression (6) replaces d^1_{flt} by the alternative minimum-distance measure d^2_{flt}, suggesting that the results are fairly robust to the measure of physical distance. The same regressions were then repeated, but using variables q_{flt}, d_{flt} and Y_{lt} linearly rather than their logarithmic transformations. Figure 3-4 displays estimates for regression (7), the linear counterpart to regression (5). A 100 km increase in distance is on average associated with a 31.5 kt (315.08 t/km x 100 km) reduction in a firm's shipments to a market[41]. Notice that Figure 3-4 also displays "clustered" standard errors for regressions (5) and (7) where a same firm's shipments to a same market over time are treated as a cluster (i.e. clustering by firm and market). This allows for observations which are not independent within a cluster, corresponding to shipments in different years (though still requiring observations to be independent between clusters).

The fitted gravity equation can then be used to establish outliers in the data. Consider shipments by the eight largest firms to each of the six largest markets in 1999: São Paulo (*SP*), Minas Gerais (*MG*), Rio de Janeiro (*RJ*), Bahia (*BA*), Paraná (*PR*) and Rio Grande do Sul (*RS*)[42]. Figure 3-5 compares actual shipments to shipments predicted by the linear specification (7) of the gravity model in the year 1999.[43]

[41] Given the large variation in population density across the different "regions" of Brazil (North, Northeast, Southeast, South and Centrewest), regressions (5) and (7) – variables in logs and in linear form respectively – were modified allowing the coefficient on the distance variable to vary across regions. In the modified log-form specification (5'), estimated distance elasticities range from -3.2 for shipments to markets in the North and Northeast, to -1.9 for shipments to the Centrewest. However, equality of estimated distance elasticities across regions cannot be rejected in all but one pairwise test. I thus maintain specification (5) where the same distance elasticity applies for shipments to all markets in the country (estimated at -2.6, such that a 1% increase in average distance is associated with a 2.6% reduction in shipments). As for the modified linear form specification (7'), a 100 km increase in distance is associated with a statistically-different reduction in shipments to markets across regions: this varies from a 10-20 kt reduction in shipments to markets in the less-densely populated North, Northeast and Centrewest regions, to a 800-1100 kt reduction in shipments to markets in the relatively developed and densely populated Southeast and South.

[42] These states are adjacent to one another and are located around the Southeast region of the country – see Figure 2-6.

[43] This is done for selected firms – the firms remaining following the acquisitions over the 90s. It is also interesting to plot actual shipments against predicted shipments for these firms in each

Figure 3-4: Gravity equation estimates

	Log-linear							Linear	
	(1)	(2)	(3)	(4)	(5)	(6)	(5) with cluster. s.e.[1]	(7)	(7) with cluster. s.e.[1]
Average distance	-3.741 ***	-2.701 ***	-3.293 ***	-2.089 ***	-2.604 ***		-2.604 ***	-315.1 ***	-315.1 ***
	(0.132)	(0.159)	(0.181)	(0.131)	(0.142)		(0.348)	(36.2)	(85.2)
Market size	1.734 ***	1.547 ***	9.741 ***	2.176 ***	6.848 ***	6.607 ***	6.848 ***	738.6 ***	738.6 ***
	(0.146)	(0.143)	(1.504)	(0.122)	(1.289)	(1.310)	(1.255)	(235.8)	(238.8)
Firm-located-in-market dummy		6.295 ***	4.276 ***	3.281 ***	1.490 ***	1.098 *	1.490	353255 ***	353255 ***
		(0.535)	(0.582)	(0.544)	(0.503)	(0.614)	(1.247)	(21036)	(51592)
Alternative: Minimum distance						-2.080 ***			
						(0.121)			
Constant	14.216 ***	7.496 ***	-31.426 ***	2.296	-18.617 ***	-21.169 ***	-18.617 ***	-121876	-121876
	(1.372)	(1.452)	(7.804)	(2.343)	(6.821)	(6.902)	(7.010)	(82962)	(115270)
Market fixed effects	NO	NO	YES	NO	YES	YES	YES	YES	YES
Firm fixed effects	NO	NO	NO	YES	YES	YES	YES	YES	YES
No. observations Uncensored	1187	1187	1187	1187	1187	1187	1187	1187	1187
Total	3484	3484	3484	3484	3484	3484	3484	3484	3484
Log likelihood	-5299.0	-5256.3	-5114.9	-4909.4	-4751.3	-4793.8	-4751.3	-17321.1	-17321.1

Note: Cement flow as the dependent variable. Heteroskedasticity-robust standard errors in parentheses.
(Two-tailed tests) *** Significant(ly different from zero) at the 1% level; ** Significant at the 5% level; * Significant at the 10% level
[1] Firm-market pairs are defined as clusters; allowing for observations which are not independent within cluster (although they must be independent between clusters)

Firm \ Market	SP Actual	SP Predict	SP Diff	SP A/P	MG Actual	MG Predict	MG Diff	MG A/P	RJ Actual	RJ Predict	RJ Diff	RJ A/P	BA Actual	BA Predict	BA Diff	BA A/P	PR Actual	PR Predict	PR Diff	PR A/P	RS Actual	RS Predict	RS Diff	RS A/P
Votorantim	5,828	1,614 (537)	4,214	3.6	402	907 (229)	-505	0.4	796	917 (211)	-120	0.9	995	912 (228)	83	1.1	1,464	930 (226)	534	1.6	1,228	910 (203)	318	1.3
GJS	158	792 (346)	-634	0.2	298	266 (98)	32	1.1	328	184 (78)	144	1.8	597	253 (87)	344	2.4	30	0	30		3	0 (129)	3	
Holcim	1,068	1,292 (371)	-224	0.8	1,240	822 (124)	418	1.5	1,030	644 (91)	386	1.6	5	69 (78)	-65	0.1	23	221 (92)	-198	0.1	0	64 (103)	-64	0.0
Lafarge	555	1,063 (308)	-508	0.5	1,267	584 (114)	683	2.2	943	488 (105)	456	1.9	486	436 (91)	50	1.1	83	42 (126)	41	2.0	0	0 (219)	0	
CamCorrea	1,173	1,229 (349)	-57	1.0	998	721 (102)	277	1.4	0	101 (99)	-101	0.0	3	0	3		273	45 (109)	228	6.1	105	0 (127)	105	
CPCim	1,628	1,039 (287)	589	1.6	400	408 (95)	-8	1.0	464	312 (116)	152	1.5	0	0	0		0	0	0		0	0 (195)	0	
Brennand	44	736 (328)	-692	0.1	19	150 (91)	-131	0.1	0	0	0		126	197 (75)	-72	0.6	0	0	0		0	0 (137)	0	
Cimpor	692	1,181 (361)	-488	0.6	0	50 (103)	-50	0.0	0	0	0		219	415 (92)	-196	0.5	2	0	2		630	491 (96)	139	1.3
Soeicom	551	736 (292)	-185	0.7	437	682 (99)	-245	0.6	239	63 (101)	177	3.8	3	0	3		40	0	40		0	0 (176)	0	
Itambe	7	771 (333)	-763	0.0	3	0	3		0	0	0		0	0	0		402	520 (115)	-118	0.8	255	1 (128)	254	220.5
Ciplan	0	662 (309)	-662	0.0	24	157 (95)	-133	0.2	0	0	0		27	0	27		0	0	0		0	0 (164)	0	

Note: Actual and predicted shipments in kt. Standard errors of prediction in parentheses. Bold format for difference denotes significance at the 5% level.

Figure 3-5: Comparison of actual shipments to shipments predicted by the gravity model. In 1999, for the 6 largest markets, using the linear specification.

There are several instances where actual shipments are significantly different from predicted shipments (at the 5% level). Take shipments to the state of *MG*, for example. On the basis of plant location and market size, the gravity model predicts that Votorantim's shipments in 1999 should approximately fall within the 450-1350 kt confidence interval. Yet they amount to a mere 402 kt, corresponding to an 8% volume share, despite *MG* being the second largest market and Votorantim owning a plant located in *MG* (in addition to five plants located in the neighbouring states of *SP* and *RJ*)[44]. On the other hand, actual shipments by both Lafarge (1,267 kt, or a 25% share) and Holcim (1,240kt, or a 24% share) are significantly higher than the gravity model would predict on the basis of plant location and market size (conditional on observed plant location). This pattern is somewhat similar in the neighbouring state of *RJ*, the third largest market. Votorantim, with two plants located within the state, ships an actual 796 kt (21% share) against a predicted (point estimate of) 917 kt (though the difference is not significant at the two-tailed 5% level). Again, actual shipments by both Lafarge (943 kt, or a 25% share) and Holcim (1,030 kt, or a 27% share) significantly exceed predicted shipments (respectively 488 kt and 644 kt). However, when it comes to the state of *SP*, the largest market, this pattern is inverted. Votorantim, with three plants located in that state, ships an amount vastly higher than that predicted by the gravity model: 5,828 kt (corresponding to a 50% share) are shipped as against a predicted interval between 560 kt and 2,670 kt. Actual shipments for both Lafarge (555 kt, or a 5% share) and Holcim (1,068 kt, or a 9% share) fall short of their predicted shipments, though these differences are not significant at the 5% level. Further relating to the large *SP* market, shipments by Brennand (with a plant located in the south of the state of *GO*, almost adjacent to the state of *SP*) and Itambé (with a plant located in the state of *PR*, close to *SP*; firm 10 in Figure 2-6) fall far short of that predicted by the gravity model.

One could argue that actual shipments differ from predicted levels owing to price

of the nine years 1991 to 1999, with the 45^0 diagonal depicting where actual shipments equal predicted shipments. These plots are not included for the sake of brevity but can be obtained from the author.

[44]Note further that, in stark contrast, Votorantim's share in the neighbouring state of *SP* amounts to 50% (5,828 kt) – see below. Of these 5,828 kt, Votorantim's plant in *MG* contributes 1,182 kt, while it ships only 340 kt within *MG*.

effects, to a large extent not captured by the gravity model despite the inclusion of fixed effects. By this argument, for example, Itambé's marginal cost in serving the *SP* market (from a plant located in *PR*) could be sufficiently high, or cement prices in *SP* could be sufficiently low, that it is not profitable for Itambé to increase shipments to *SP*. But as I show in Chapter 2, where I directly calculate costs based on engineering estimates and factor prices, cement prices exceed marginal cost and the hypothesis of Cournot behaviour is rejected[45]. By way of the example, there appears to be "something else" holding back Itambé's shipments to *SP* as compared to an "average" pattern of shipments[46]. This could be, say, a quality incident in the past which tarnished its corporate image in the *SP* market. More convincingly, however, Itambé's shying away from the huge *SP* market next door to it is consistent with a market division or a multimarket contact story à la Bernheim and Whinston (1990)[47].

A similar benchmarking exercise can be applied by fitting the following alternative gravity-type model to (3.9):

$$\begin{aligned} s_{flt} &= s^*_{flt} \quad \text{if} \quad s^*_{flt} = \gamma_1 + \gamma_2 d_{flt} + \gamma_3 I_{flt} + \epsilon_{flt} > 0 \\ &= 0 \qquad \text{otherwise} \end{aligned} \tag{3.10}$$

[45]To provide a flavour of the test of conduct of the preceding chapter, if Itambé were to behave in Cournot fashion, it would raise shipments to *SP*. Assuming perfectly-competitive resellers, and considering (i) a market price elasticity of demand of -0.5, (ii) Itambé's share of shipments in *SP* of only 0.06%, and (iii) consumer prices of approximately R$ 10 / bag at the end of 1999, Itambé would increase shipments as long as its marginal cost (including the reseller's margin) fell short of its perceived marginal revenue: $10 + \frac{10}{-0.5}0.06\% \simeq 10$. Similar reasoning applies to Lafarge and Holcim's shipments to *SP*, respectively with shares of 7% and 9%.

[46]See the following footnote and the discussion of the identification problem of conduct and cost in the introduction to this chapter.

[47]Clearly the gravity model is an imperfect statistical device for detecting market division. To illustrate, consider a hypothetical situation where producers were to divide markets in the extreme and, as a result, cement were hardly to travel. All that the gravity-equation estimates would suggest is that the elasticity of shipments with respect to distance would be very high, and this would be consistent with competitive behaviour. Nevertheless, to the extent that market division is "incomplete" and consumers and producers are unevenly distributed over space, estimates from the gravity equation may point at some interesting outliers, where cement travels less (or more) than the "average".

Ideally a researcher would wish to observe a natural experiment in the degree of antitrust enforcement – say, different regional markets subject to different degrees of enforcement, or a given region seeing an increase in the degree of enforcement over time – and hope to learn something from (cross-sectional or times-series) changes to the distance elasticity. This is a possible avenue for future research.

where $s^*_{flt} := \frac{q^*_{flt}}{\sum_f q^*_{flt}}$ is firm f's share of shipments to market l in year t. The market size variable Y_{lt} is dropped in this shipment-share version (3.10) as compared to specification (3.9); market effects and firm effects are again included. Estimated coefficients (not shown) on both the distance variable and the within-state border effect have the expected sign and are significantly different from zero at the 1% level. A 100 km increase in the distance to a market is associated with a 2.6% reduction in the shipment share (clustered standard error of 0.25%). Controlling for distance, having a plant located within a market is associated with a 19.5% increase in shipment share (clustered standard error of 2.9%): this again suggests that state borders may be playing an important role in the (tacit) coordination of firms' strategies.

Figure 3-6 compares actual shipments in 1996 against predicted shipments to markets located in the northeast region, ranging from the state of Maranhão (*MA*) to the state of Sergipe (*SE*). Three producers have been dominant in this region throughout the 1990s: Votorantim, GJS and Brennand. To illustrate, consider the relatively large market of Ceará (*CE*). Both Votorantim and GJS own a plant located within *CE*, with the latter firm owning a further two plants nearby (see Figure 2-6). Yet Votorantim has over three times GJS's share of shipments in *CE*. While the share-version of the gravity equation predicts a share of 53%, Votorantim commands an actual 71% share in 1996, with GJS accounting for an actual share of only 22% against a predicted 46%[48]. Note further that Votorantim also ships to *CE* from its plant in the state of Paraíba, *PB*, 700 km away. In contrast, Brennand, operating a plant in *PB* next door to Votorantim's plant in *PB*, hardly ships to *CE*. While Brennand's actual share of shipments to *CE* amounts to only 3%, the gravity equation predicts a much higher share, of 13%. In *PB*, however, Brennand appears to enjoy a free hand: countering prediction, Brennand has an actual share of shipments in *PB* that is double that of Votorantim, also located within the state. A story where Votorantim tacitly agrees to give Brennand the upper hand

[48] Interestingly, in 1991 GJS hauled significant quantities of cement through *CE* from its plants located in the states of *RN* and *PE* in the east, bound for the state of *PA* in the west, approximately 2000 km away. Consumer prices in *PA* at the time were only slightly higher than those in *CE*, and GJS's 20% share of shipments in *CE* already paled in comparison to Votorantim's 76%. (In *PA*, by contrast, GJS's shipment share of 77% dwarfed Votorantim's 23% share.)

in *PB* in exchange for the latter staying away from *CE* would help explain the observed flows. A similar situation occurs between Votorantim and Brennand in the small and adjacent states of Alagoas, *AL* (Brennand commands an 83% share and operates a plant within the state) and Sergipe, *SE* (Votorantim commands an 89% share and operates a plant within the state, though also shipping from as far as its plant in *PB*, 600 km away). Of note, prices in both *AL* and *SE* are similar to that in other markets to which cement is shipped over greater distances[49].

3.6 Concluding remarks

This chapter has explored the robustness and validity of different theoretical and estimation-related modelling aspects and findings which concern the preceding chapter. By providing examples of simple dynamic multimarket games where market division can be supported in equilibrium, it indicates the rationality of the type of conduct in the Brazilian cement industry that was identified previously. On devising strategies which take multimarket contact into account, firms can achieve complete market division, limiting cross-hauling and maximising aggregate profit. The chapter then reviews some seminal models from the spatial competition literature to validate the theoretical framework of Chapter 2. A variety of additional robustness and specification tests of the demand estimation of that chapter are then performed. Finally, a gravity model is estimated to analyse cement shipments. While gravity equations have been extensively used in the trade literature to analyse the effect of distance, broadly defined, on cross-industry trade and investment flows, this chapter's gravity exercise using a rich *industry-specific* dataset – complemented by the structural estimation of the preceding chapter – points to a potential identification problem in the trade literature's use of gravity models: that estimated distance effects may also be proxying for firms' strategies of dividing

[49]This example was provided as a "case in point" in Section 2.4.4. If Brennand behaved in Cournot fashion, its perceived marginal revenue in the state of *SE*, next door to its plant in *AL*, would be $p + \frac{p}{-0.5}.0\% = p$, where the price elasticity is -0.5 and its share of shipments is 0% (it does not ship to *SE*). As seen in Chapter 2, its marginal cost is lower than p, if only because its plant in *AL* ships to the states of *PB*, *PE* and *BA*, located at further distances than *SE* and where prices are similar to those in *SE*. Thus the hypothesis of Cournot behaviour can be rejected in favour of more collusive conduct.

Market / Firm	*MA* Actual	*MA* Predict	*MA* Diff	*MA* A/P	*PI* Actual	*PI* Predict	*PI* Diff	*PI* A/P	*CE* Actual	*CE* Predict	*CE* Diff	*CE* A/P	*RN* Actual	*RN* Predict	*RN* Diff	*RN* A/P
Votorantim	34	31 (4)	4	1.1	55	39 (5)	16	1.4	71	53 (8)	18	1.3	15	35 (8)	-21	0.4
GJS	55	61 (4)	-7	0.9	30	30 (5)	1	1.0	22	46 (9)	-24	0.5	49	46 (7)	3	1.1
Brennand	6	0	6		0	6 (5)	-6	0.0	3	13 (7)	-10	0.3	36	24 (7)	13	1.5
Matsulfur	0	0	0		0	0	0		0	0	0		0	0	0	
Cisafra	3	8 (4)	-4	0.4	15	7 (5)	8	2.2	4	0	4		0	0	0	
Lafarge	0	0	0		0	0	0		0	0	0		0	0	0	
Ciplan	1	0	1		0	0	0		0	0	0		0	0	0	

Market / Firm	*PB* Actual	*PB* Predict	*PB* Diff	*PB* A/P	*PE* Actual	*PE* Predict	*PE* Diff	*PE* A/P	*AL* Actual	*AL* Predict	*AL* Diff	*AL* A/P	*SE* Actual	*SE* Predict	*SE* Diff	*SE* A/P
Votorantim	29	50 (8)	-21	0.6	57	53 (7)	4	1.1	7	27 (8)	-20	0.2	89	54 (7)	35	1.7
GJS	18	17 (7)	2	1.1	20	46 (8)	-26	0.4	11	21 (8)	-10	0.5	10	45 (7)	-35	0.2
Brennand	53	38 (8)	15	1.4	17	19 (6)	-2	0.9	83	43 (10)	40	1.9	0	17 (6)	-17	0.0
Matsulfur	0	0	0		0	0	0		0	0	0		0	1 (6)	-1	0.0
Cisafra	0	0	0		6	0	6		0	0	0		1	0	1	
Lafarge	0	0	0		0	0	0		0	0	0		0	0	0	
Ciplan	0	0	0		0	0	0		0	0	0		0	0	0	

Note: Actual and predicted market shares. Standard errors of prediction in parentheses. Bold format for difference denotes significance at the 5% level.

Figure 3-6: Comparison of actual shipments to shipments predicted by the gravity model. In 1996, for markets in the northeast (*MA* to *SE*), using the volume-share version of the linear specification.

(geographic) markets.

A. Appendix to Chapter 1

A.1 Proof of Corollary 1.1 and other sufficient conditions that rule out Nash-strategy sets (d), (e) and (f) as candidate equilibria, in the case of symmetry

I begin by proving Corollary 1.1. Consider the first part of the sufficient condition, namely that $\{\theta \mid \Phi_I(\theta) \geq 0\} \subseteq \{\theta \mid \Phi_{II}(\theta) \geq 0\}$. It rules out sets (e) and (f) as candidate equilibria. To see this, notice that if either set (e) or set (f) forms an equilibrium, necessary conditions from Proposition 1.2 are that $\Phi_I \geq 0$ and $\Phi_{II} < 0$. Hence if whenever $\Phi_I \geq 0$ it happens that $\Phi_{II} \geq 0$, i.e. if $\Phi_{II} \geq 0 \mid \Phi_I \geq 0$, then neither set (e) nor set (f) can form an equilibrium. (Recall that, from the symmetry property, $\Phi_I = \Phi_V$ and $\Phi_{II} = \Phi_{VI}$. Φ_I is thus the surplus from any merger in isolation. Φ_{II} is thus the surplus from any merger conditional on the merger of the rival set of firms.) Now consider the second part of the sufficient condition, namely that $\{\theta \mid \Phi_{II}(\theta) \geq 0\} \subseteq \{\theta \mid \Phi_{III}(\theta) \geq 0\}$. It rules out set (d) as a candidate equilibrium. To see this, notice that if set (d) forms an equilibrium, necessary conditions are that $\Phi_{II} \geq 0$ and $\Phi_{III} < 0$. Hence if whenever $\Phi_{II} \geq 0$ it happens that $\Phi_{III} \geq 0$, i.e. if $\Phi_{III} \geq 0 \mid \Phi_{II} \geq 0$, then set (d) cannot form an equilibrium. (Recall that, from the symmetry property, $\Pi_{M_1,s_3} = \Pi_{M_2,s_3}$ and $\sum_{i\in\mathcal{M}_1} \Pi_{i,s_0} = \sum_{i\in\mathcal{M}_2} \Pi_{i,s_0}$, so that $\Phi_{III} = \Pi_{M_1,s_3} - \sum_{i\in\mathcal{M}_1} \Pi_{i,s_0} = \Pi_{M_2,s_3} - \sum_{i\in\mathcal{M}_2} \Pi_{i,s_0}$. Φ_{III} is thus $\mathcal{M}_1$-firms' or $\mathcal{M}_2$-firms' difference in profits when two mergers are undertaken over no merger being undertaken.) *Q.E.D.*

Inspection of Proposition 1.2 indicates other sufficient conditions that rule out Nash-strategy sets (d), (e) and/or (f) as candidate equilibria. I provide three.

1. Set (d) is ruled out as an equilibrium if the sum of payoffs of independent firms in $\mathcal{M}_1$ ($\mathcal{M}_2$) is always lower under the fragmented market structure s_0 (s_0) than under the more concentrated market structure s_1 (s_2), where the rival set of firms has merged.

   Proof. If set (d) forms an equilibrium, necessary conditions from Proposition 1.2 are that $\Phi_{II} \geq 0$ and $\Phi_{III} < 0$, i.e. $\Pi_{M_2,s_3} - \sum_{i\in\mathcal{M}_2} \Pi_{i,s_2} \geq 0$ and $\Pi_{M_1,s_3} - \sum_{i\in\mathcal{M}_1} \Pi_{i,s_0} < 0$. From the definition of symmetry, these two conditions can be combined into $\sum_{i\in\mathcal{M}_1} \Pi_{i,s_0} = \sum_{i\in\mathcal{M}_2} \Pi_{i,s_0} > \Pi_{M_1,s_3} = \Pi_{M_2,s_3} \geq \sum_{i\in\mathcal{M}_2} \Pi_{i,s_2} = \sum_{i\in\mathcal{M}_1} \Pi_{i,s_1}$. Hence if $\sum_{i\in\mathcal{M}_1} \Pi_{i,s_0} = \sum_{i\in\mathcal{M}_2} \Pi_{i,s_0} \leq \sum_{i\in\mathcal{M}_1} \Pi_{i,s_1} = \sum_{i\in\mathcal{M}_2} \Pi_{i,s_2}$ (note that this occurs iff $\Phi_{II} \leq \Phi_{III}$) then set (d) cannot form an equilibrium. Note that this condition and the second part of the sufficient condition of Corollary 1.1 are based on the same necessary conditions from Proposition 1.2; they are not, however, equivalent. ■

2. Set (e) is ruled out as an equilibrium if the payoff to merged firm M_1 (M_2) is always lower under market structure s_2 (s_1), where the rival set of firms has not merged, than under the more concentrated market structure s_3 (s_3), where the two mergers have been undertaken.

   Proof. If set (e) forms an equilibrium, necessary conditions are that $\Phi_{II} < 0$ and $\Phi_{IV} \geq 0$, i.e. $\Pi_{M_2,s_3} - \sum_{i\in\mathcal{M}_2} \Pi_{i,s_2} < 0$ and $\Pi_{M_1,s_2} - \sum_{i\in\mathcal{M}_1} \Pi_{i,s_1} \geq 0$. Again, from the definition of symmetry, these are jointly equivalent to $\Pi_{M_1,s_2} = \Pi_{M_2,s_1} \geq \sum_{i\in\mathcal{M}_1} \Pi_{i,s_1} = \sum_{i\in\mathcal{M}_2} \Pi_{i,s_2} > \Pi_{M_2,s_3} = \Pi_{M_1,s_3}$. Hence if $\Pi_{M_1,s_2} = \Pi_{M_2,s_1} \leq \Pi_{M_1,s_3} = \Pi_{M_2,s_3}$ (note that this occurs iff $\Phi_{IV} \leq \Phi_{II}$) then set (e) cannot form an equilibrium. ■

3. Both set (e) and set (f) are ruled out as equilibria if the surplus from merging in isolation, when positive, is always lower than when the rival set of firms also merges.

   Proof. If either set (e) or set (f) forms an equilibrium, necessary conditions are that $\Phi_I \geq 0$ and $\Phi_{II} < 0$. Hence if whenever $\Phi_I \geq 0$ it happens that

$\Phi_I \leq \Phi_{II}$, i.e. if $\Phi_I \leq \Phi_{II} \mid \Phi_I \geq 0$, then neither set (e) nor set (f) can form an equilibrium. (Recall that, from the symmetry property, $\Phi_I = \Phi_V$ and $\Phi_{II} = \Phi_{VI}$. Φ_I is thus the surplus from any merger in isolation, i.e. when the rival set of firms does not merge. Φ_{II} is thus the surplus from any merger conditional on the merger of the rival set of firms.) Notice that when this condition is satisfied, then the first part of the sufficient condition of Corollary 1.1 is necessarily satisfied (but the reverse is not true). ■

In addition to satisfying the sufficient condition of Corollary 1.1, which alone already rules out sets (d), (e) and (f), it can be shown that the models of Examples 1 and 2 also satisfy sufficient conditions 1-3 just stated. That is, for all admissible parameter values, it holds that (1) $\Phi_{II} \leq \Phi_{III}$, (2) $\Phi_{IV} \leq \Phi_{II}$, and (3) whenever $\Phi_I \geq 0$ it happens that $\Phi_I \leq \Phi_{II}$. For instance, consider the Perry and Porter (1985) model of Example 2 and sufficient conditions 1 and 2 just stated. For sets (d) and (e) respectively to be supported in equilibrium, for some parameter values firms would need to be better off under less concentrated market structures as against more concentrated ones: firm M_1 would need to be better off under $\{M_1, 2, 4\}$ than under $\{M_1, M_2\}$, or firm 1 would need to be better off under $\{1, 2, 3, 4\}$ than under $\{M_2, 1, 3\}$. This is not the case for any combination of parameters and hence sets (d) and (e) can be ruled out as candidate equilibria.

Another example, not mentioned elsewhere in this chapter, is afforded by Deneckere and Davidson (1985)'s model of symmetrically differentiated goods with *Bertrand competition.* It is straightforward to solve the simple sequential merger game with two merger stages for an industry initially with four homogeneous firms and, say, $\mathcal{M}_1 = \{1, 3\}$ and $\mathcal{M}_2 = \{2, 4\}$. For the entire range of the substitutability parameter, it can be shown that $\Phi_{IV} < 0 < \Phi_I < \Phi_{II} < \Phi_{III}$. Therefore, not only is the sufficient condition of Corollary 1.1 satisfied, but the other sufficient conditions 1-3 just stated also hold. As such, sets (d), (e) and (f) cannot be supported in equilibrium. Indeed, the Bertrand assumption ensures that only set (a) obtains in equilibrium, with mergers taking place from each node along the equilibrium path of the game and also from each node lying outside the equilibrium path. Again, this result hinges on the response of non-participating firms to a merger. As Deneckere and Davidson (1985) have shown, reaction schedules under competition in prices are typically upward-sloping, with non-participants to a merger reacting to the participants' price increase by themselves raising prices. (See also Levy and Reitzes 1992.)

A.2 Cross-border mergers in a vertically-differentiated industry

A.2.1 Historical motivation: Derivation of the autarky equilibrium

In order to motivate the number of firms in each country (three) and their respective qualities (high quality u_A for firms located in country A and low quality u_B for firms located in country B, such that $v = \frac{u_A}{u_B} \geq 1$), the equilibrium to the autarkic (long-term) entry and investment game is derived[1]. Recall the two-stage game played in each country: in a first stage, firms simultaneously make entry and quality investment decisions while, in a second stage, they engage in Cournot competition.

In the second (market) stage, given that n firms have entered in the first (entry and investment) stage with qualities $u = (u_j)$, $j = 1, ..., n$, the gross profit of firm i (recalling the price-to-quality ratio λ) is

$$\Pi_i = p_i x_i - c x_i = \lambda u_i x_i - c x_i$$

Firm i maximises Π_i taking the vector of qualities u from the earlier stage and x_j, $j \neq i$, as given. The first order condition is $\lambda u_i + u_i x_i \frac{d\lambda}{dx_i} - c = 0$, where from (1.2) we have $\frac{d\lambda}{dx_i} = -\frac{S}{(\sum_{j=1}^{n} u_j x_j)^2} u_i = -\frac{u_i}{S}\lambda^2$, or

$$u_i x_i = \frac{c - \lambda u_i}{\frac{d\lambda}{dx_i}} = \frac{S}{\lambda} - \frac{cS}{\lambda^2}\frac{1}{u_i} \tag{4.1}$$

[1] The specification here closely resembles Motta (1992) and the derivation follows Sutton (1998, Appendix 15.1).

Summing over j, we obtain $\sum_{j=1}^{n} u_j x_j = n\frac{S}{\lambda} - \frac{cS}{\lambda^2}\sum_{j=1}^{n}(\frac{1}{u_j})$. Using expression (1.2), I can solve for the price-to-quality ratio, $\lambda = \frac{c}{n-1}\sum_{j=1}^{n}(\frac{1}{u_j})$. Substituting for λ in FOC (4.1), I solve for the output of firm (variety) i:

$$x_i = \frac{S}{\lambda u_i}(1 - \frac{c}{\lambda}\frac{1}{u_i}) = \frac{S}{c}\frac{n-1}{u_i \sum_{j=1}^{n}(\frac{1}{u_j})}(1 - \frac{n-1}{u_i \sum_{j=1}^{n}(\frac{1}{u_j})}) \tag{4.2}$$

Note that by labelling the firm offering the lowest quality as firm 1, a necessary and sufficient condition for all n firms to command positive sales in equilibrium is $u_1 \sum_{j=1}^{n}(\frac{1}{u_j}) > n - 1$[2]. I can further solve for the price and gross profit:

$$p_i = \lambda u_i = c\frac{u_i}{n-1}\sum_{j=1}^{n}(\frac{1}{u_j}) \tag{4.3}$$

$$\Pi_i = (p_i - c)x_i = (1 - \frac{n-1}{u_i \sum_{j=1}^{n}(\frac{1}{u_j})})^2 S \tag{4.4}$$

Note that price does not depend on market size and that profit is increasing in quality and does not depend on marginal cost at equilibrium.

I now turn to the entry and investment stage. Sutton (1991, c.3) proves that at the unique Nash equilibrium, firms choose the same quality level u. In this case, $u_i \sum_{j=1}^{n}(\frac{1}{u_j}) = n$, and for every firm choosing to enter, output, price and gross profit collapse to:

$$x_i = x = \frac{S}{c}\frac{n-1}{n^2} \tag{4.5}$$

$$p_i = p = c\frac{n}{n-1} \tag{4.6}$$

$$\Pi_i = \Pi = \frac{S}{n^2} \tag{4.7}$$

Recall from Section 1.3 that $F(u) = u^3$, $u \geq 1$. (The convexity of the fixed cost function is chosen so that, in equilibrium, three firms find it profitable to enter each country.) Net profit per firm is given by $\pi(n, u) = \Pi(n) - F(u) = \frac{S}{n^2} - u^3$. The industry equilibrium, where n firms enter with quality u, is characterised by two conditions (Motta 1992):

[2] That is, the lowest quality offering with positive sales has a quality that exceeds the harmonic mean of the qualities of all offerings multiplied by $\frac{n}{n-1}$ (the latter approximates 1 for large n).

(I) (Free entry) $\pi(n,u) \geq 0$ (viability) and $\pi(n+1,u) < 0$ (stability)

(II) (Optimal quality[3]) $\frac{d\Pi}{du} = \frac{dF}{du}$

By considering a deviant firm i offering quality u_i when all its rivals $j \neq i$ offer a common quality u, condition (II) becomes[4]:

$$\frac{d\pi_i}{du_i}|_{u_i=u} = \frac{d\Pi_i}{du_i}|_{u_i=u} - \frac{dF_i}{du_i}|_{u_i=u} = \frac{2S(n-1)^2}{un^3} - 3u^2 = 0$$

which can be rearranged to

$$u^3 = \frac{2S(n-1)^2}{3n^3} \tag{4.8}$$

Given $u = u(n)$ by condition (II), and that condition (I) can be expressed as $\sqrt{\frac{S}{[u(n+1)]^3}} - 1 < n \leq \sqrt{\frac{S}{[u(n)]^3}}$, condition (I) may then be rewritten as

$$\sqrt{\frac{3(n+1)^3}{2n^2}} - 1 < n \leq \sqrt{\frac{3n^3}{2(n-1)^2}}$$

The only possible solution to this inequality is $n = 3$ and this does not depend on S[5].

The autarky equilibrium is then given by equations (4.5) to (4.8), where $n = n_l = 3$ firms enter in each country $l \in \{A, B\}$ and the common quality, price, firm output and firm (gross and net) profits are given by:

$$\begin{aligned} u_l &= \frac{2}{3}\sqrt[3]{\frac{S_l}{3}}, \quad p_l = \frac{3}{2}c, \quad x_l = \frac{2}{9}\frac{S_l}{c}, \\ \Pi_l &= \frac{S_l}{9}, \quad \pi_l = \frac{S_l}{9} - (u_l)^3 = \frac{S_l}{81} \end{aligned} \tag{4.9}$$

Given the assumption $u_l \geq 1$, I further assume that $S_l > \frac{81}{8}$. Quality is an in-

[3] As in Motta (1992), I consider only internal solutions.

[4] From equation (4.4), when all firm i's rivals offer a common quality u, I obtain $\Pi_i = (1 - \frac{1}{\frac{1}{n-1}+\frac{u_i}{u}})^2 S$. Then $\frac{d\Pi_i}{du_i} = \frac{2S}{u}\frac{\frac{1}{n-1}+\frac{u_i}{u}-1}{(\frac{1}{n-1}+\frac{u_i}{u})^3}$, and by evaluating this expression at $u_i = u$ the marginal benefit for the deviant firm i of increasing quality when all firms choose a common quality u is $\frac{d\Pi_i}{du_i}|_{u_i=u} = \frac{2S(n-1)^2}{un^3}$. The SOC is satisfied at the solution $n = 3$ (see below).

[5] This "non-convergence" result is consistent with the finiteness property of many vertical product differentiation models. The symmetry of the quality chosen by firms at equilibrium is, however, less common in the literature and hinges on the symmetry of consumer preferences and the assumption of quantity as opposed to price competition (Shaked and Sutton 1983, Sutton 1991, Motta 1992).

creasing function of market size so in view of the assumption that $S_A \geq S_B (> \frac{81}{8})$, quality offered by firms in country A is at least equal to that offered by their counterparts in country B: $u_A \geq u_B (\geq 1)$.

A.2.2 Reduced-form profit functions in game with T merger stages and (initially) T independent firms in each country

Consider the sequential cross-border merger game spelled out at the beginning of Section 1.3 yet embed each country initially with T independent firms, $T \geq 2$, and extend the game to T merger stages. (The setup in Section 1.3 refers to the case $T = 3$. See below.) Thus, the firms located in country A (producing with quality u_A) are labelled $1, 2, 3, ..., T$ and the firms located in country B (quality u_B) are labelled $T + 1, T + 2, T + 3, ..., 2T$. In the same vein, in the first stage firms 1 and $T + 1$ decide whether to merge (if formed, merged firm is labelled M_1), in the second stage firms 2 and $T + 2$ decide whether to merge (if formed, merged firm is labelled M_2), and so on, in sequence until stage T, where firms T and $2T$ are the last pair of firms to undertake a merger decision (if formed, merged firm is labelled M_T). The game ends at stage $T + 1$, the market competition stage. Given the symmetry, the $T + 1$ possible market structures coming out of the T merger stages are labelled $r_0, r_1, r_2, ..., r_T$ where, as in Section 1.3, the subscript denotes the number of cross-border mergers undertaken.

To solve for market competition equilibrium outcomes as a function of market structure, begin by considering market structure r_i, where $i = 1, 2, 3, ..., T - 1$, i.e. both independent and merged (multinational) firms exist (I will return to the "corner" structures r_0 and r_T shortly). Specifically, under market structure r_i there are i multinational firms, $T - i$ independent A-country firms and $T - i$ independent B-country firms.

By the merger-technology assumption, a multinational firm produces at quality level $\max(u_A, u_B) = u_A$ not only in country A but also in country B. Clearly, given the unit trade cost $t \geq 0$, it will no longer trade between countries, meeting the demand for its (high-quality) product in each country through domestic production[6].

From the demand setup, in equilibrium consumer prices in country $l \in \{A, B\}$ are such that $\frac{p_m^l}{u_A} = \frac{p_a^l}{u_A} = \frac{p_b^l}{u_B}$, where p_j^l denotes the price of the good produced by firm j and sold in country l. (Recall that the subscripts m, a and b denote a multinational firm, an independent A-country firm and an independent B-country firm, respectively.) This may be written in terms of the quality ratio $v = \frac{u_A}{u_B} \geq 1$:

$$p_m^l = p_a^l = vp_b^l \qquad l \in \{A, B\} \tag{4.10}$$

Thus, to illustrate, consumer prices in country B are given by p_m^B, p_a^B and p_b^B, where high-quality goods (produced by multinational firms and independent A-country firms) command a price premium relative to low-quality goods (produced by independent B-country firms).

With no loss of generality (solely for the purpose of labelling firms), assume that the i multinational firms in market structure r_i were formed in the first i merger stages. Similar to equation (1.3), the inverse demand functions for a low-quality product in country A and country B are, respectively:

$$p_b^A = \frac{S_A}{v\sum_{j=M_1}^{M_i} x_j^A + v\sum_{j=i+1}^{T} x_j^A + \sum_{j=T+i+1}^{2T} x_j^A}$$

$$p_b^B = \frac{S_B}{v\sum_{j=M_1}^{M_i} x_j^B + v\sum_{j=i+1}^{T} x_j^B + \sum_{j=T+i+1}^{2T} x_j^B}$$

The first term in the denominator corresponding to each market is the quality-adjusted sales of the i multinational firms in that market; the second and third terms respectively refer to the quality-adjusted sales of the $T - i$ independent A-country firms and the $T - i$ independent B-country firms in that market. The

[6]This would not necessarily be the case were a merger *not* to lead to a quality upgrade in the production facilities in country B. For a large enough quality ratio relative to the trade cost, it can be shown that a multinational firm would continue exporting high-quality product produced by its A-country plant to country B, discontinuing production operations in country B. For example, when $T = 3$ under market structure r_1 (one multinational firm and two independent firms in each country), this would happen if $v > 1 + \frac{t}{c}$. Otherwise, for $v \leq 1 + \frac{t}{c}$, the multinational firm would produce high-quality product in country A for consumption in country A and produce low-quality product in country B for consumption in country B.

inverse demand functions for high-quality product in both countries, $p^l_m = p^l_a$, $l \in \{A, B\}$, can then be obtained from equation (4.10).

Each firm's optimisation problem can now be written. A merged multinational firm m, $m = M_1, M_2, ..., M_i$, maximises profits by setting x^A_m and x^B_m to solve

$$\max_{x^A_m \geq 0, x^B_m \geq 0} \; p^A_m x^A_m + p^B_m x^B_m - c(x^A_m + x^B_m)$$

taking all other firms' outputs as given. As in equation (1.2), writing the price-to-quality ratio in country A as $\lambda^A = \frac{S_A}{v\sum_{j=M_1}^{M_i} x^A_j + v\sum_{j=i+1}^{T} x^A_j + \sum_{j=T+i+1}^{2T} x^A_j}$, and the price-to-quality ratio in country B as $\lambda^B = \frac{S_B}{v\sum_{j=M_1}^{M_i} x^B_j + v\sum_{j=i+1}^{T} x^B_j + \sum_{j=T+i+1}^{2T} x^B_j}$, the optimisation problem for the multinational firm may be rewritten

$$\max_{x^A_m \geq 0, x^B_m \geq 0} \; v(\lambda^A x^A_m + \lambda^B x^B_m) - c(x^A_m + x^B_m)$$

Since $\frac{d\lambda^l}{dx^l_m} = -v\frac{(\lambda^l)^2}{S_l}$, $l \in \{A, B\}$, the two FOCs are

$$vx^A_m = \frac{S_A}{\lambda^A}\left(1 - \frac{1}{\lambda^A}\frac{c}{v}\right) \tag{4.11}$$

$$vx^B_m = \frac{S_B}{\lambda^B}\left(1 - \frac{1}{\lambda^B}\frac{c}{v}\right) \tag{4.12}$$

An independent A-country firm a, $a = i+1, i+2, ..., T$, located in country A, sets x^A_a and x^B_a to solve

$$\max_{x^A_a \geq 0, x^B_a \geq 0} \; p^A_a x^A_a + (p^B_a - t)x^B_a - c(x^A_a + x^B_a)$$

where sales in country B are subject to the unit trade cost. This may be rewritten in terms of the price-to-quality ratios in each market:

$$\max_{x^A_a \geq 0, x^B_a \geq 0} \; v(\lambda^A x^A_a + \lambda^B x^B_a) - cx^A_a - (c+t)x^B_a$$

Since $\frac{d\lambda^l}{dx_a^l} = -v\frac{(\lambda^l)^2}{S_l}$, $l \in \{A, B\}$, the two FOCs are

$$vx_a^A = \frac{S_A}{\lambda^A}\left(1 - \frac{1}{\lambda^A}\frac{c}{v}\right) \tag{4.13}$$

$$vx_a^B = \frac{S_B}{\lambda^B}\left(1 - \frac{1}{\lambda^B}\frac{c+t}{v}\right) \tag{4.14}$$

Similarly, an independent B-country firm b, $b = T+i+1, T+i+2, ..., 2T$, located in country B, solves

$$\max_{x_b^A \geq 0, x_b^B \geq 0} (p_b^A - t)x_b^A + p_b^B x_b^B - c(x_b^A + x_b^B)$$

or,

$$\max_{x_b^A \geq 0, x_b^B \geq 0} \lambda^A x_b^A + \lambda^B x_b^B - (c+t)x_b^A - cx_b^B$$

Now $\frac{d\lambda^l}{dx_b^l} = -\frac{(\lambda^l)^2}{S_l}$, $l \in \{A, B\}$, and the two FOCs are

$$x_b^A = \frac{S_A}{\lambda^A}\left(1 - \frac{1}{\lambda^A}(c+t)\right) \tag{4.15}$$

$$x_b^B = \frac{S_B}{\lambda^B}\left(1 - \frac{1}{\lambda^B}c\right) \tag{4.16}$$

Adding across all FOCs pertaining to market A (i FOCs (4.11) for the multinational firms, $T-i$ FOCs (4.13) for the independent A-country firms and $T-i$ FOCs (4.15) for the independent B-country firms), I obtain:

$$v\sum_{j=M_1}^{M_i} x_j^A + v\sum_{j=i+1}^{T} x_j^A + \sum_{j=T+i+1}^{2T} x_j^A = \frac{S_A}{\lambda^A}\left(2T - i - \frac{1}{\lambda^A}\left(c\frac{T+(T-i)v}{v} + (T-i)t\right)\right)$$

and noting that the LHS is simply $\frac{S_A}{\lambda^A}$ I can solve for λ^A:

$$\lambda^A = \frac{1}{2T-(i+1)}\left(\frac{c(T+(T-i)v) + (T-i)tv}{v}\right)$$

Similarly adding across all FOCs pertaining to market B (i FOCs (4.12), $T-i$

FOCs (4.14) and $T-i$ FOCs (4.16), λ^B can be obtained:

$$\lambda^B = \frac{1}{2T-(i+1)}\left(\frac{c(T+(T-i)v)+(T-i)t}{v}\right)$$

Substituting for the price-to-quality ratios in FOCs (4.11) through (4.16), sales per firm in each market are obtained:

$$\begin{aligned}
x_m^A &= (2T-(i+1))\,S_A\frac{c((T-i)v-(T-i-1))+(T-i)tv}{(c\,(T+(T-i)\,v)+(T-i)tv)^2} \\
x_a^A &= x_m^A \\
x_b^A &= (2T-(i+1))\,S_A v\frac{c(T-(T-1)v)-(T-1)tv}{(c\,(T+(T-i)\,v)+(T-i)tv)^2} \\
x_m^B &= (2T-(i+1))\,S_B\frac{c((T-i)v-(T-i-1))+(T-i)t}{(c\,(T+(T-i)\,v)+(T-i)t)^2} \\
x_a^B &= (2T-(i+1))\,S_B\frac{c((T-i)v-(T-i-1))-(T-1)t}{(c\,(T+(T-i)\,v)+(T-i)t)^2} \\
x_b^B &= (2T-(i+1))\,S_B v\frac{c(T-(T-1)v)+(T-i)t}{(c\,(T+(T-i)\,v)+(T-i)t)^2}
\end{aligned}$$

Non-negativity constraints – ensuring that despite trade costs $t \geq 0$ and quality asymmetries $v \geq 1$, all firms command positive sales in both countries – can be summarised as two parameter restrictions. The "low-enough-quality-ratio" restriction follows from requiring that low-quality B-country firms are still able to sell in their home country when imports are most competitive ($t = 0$): $x_b^B \geq 0 \iff v \leq \frac{T}{T-1} + \frac{T-i}{T-1}\tilde{t}$ which implies $v \leq \frac{T}{T-1}$ when $\tilde{t} = 0$ (recall $\tilde{t} = \frac{t}{c}$). The "low-enough-trade-cost" restriction ensures that low-quality B-country firms' exports to country A are not priced out of the market: $x_b^A \geq 0 \iff \tilde{t} \leq \frac{T-(T-1)v}{(T-1)v}$. Denote by $\mathcal{P}$ the set of parameter values $(v, \tilde{t})$ satisfying these two conditions[7].

[7] Formally, space $\mathcal{P}$ is defined as $\{(v,\tilde{t}) \in \Re^2 \mid 1 \leq v \leq \frac{T}{T-1}$ and $0 \leq \tilde{t} \leq \frac{T-(T-1)v}{(T-1)v}\}$. Alternatively, $\mathcal{P}$ can be defined by the restrictions $0 \leq \tilde{t} \leq \frac{1}{T-1}$ and $1 \leq v \leq \frac{T}{T-1}\frac{1}{1+\tilde{t}}$.

The reason why the analysis in Section 1.3 is confined to $\mathcal{P}$ is simplicity. Lifting these parameter restrictions adds little insight: extending the space of parameter values to $\{(v,\tilde{t}) \in \Re^2 \mid v \geq 1$ and $\tilde{t} \geq 0\}$ would enlarge zone (a), the zone where mergers are always profitable, both along and off the equilibrium path of the game.

Proof of this claim is available from the author, but headway can be made by noting that, say, for the case considered in Section 1.3 (i.e. $T = 3$): (i) independent B-country firms do not command positive sales *even* in their home country, $x_b^A = x_b^B = 0$, when $v \geq \frac{3}{2} + \frac{3}{2}\tilde{t}$ and $\tilde{t} \geq 0$ under market structure r_0, or when $v \geq \frac{3}{2} + \tilde{t}$ and $\tilde{t} \geq 0$ under r_1, or when $v \geq \frac{3}{2} + \frac{1}{2}\tilde{t}$

Prices are obtained noting that $p_b^l = \lambda^l$, and from equation (4.10), $p_m^l = p_a^l = v\lambda^l$, $l \in \{A, B\}$. Finally, evaluating the objective functions at these sales and prices, profits per firm in each market are obtained:

$$
\begin{array}{llll}
\Pi_{m,r_i}^A = \left(\frac{(T-i)v-(T-i-1)+(T-i)\tilde{t}v}{T+(T-i)v+(T-i)\tilde{t}v}\right)^2 S_A & \Pi_{m,r_i}^B = \left(\frac{(T-i)v-(T-i-1)+(T-i)\tilde{t}}{T+(T-i)v+(T-i)\tilde{t}}\right)^2 S_B \\
\Pi_{a,r_i}^A = x_{m,r_i}^A & \Pi_{a,r_i}^B = \left(\frac{(T-i)v-(T-i-1)-(T-1)\tilde{t}}{T+(T-i)v+(T-i)\tilde{t}}\right)^2 S_B \\
\Pi_{b,r_i}^A = \left(\frac{T-(T-1)v-(T-1)\tilde{t}v}{T+(T-i)v+(T-i)\tilde{t}v}\right)^2 S_A & \Pi_{b,r_i}^B = \left(\frac{T-(T-1)v+(T-i)\tilde{t}}{T+(T-i)v+(T-i)\tilde{t}}\right)^2 S_B
\end{array}
\quad (4.17)
$$

The reduced-form profit functions per firm are thus the sum of the profit components in the two markets, e.g. for a multinational firm, $\Pi_{m,r_i} = \Pi_{m,r_i}^A + \Pi_{m,r_i}^B$.

While the derivation above was carried out for "intermediate" market structures where both independent and multinational firms exist, it similarly applies to structures r_0 (where there are no multinational firms) and r_T (where there are no independent firms). It is easy to see that the reduced-form profit functions (4.17) – as well as the outputs and prices derived above) – also hold where applicable. In other words, Π_{m,r_i} calculated from (4.17) holds for $i = 1, 2, 3, ..., T$, while Π_{a,r_i} and Π_{b,r_i} hold for $i = 0, 1, 2, ..., T-1$.

Market competition equilibrium outcomes as a function of market structure (case $T = 3$)

The reduced-form profit functions used to compute the merger surplus functions (Definition 1.3) follow from plugging $T = 3$ in equations (4.17) for the general case considered above (T merger stages and initially T independent firms in each country). The space of parameter values $\mathcal{P}$ follows similarly from the space derived for the general case.

In view of Lemma 1.1, I reproduce the three merger surplus functions which

---

and $\tilde{t} \geq 0$ under r_2; (ii) trade is too expensive *even* for A-country firms, $x_a^B = x_a^A = 0$, when $1 \leq v \leq \frac{2}{3} + \frac{2}{3}\tilde{t}$ under r_0, or when $1 \leq v \leq \frac{1}{2} + \tilde{t}$ under r_1, or when $1 \leq v \leq 2\tilde{t}$ under r_2. Note that for $\frac{3}{2} + \frac{1}{2}\tilde{t} \leq v \leq 2\tilde{t}$ under r_2, only two (multinational) firms command positive sales in country B, unlike all other parameter combinations and market structures where the number of firms selling into each country is at least three.

need to be signed in order to solve the three-merger-stage game in Section 1.3:

$$\Psi_I(v,\tilde{t}) = \left[\left(\frac{2v-1+2\tilde{t}v}{3+2v+2\tilde{t}v}\right)^2 - \frac{\left(3v-2+3\tilde{t}v\right)^2+\left(3-2v-2\tilde{t}v\right)^2}{9\left(1+v+\tilde{t}v\right)^2}\right] S_A + \\ + \left[\left(\frac{2v-1+2\tilde{t}}{3+2v+2\tilde{t}}\right)^2 - \frac{\left(3v-2-2\tilde{t}\right)^2+\left(3-2v+3\tilde{t}\right)^2}{9\left(1+v+\tilde{t}\right)^2}\right] S_B$$

$$\Psi_{II}(v,\tilde{t}) = \left[\left(\frac{v+\tilde{t}v}{3+v+\tilde{t}v}\right)^2 - \frac{\left(2v-1+2\tilde{t}v\right)^2+\left(3-2v-2\tilde{t}v\right)^2}{\left(3+2v+2\tilde{t}v\right)^2}\right] S_A + \\ + \left[\left(\frac{v+\tilde{t}}{3+v+\tilde{t}}\right)^2 - \frac{\left(2v-1-2\tilde{t}\right)^2+\left(3-2v+2\tilde{t}\right)^2}{\left(3+2v+2\tilde{t}\right)^2}\right] S_B$$

$$\Psi_{III}(v,\tilde{t}) = \left[\frac{1}{9} - \frac{\left(v+\tilde{t}v\right)^2+\left(3-2v-2\tilde{t}v\right)^2}{\left(3+v+\tilde{t}v\right)^2}\right] S_A + \\ + \left[\frac{1}{9} - \frac{\left(v-2\tilde{t}\right)^2+\left(3-2v+\tilde{t}\right)^2}{\left(3+v+\tilde{t}\right)^2}\right] S_B$$

A.2.3 Proof of Lemma 1.1

Part (i) This proof amounts to verifying that for all $(v,\tilde{t}) \in \mathcal{P}$:

- $\Psi_{II}(v,\tilde{t}) \geq 0$ whenever $\Psi_I(v,\tilde{t}) \geq 0$,
- $\Psi_{III}(v,\tilde{t}) \geq 0$ whenever $\Psi_{II}(v,\tilde{t}) \geq 0$.

This is done separately for the merger surplus emanating from each market: a second subscript is added to denote the market. Thus, for example, $\Psi_{I,A}(v,\tilde{t}) = \Pi^A_{m,r_1} - \Pi^A_{a,r_0} - \Pi^A_{b,r_0}$ denotes the terms in $\Psi_I(v,\tilde{t})$ corresponding to market A, obtained from equations (4.17), for $T=3$, and reproduced at the end of A.2.2.

I begin with $\Psi_{I,A}(v,\tilde{t})$, $\Psi_{II,A}(v,\tilde{t})$ and $\Psi_{III,A}(v,\tilde{t})$. Consider a straight line segment going from $(v,\tilde{t}) = (1,0)$ to any point on $\tilde{t} = \frac{3-2v}{2v}$, the boundary of $\mathcal{P}$ where the "low-enough-trade-cost" restriction binds (i.e. $x^A_b = 0$). By writing this line segment as $\tilde{t} = \rho(v-1)$, where $0 \leq \rho \leq \infty$ and $0 \leq \rho(v-1) \leq \frac{3-2v}{2v}$, changes in v and $\tilde{t}$ along this line segment parameterised by ρ may be referred to simply as changes in v (for $\rho = 0$ the line segment lies on the v-axis; for $\rho \to \infty$ the line segment lies on the $\tilde{t}$-axis). I seek to uncover how the functions[8] $\Psi_{I,A}$, $\Psi_{II,A}$ and $\Psi_{III,A}$ change as I increase v along line segment ρ (i.e. jointly increasing $\tilde{t}$ such that $\tilde{t} = \rho(v-1)$) from the lower end $v = 1$ (i.e. $(v, \tilde{t}) = (1, 0)$) to the upper end defined implicitly by $\rho(v - 1) = \frac{3-2v}{2v}$ (label this value $v = \bar{v}$; formally this label should carry the parameter ρ, omitted for simplicity). The following may be verified. At the lower end of the line segment, when $v = 1$, all three functions are negative. Intuitively, merging firms' surplus arising from any stand-alone cross-border merger (be this the first, the second or the third) is negative: since both v and $\tilde{t}$ are low, non-participating firms respond to the merger by increasing output considerably. At the upper end of the line segment, when $v = \bar{v}$, all three functions are equal to zero. Intuitively, since $x_b^A = 0$, there is no surplus to be enjoyed on sales in country A from cross-border mergers when B-country firms' exports to country A are (just) priced out of the market. Prior to any merger, there are three firms offering quality v commanding positive sales in country A; after a merger this is unchanged. Now, starting at $v = 1$ and increasing v along the line segment, $\Psi_{I,A}$, $\Psi_{II,A}$ and $\Psi_{III,A}$ each increase continuously from negative values toward positive values, reaching a maximum, then decreasing continuously toward zero when $v = \bar{v}$. Label the first value of v at which $\Psi_{I,A}$ is zero as v'_A, the first value of v at which $\Psi_{II,A}$ is zero as v''_A, and the first value of v at which $\Psi_{III,A}$ is zero as v'''_A (again the labels omit the reference to ρ for simplicity). One can verify that $0 < v'''_A < v''_A < v'_A < \bar{v}$. Since this is true along any line segment parameterised by ρ, $0 \leq \rho \leq \infty$, the following result holds:

$$\left\{(v,\tilde{t}) \in \mathcal{P} \mid \Psi_{I,A}(v,\tilde{t}) \geq 0\right\} \subset \left\{(v,\tilde{t}) \in \mathcal{P} \mid \Psi_{II,A}(v,\tilde{t}) \geq 0\right\}$$

$$\subset \left\{(v,\tilde{t}) \in \mathcal{P} \mid \Psi_{III,A}(v,\tilde{t}) \geq 0\right\} \subset \mathcal{P} \tag{4.18}$$

It may also be verified that, in addition to intersecting at $v = \bar{v}$, $\Psi_{II,A}$ and $\Psi_{I,A}$ cross as they slope upwards at a point, labelled $v_A^{''-'}$, which lies between 0 and v''_A. In other words, $\Psi_{II,A} - \Psi_{I,A} = 0$ at $v = v_A^{''-'}$, where $0 < v_A^{''-'} < v''_A$. To the right of

[8] For simplicity I drop the arguments of the functions.

this point, for $v_A^{\prime\prime-\prime} < v < \bar{v}$, $\Psi_{II,A} - \Psi_{I,A} > 0$, whereas to its left, for $0 \leq v < v_A^{\prime\prime-\prime}$, $\Psi_{II,A} - \Psi_{I,A} < 0$. Since $0 < v_A^{\prime\prime-\prime} < v_A^{\prime\prime} < v_A^{\prime} < \bar{v}$, the following (stronger) result holds:

$$\{(v,\tilde{t}) \in \mathcal{P} \mid \Psi_{I,A}(v,\tilde{t}) \geq 0\} \subset \{(v,\tilde{t}) \in \mathcal{P} \mid \Psi_{II,A}(v,\tilde{t}) - \Psi_{I,A}(v,\tilde{t}) \geq 0\} \quad (4.19)$$

It can further be verified that $\Psi_{III,A} - \Psi_{II,A} > 0$ for $0 \leq v < \bar{v}$ (recall $\Psi_{III,A} = \Psi_{II,A}$ when $v = \bar{v}$) and thus

$$\forall\ (v,\tilde{t}) \in \mathcal{P},\ \Psi_{III,A}(v,\tilde{t}) - \Psi_{II,A}(v,\tilde{t}) \geq 0 \quad (4.20)$$

Turning now to the merger surplus terms in market B, the same results come through, despite some differences in the corresponding functions which are detailed next. The values of the functions $\Psi_{I,B}$, $\Psi_{II,B}$ and $\Psi_{III,B}$ are no longer zero for *all* $(v,\tilde{t})$ along the border of $\mathcal{P}$ for which the "low-enough-trade-cost" restriction binds. For these parameter values, unlike in country A, B-country firms still command positive sales in their home country and hence there is surplus from merger to be made on sales in country B (i.e. the functions are strictly positive), with two exceptions. One is when imports are at their most competitive, $(v,\tilde{t}) = (\frac{3}{2},0)$, and B-country firms command zero sales in their home country (as they do abroad), in which case $\Psi_{I,B}$, $\Psi_{II,B}$ and $\Psi_{III,B}$ again equal zero. The other situation along this border of $\mathcal{P}$ where these functions equal zero occurs when $(v,\tilde{t}) = (1,\frac{1}{2})$. Here, quality is symmetric and markets are effectively autarkic: cross-border merger does not change the (effective) number of competitors in each market. Now, starting at $v = 1$ and increasing v along any line segment parameterised by ρ, $\Psi_{I,B}$, $\Psi_{II,B}$ and $\Psi_{III,B}$ each increase continuously from negative values toward positive values; whether the functions increase monotonically or whether they reach a (positive) maximum before decreasing to a lower albeit positive value or zero when $v = \bar{v}$ depends on the line segment parameterised by ρ. Labelling the lower and possibly only value of v at which $\Psi_{I,B}$ ($\Psi_{II,B}$, $\Psi_{III,B}$) is zero as $v_B^{\prime}$ ($v_B^{\prime\prime}$, $v_B^{\prime\prime\prime}$ respectively), one verifies that $0 < v_B^{\prime\prime\prime} < v_B^{\prime\prime} < v_B^{\prime} < \bar{v}$. Since this is true along any line segment parameterised by ρ, $0 \leq \rho \leq \infty$, a result analogous to (4.18) holds:

$$\left\{(v,\tilde{t}) \in \mathcal{P} \mid \Psi_{I,B}(v,\tilde{t}) \geq 0\right\} \subset \left\{(v,\tilde{t}) \in \mathcal{P} \mid \Psi_{II,B}(v,\tilde{t}) \geq 0\right\}$$

$$\subset \left\{(v,\tilde{t}) \in \mathcal{P} \mid \Psi_{III,B}(v,\tilde{t}) \geq 0\right\} \subset \mathcal{P} \tag{4.21}$$

Similar to their country A counterparts, $\Psi_{II,B}$ and $\Psi_{I,B}$ cross as they slope upwards at a point, labelled $v_B^{\prime\prime-\prime}$, which lies between 0 and $v_B^{\prime\prime}$. To the right of this point, for $v_B^{\prime\prime-\prime} < v < \bar{v}$, $\Psi_{II,B} - \Psi_{I,B} > 0$[9], whereas to its left, for $0 \leq v < v_B^{\prime\prime-\prime}$, $\Psi_{II,B} - \Psi_{I,B} < 0$. Since $0 < v_B^{\prime\prime-\prime} < v_B^{\prime\prime} < v_B^{\prime} < \bar{v}$, the following (stronger) result holds:

$$\left\{(v,\tilde{t}) \in \mathcal{P} \mid \Psi_{I,B}(v,\tilde{t}) \geq 0\right\} \subset \left\{(v,\tilde{t}) \in \mathcal{P} \mid \Psi_{II,B}(v,\tilde{t}) - \Psi_{I,B}(v,\tilde{t}) \geq 0\right\} \tag{4.22}$$

It can further be verified that $\Psi_{III,B} - \Psi_{II,B} > 0$ for $0 \leq v < \bar{v}$[10] and thus

$$\forall\ (v,\tilde{t}) \in \mathcal{P},\ \Psi_{III,B}(v,\tilde{t}) - \Psi_{II,B}(v,\tilde{t}) \geq 0 \tag{4.23}$$

Finally, it can be verified that along all line segments parameterised by ρ, $0 \leq \rho \leq \infty$, $v_B^{\prime\prime-\prime} < v_A^{\prime}$ and $v_A^{\prime\prime-\prime} < v_B^{\prime}$.

The results for both markets are now combined to conclude the proof. I wish to show that for $(v,\tilde{t})$ such that $\Psi_I(v,\tilde{t}) = \Psi_{I,A}(v,\tilde{t}) + \Psi_{I,B}(v,\tilde{t}) \geq 0$, then $\Psi_{II}(v,\tilde{t}) = \Psi_{II,A}(v,\tilde{t}) + \Psi_{II,B}(v,\tilde{t}) \geq 0$. (I also need to show that for $(v,\tilde{t})$ such that $\Psi_{II}(v,\tilde{t}) = \Psi_{II,A}(v,\tilde{t}) + \Psi_{II,B}(v,\tilde{t}) \geq 0$, then $\Psi_{III}(v,\tilde{t}) = \Psi_{III,A}(v,\tilde{t}) + \Psi_{III,B}(v,\tilde{t}) \geq 0$. This is postponed briefly.) For $(v,\tilde{t})$ such that $\Psi_I \geq 0$ this may be due to either of three possibilities:

- $\Psi_{I,A} \geq 0$ and $\Psi_{I,B} \geq 0$. From (4.18), $\Psi_{I,A} \geq 0 \Longrightarrow \Psi_{II,A} \geq 0$, while from (4.21), $\Psi_{I,B} \geq 0 \Longrightarrow \Psi_{II,B} \geq 0$. Hence

$$\left\{(v,\tilde{t}) \in \mathcal{P} \mid \Psi_I(v,\tilde{t}) \geq 0\right\} \subset \left\{(v,\tilde{t}) \in \mathcal{P} \mid \Psi_{II}(v,\tilde{t}) \geq 0\right\} \tag{4.24}$$

[9] $\Phi_{II,B} - \Phi_{I,B} > 0$ for $v_B^{\prime\prime-\prime} < v \leq \bar{v}$ when $0 < \rho < \infty$; i.e. only when $\rho = 0$ or $\rho \to \infty$ does $\Phi_{II,B} - \Phi_{I,B} = 0$ when $v = \bar{v}$.

[10] The previous footnote similarly applies: only when $\rho = 0$ or $\rho \to \infty$ does $\Phi_{III,B} - \Phi_{II,B} = 0$ when $v = \bar{v}$.

from which the first half of Lemma 1.1, part (i) follows.

- $\Psi_{I,A} > 0$, $\Psi_{I,B} < 0$ but $\Psi_{I,A} + \Psi_{I,B} \geq 0$. Here, since $\Psi_{I,A} > 0$ and $\Psi_{I,B} < 0$, v must be greater than v'_A but less than v'_B, i.e. $v'_A < v < v'_B$. As verified earlier, $v_B^{\prime\prime-\prime} < v'_A$, whence it follows that $v_B^{\prime\prime-\prime} < v$ and thus $\Psi_{II,B} - \Psi_{I,B} > 0$. From (4.19), $\Psi_{I,A} \geq 0$ implies that $\Psi_{II,A} - \Psi_{I,A} \geq 0$. Since $\Psi_{I,A} + \Psi_{I,B} \geq 0$ it must be that $\Psi_{II,A} + \Psi_{II,B} > 0$ and result (4.24) follows.

- $\Psi_{I,A} < 0$, $\Psi_{I,B} > 0$ but $\Psi_{I,A} + \Psi_{I,B} \geq 0$. Here, since $\Psi_{I,A} < 0$ and $\Psi_{I,B} > 0$, v must be greater than v'_B but less than v'_A, i.e. $v'_B < v < v'_A$. As verified earlier, $v_A^{\prime\prime-\prime} < v'_B$, whence it follows that $v_A^{\prime\prime-\prime} < v$ and thus $\Psi_{II,A} - \Psi_{I,A} > 0$. Again, since $\Psi_{I,A} + \Psi_{I,B} \geq 0$ it must be that $\Psi_{II,A} + \Psi_{II,B} > 0$ and result (4.24) follows.

I now show that for $(v, \tilde{t})$ such that $\Psi_{II} = \Psi_{II,A} + \Psi_{II,B} \geq 0$, it must be that $\Psi_{III} = \Psi_{III,A} + \Psi_{III,B} \geq 0$. The proof is a simplified version of the previous one, and follows from noting by (4.20) and (4.23) that $\forall\ (v, \tilde{t}) \in \mathcal{P}$, $\Psi_{III,A} - \Psi_{II,A} \geq 0$ and $\Psi_{III,B} - \Psi_{II,B} \geq 0$. Thus the counterpart to result (4.24) for the second half of Lemma 1.1 is obtained:

$$\left\{(v,\tilde{t}) \in \mathcal{P} \mid \Psi_{II}(v,\tilde{t}) \geq 0\right\} \subset \left\{(v,\tilde{t}) \in \mathcal{P} \mid \Psi_{III}(v,\tilde{t}) \geq 0\right\} \qquad (4.25)$$

Summarising results (4.24) and (4.25),

$$\left\{(v,\tilde{t}) \in \mathcal{P} \mid \Psi_{I}(v,\tilde{t}) \geq 0\right\} \subset \left\{(v,\tilde{t}) \in \mathcal{P} \mid \Psi_{II}(v,\tilde{t}) \geq 0\right\} \subset \left\{(v,\tilde{t}) \in \mathcal{P} \mid \Psi_{III}(v,\tilde{t}) \geq 0\right\} \subset \mathcal{P} \qquad (4.26)$$

from which Lemma 1.1, part (i) follows. *Q.E.D.*

Part (ii) Recall from Definition 1.3 that $\Psi_{II}(v, \tilde{t}) \geq 0$ is equivalent to $\Pi_{m,r_2} \geq \Pi_{a,r_1} + \Pi_{b,r_1}$. It can be verified that (for $(v, \tilde{t}) \in \mathcal{P}$) $\Pi_{a,r_1} + \Pi_{b,r_1} \geq \Pi_{a,r_0} + \Pi_{b,r_0}$ (i.e. non-participating independent firms gain from a first stand-alone merger). Hence if $\Psi_{II} \geq 0$, it follows that $\Pi_{m,r_2} \geq \Pi_{a,r_0} + \Pi_{b,r_0}$, which is equivalent to $\Psi_{IV} \geq 0$.

Similarly, $\Psi_{III} \geq 0$ is equivalent to $\Pi_{m,r_3} \geq \Pi_{a,r_2} + \Pi_{b,r_2}$. It can be verified that $\Pi_{a,r_2} + \Pi_{b,r_2} \geq \Pi_{a,r_1} + \Pi_{b,r_1}$ (i.e. non-participating independent firms gain from a second stand-alone merger). Hence if $\Psi_{III} \geq 0$, it follows that $\Pi_{m,r_3} \geq \Pi_{a,r_1} + \Pi_{b,r_1}$, which is equivalent to $\Psi_V \geq 0$. From $\Pi_{a,r_1} + \Pi_{b,r_1} \geq \Pi_{a,r_0} + \Pi_{b,r_0}$, it follows from $\Psi_V \geq 0$ that $\Pi_{m,r_3} \geq \Pi_{a,r_0} + \Pi_{b,r_0}$, which is equivalent to $\Psi_{VI} \geq 0$. *Q.E.D.*

A.2.4 (Sketch of) Proof of Proposition 1.3

This proof follows largely from the proof of Lemma 1.1, noting the following definitions:

- Zone (a):= $\{(v, \tilde{t}) \in \mathcal{P} \mid \Psi_I(v,\tilde{t}) \geq 0,\ \Psi_{II}(v,\tilde{t}) \geq 0,\ \Psi_{III}(v,\tilde{t}) \geq 0\}$
- Zone (c1):= $\{(v, \tilde{t}) \in \mathcal{P} \mid \Psi_I(v,\tilde{t}) < 0,\ \Psi_{II}(v,\tilde{t}) \geq 0,\ \Psi_{III}(v,\tilde{t}) \geq 0\}$
- Zone (c2):= $\{(v, \tilde{t}) \in \mathcal{P} \mid \Psi_I(v,\tilde{t}) < 0,\ \Psi_{II}(v,\tilde{t}) < 0,\ \Psi_{III}(v,\tilde{t}) \geq 0\}$
- Zone (h):= $\{(v, \tilde{t}) \in \mathcal{P} \mid \Psi_I(v,\tilde{t}) < 0,\ \Psi_{II}(v,\tilde{t}) < 0,\ \Psi_{III}(v,\tilde{t}) < 0\}$

That the zones and the boundaries between them are as stated (and depicted in Figure 1-6) follows additionally from verifying that:

- $\bar{v}$, v'_A, v''_A, v'''_A, v'_B, v''_B and v'''_B are decreasing in ρ, the parameter of the straight line segment, as this increases from 0 (line segment lies on top of the v-axis) to ∞ (line segment lies on top of the $\tilde{t}$-axis).
- $v'''_B \leq v'''_A < v''_B \leq v''_A$ for $0 \leq \rho \leq \infty$.
- $v''_B \leq v''_A \leq v'_B \leq v'_A$ for $0 \leq \rho \leq \check{\rho} < \infty$ (line segment steep enough) and $0 < \hat{\rho} \leq \rho \leq \infty$ (line segment flat enough), where $\check{\rho} < \hat{\rho}$, while $v''_B < v'_B < v''_A < v'_A$ for $\check{\rho} < \rho < \hat{\rho}$.
- $v''^{-\prime}_A < v''_B$ and $v''^{-\prime}_B < v''_B$ so that $\Psi_{II,A} > \Psi_{I,A}$ and $\Psi_{II,B} > \Psi_{I,B}$ for $v''_B < v < v'_A$ along any line segment ρ.

These elements suffice to prove Proposition 1.3. *Q.E.D.*

A.2.5 Proof of results depicted in Figure 1-8

I consider the equilibrium in each of zones (a), (c1) and (c2) of parameter space $\mathcal{P}$ in turn, starting from a fixed cost G associated with implementing a cross-border merger equal to zero (equilibria as in Figure 1-5, in the absence of fixed costs). I analyse how these equilibria change as G increases from zero.

Notice that the fixed cost G of implementing a merger changes the conditions for any merger to be profitable: whereas in the absence of the fixed cost this was given by $\Psi_X(v,\tilde{t}) \geq 0$, for $X \in \{I, II, III, IV, V, VI\}$, where X represents the relevant merger surplus function, the introduction of the fixed cost changes this condition to $\Psi_X(v,\tilde{t}) - G \geq 0$. (Refer to Definition 1.3. Recall that, for example, $\Psi_V(v,\tilde{t})$ reflects the profitability of a merger which if carried through would induce a change in market structure from r_1 (one cross-border merger) to r_3 (three cross-border mergers).

Note from the Proof of Lemma 1.1, part (ii) above[11] that for all $(v,\tilde{t}) \in \mathcal{P}$, $\Psi_{IV} \geq \Psi_{II}$ and $\Psi_{VI} \geq \Psi_V \geq \Psi_{III}$.

I begin with $(v,\tilde{t})$ in zone (a). From the Proof of Lemma 1.1, part (i), in this zone $\Psi_{III} \geq \Psi_{II} \geq \Psi_I \geq 0$. For $G \leq \Psi_I$, the equilibrium (solved-out game tree) replicates the equilibrium for zone (a) in the absence of fixed costs (Figure 1-5), where firms choose to merge from all nodes in the game tree. For $\Psi_I < G \leq \Psi_{II}$, we have that $\Psi_I - G < 0$ and thus firms 3 and 6 will now choose not to merge conditional on $\overline{\mathbf{14}}$ and $\overline{\mathbf{25}}$. By $\Psi_{II} - G \geq 0$, $\Psi_{III} \geq \Psi_{II}$ and $\Psi_{IV} \geq \Psi_{II}$, firms continue choosing to merge from all other nodes in the game tree. For this range of values of G, the equilibrium then replicates the equilibrium for zone (c1) in the absence of fixed costs (Figure 1-5). For $\Psi_{II} < G \leq \Psi_{III}$, $\Psi_{II} - G < 0$ implies that firms 3 and 6 will now choose not to merge conditional on either **14** and $\overline{\mathbf{25}}$, or $\overline{\mathbf{14}}$ and **25**. Since $\Psi_I - G < 0$, firms 3 and 6 will still choose not to merge conditional on $\overline{\mathbf{14}}$ and $\overline{\mathbf{25}}$, and firms 2 and 5 will now choose not to merge conditional on $\overline{\mathbf{14}}$. By $\Psi_{III} - G \geq 0$, $\Psi_V \geq \Psi_{III}$ and $\Psi_{VI} \geq \Psi_{III}$, firms continue choosing to merge

[11] For ease of exposition, I again suppress the arguments $(v,\tilde{t})$ of the merger surplus functions. I also introduce the notation **ij** to depict the outcome where firms i and j merge and $\overline{\mathbf{ij}}$ as the complementary (no-merger) outcome.

from all other nodes in the game tree. For this range of values of G, the equilibrium then replicates the equilibrium for zone (c2) in the absence of fixed costs (Figure 1-5). For $G > \Psi_{III}$, firms 3 and 6 will now choose not to merge conditional on **14** and **25**. Since $\Psi_{II} - G < 0$, firms 3 and 6 will still choose not to merge conditional on either **14** and $\overline{\mathbf{25}}$, or $\overline{\mathbf{14}}$ and **25**, and firms 2 and 5 will now choose not to merge conditional on **14**. Since $\Psi_{I} - G < 0$, firms 3 and 6 will still choose not to merge conditional on $\overline{\mathbf{14}}$ and $\overline{\mathbf{25}}$, firms 2 and 5 will still choose not to merge conditional on $\overline{\mathbf{14}}$, and firms 1 and 4 will now choose not to merge. For this range of values of G, along the equilibrium path no mergers occur and the equilibrium then replicates the equilibrium for zone (h) in the absence of fixed costs (Figure 1-5).

The proofs of the equilibria for $(v, \tilde{t})$ in zones (c1) and (c2) follow from that of zone (a). In zone (c1), by definition, $\Psi_I < 0 \leq \Psi_{II}$, and only the equilibria for zones (b), (c2) and (h) in the absence of fixed costs can be replicated as G increases from zero, as analysed for zone (a). In zone (c2), by definition, $\Psi_{III} \geq 0$ and $\Psi_I, \Psi_{II} < 0$, and only the equilibria for zones (c2) and (h) in the absence of fixed costs can be replicated as G increases from zero, as analysed for zone (a). *Q.E.D.*

A.3 Cross-border mergers in the Perry and Porter (1985) model

A.3.1 Derivation of reduced-form profit functions

Coming out of the (two) merger stages, there are three possible market structures: (i) r_0, where no cross-border is undertaken; (ii) r_1, where one merger decision is favourable but the other is not; and (iii) r_2, where both mergers take place. I now turn to each.

(i) Under r_0, independent A-country firm 1 (say), owning capital stock k, sets outputs in both countries A and B, respectively x_1^A and x_1^B, to solve

$$\max_{x_1^A\geq 0, x_1^B\geq 0} P^A(X^A)x_1^A + \left(P^B(X^B) - t\right)x_1^B - C(x_1^A + x_1^B, k)$$

where its exports to country B are subject to the unit trade cost t. (Recall that $P^l(.)$ denotes the inverse demand function for the homogeneous good in country l, $l \in \{A, B\}$, $X^l = \sum_{i=1}^4 x_i^l$, and $C(x,k)$ denotes the firm's cost function.) Given the functional forms laid out in Section 1.4, the FOCs may be written:

$$a - X^A - x_1^A - \left(d + \frac{e}{k}(x_1^A + x_1^B)\right) = 0 \qquad (4.27)$$

$$a - X^B - t - x_1^B - \left(d + \frac{e}{k}(x_1^A + x_1^B)\right) = 0 \qquad (4.28)$$

The FOCs for the other independent firms, namely A-country firm 2 and B-country firms 3 and 4, can be similarly written, adjusting for the trade cost being incurred by the latter two on their exports to country A. Solving the system of FOCs, and recalling $\tilde{e} = \frac{e}{k}$ (the rate of change of marginal cost) and $\tilde{t} = \frac{t}{a-d}$ (the normalised trade cost), one obtains $X^A = X^B = \frac{(a-d)(4-2\tilde{t})}{5+2\tilde{e}}$. From FOCs (4.27) and (4.28), it follows that $\left|x_i^A - x_i^B\right| = t$ (firm i's sales in its home market exceed its foreign-market sales by t, where $i = 1, ..., 4$) and each firm's foreign-market sales are $x_a^B = x_b^A = \frac{(a-d)\left(1-\tilde{t}(3+\tilde{e})\right)}{5+2\tilde{e}}$. (As before, despite the slight abuse of notation, subscripts a, b and m denote an independent A-country firm, an independent B-country firm and a multinational firm, respectively.) Clearly, for trade between countries to be feasible, the parameter restriction $\tilde{t} \leq \frac{1}{3+\tilde{e}}$ must be satisfied (this condition along with $\tilde{e} \geq 0$ and $\tilde{t} \geq 0$ define space $\mathcal{P}$). Equilibrium prices in both countries are $p^A = p^B = \frac{a(1+2\tilde{e})+4d+2t}{5+2\tilde{e}}$. The reduced-form profit function follows from evaluating each firm's objective function:

$$\Pi_{a,r_0} = \Pi_{b,r_0} = \frac{(a-d)^2}{(5+2\tilde{e})^2}\left(2(1+\tilde{e})(1-\tilde{t}) + \frac{\tilde{t}^2}{2}(2+\tilde{e})(13+4\tilde{e})\right) - gk$$

(ii) Under r_1, each of the two independent firms a and b solves the same problem as in (i). The multinational firm m, formed from the merger of an independent A-country firm and an independent B-country firm, owns capital stock $2k$ and is twice as "large" as either of its independent rivals. Clearly, given the unit trade cost $t \geq 0$ and the cost function $C(x_m^A + x_m^B, 2k)$, it will no longer trade between countries, supplying each country through domestic production; it solves

$$\max_{x_m^A \geq 0, x_m^B \geq 0} P^A(X^A)x_m^A + P^B(X^B)x_m^B - C(x_m^A + x_m^B, 2k)$$

The FOCs become

$$a - X^A - x_m^A - \left(d + \frac{e}{2k}(x_m^A + x_m^B)\right) = 0 \tag{4.29}$$

$$a - X^B - x_m^B - \left(d + \frac{e}{2k}(x_m^A + x_m^B)\right) = 0 \tag{4.30}$$

Solving the system of FOCs, one obtains sales in each country, $X^A = X^B = \frac{(a-d)(3+4\tilde{e})-\tilde{t}(1+\tilde{e})}{4+7\tilde{e}+2\tilde{e}^2}$. The multinational firm's sales in each country are $x_m^A = x_m^B = \frac{(a-d)(1+2\tilde{e})+\tilde{t}}{4+7\tilde{e}+2\tilde{e}^2}$. Independent firms' home-market sales exceed foreign-market sales by t, i.e. $\left|x_i^A - x_i^B\right| = t$, $i \in \{a, b\}$, where foreign-market sales $x_a^B = x_b^A = \frac{(a-d)(1+\tilde{e})\left(1-\tilde{t}(3+\tilde{e})\right)}{4+7\tilde{e}+2\tilde{e}^2}$. Equilibrium prices are $p^A = p^B = \frac{a(1+\tilde{e})(1+2\tilde{e})+d(3+4\tilde{e})+t(1+\tilde{e})}{4+7\tilde{e}+2\tilde{e}^2}$. Finally, the reduced-form profit functions are given by:

$$\Pi_{m,r_1} = \frac{(a-d)^2}{(4+7\tilde{e}+2\tilde{e}^2)^2}\left((1+2\tilde{e})^2(2+\tilde{e}) + 2\tilde{t}(1+2\tilde{e})(2+\tilde{e}) + \tilde{t}^2(2+\tilde{e})\right) - 2gk$$

$$\begin{aligned}\Pi_{a,r_1} &= \Pi_{b,r_1} \\ &= \frac{(a-d)^2}{(4+7\tilde{e}+2\tilde{e}^2)^2} \times \\ &\quad \times \left(2(1+\tilde{e})^3 - 2\tilde{t}(1+\tilde{e})^2(2+\tilde{e}) + \tilde{t}^2(10+32\tilde{e}+35\tilde{e}^2+\frac{29}{2}\tilde{e}^3+2\tilde{e}^4)\right) - gk\end{aligned}$$

(iii) Under r_2, there are no shipments across countries, each multinational firm solving the same problem as in (ii) (FOCs given by (4.29) and (4.30)). Equilibrium sales in each country are $X^A = X^B = \frac{2(a-d)}{3+\tilde{e}}$, while $x_m^A = x_m^B = \frac{a-d}{3+\tilde{e}}$; prices are $p^A = p^B = \frac{a(1+\tilde{e})+2d}{3+\tilde{e}}$. The reduced-form profit function for each multinational firm is

$$\Pi_{m,r_2} = \frac{(a-d)^2(2+\tilde{e})}{(3+\tilde{e})^2} - 2gk$$

By Proposition 1.2, the four merger surplus functions necessary to solve the game can then be computed:

$$
\begin{aligned}
\Phi_I &= \Pi_{m,r_1} - \Pi_{a,r_0} - \Pi_{b,r_0} \\
&= \frac{(a-d)^2(1-3\tilde{t}-\tilde{t}\tilde{e})}{(5+2\tilde{e})^2(4+7\tilde{e}+2\tilde{e}^2)^2} \times \\
&\quad \times \left(16\tilde{t}\tilde{e}^5 + 148\tilde{t}\tilde{e}^4 + 508\tilde{t}\tilde{e}^3 + 4\tilde{e}^3 + 4\tilde{e}^2 + 789\tilde{t}\tilde{e}^2 + 535\tilde{t}\tilde{e} - 23\tilde{e} + 122\tilde{t} - 14\right)
\end{aligned}
$$

$$
\begin{aligned}
\Phi_{II} &= \Pi_{m,r_2} - \Pi_{a,r_1} - \Pi_{b,r_1} \\
&= \frac{(a-d)^2(1-3\tilde{t}-\tilde{t}\tilde{e})}{(3+\tilde{e})^2(4+7\tilde{e}+2\tilde{e}^2)^2} \times \\
&\quad \times \left(4\tilde{t}\tilde{e}^5 + 41\tilde{t}\tilde{e}^4 + 157\tilde{t}\tilde{e}^3 + \tilde{e}^3 + 2\tilde{e}^2 + 274\tilde{t}\tilde{e}^2 + 212\tilde{t}\tilde{e} - 4\tilde{e} + 60\tilde{t} - 4\right)
\end{aligned}
$$

$$
\begin{aligned}
\Phi_{III} &= \Pi_{m,r_2} - \Pi_{a,r_0} - \Pi_{b,r_0} \\
&= \frac{(a-d)^2(1-3\tilde{t}-\tilde{t}\tilde{e})\left(4\tilde{t}\tilde{e}^3 + 33\tilde{t}\tilde{e}^2 + 89\tilde{t}\tilde{e} + 5\tilde{e} + 78\tilde{t} + 14\right)}{(3+\tilde{e})^2(5+2\tilde{e})^2}
\end{aligned}
$$

$$
\begin{aligned}
\Phi_{IV} &= \Pi_{m,r_1} - \Pi_{a,r_1} - \Pi_{b,r_1} \\
&= \frac{(a-d)^2(1-3\tilde{t}-\tilde{t}\tilde{e})\left(4\tilde{t}\tilde{e}^3 + 17\tilde{t}\tilde{e}^2 + 19\tilde{t}\tilde{e} - 3\tilde{e} + 6\tilde{t} - 2\right)}{(4+7\tilde{e}+2\tilde{e}^2)^2}
\end{aligned}
$$

Note that the sign of each Φ-function corresponds to the sign of the polynomial in $\tilde{t}$ and $\tilde{e}$ in the last bracket in the numerator (since $1-3\tilde{t}-\tilde{t}\tilde{e} \geq 0$ in space $\mathcal{P}$ and $1-3\tilde{t}-\tilde{t}\tilde{e} = 0$ only along the boundary $\tilde{t} = \frac{1}{3+\tilde{e}}$).

B. Appendix to Chapter 2

B.1 Data Appendix

This appendix comments on the sources of data and how I treat the data. I also perform robustness checks on my direct computations of marginal cost.

Anonymous acknowledgement I wish to express my gratitude to all the people related to the cement industry whom I have interviewed during the course of this project. This project would not have been possible without their help, particularly in regard to the data collection and validation effort. I do not name them in order to preserve their confidentiality but hereby acknowledge them by citing their professional relationship with the cement supply chain: representatives for various state-level construction sector trade associations (SINDUSCONs); representatives for the cement industry's trade association (SNIC); representatives for the technical arm of the cement industry's trade association (ABCP); sales representatives, engineers and executives of cement producers; representatives of cement buyers (resellers, ready-mixed concrete firms, construction firms and producers of construction aggregates); representatives for equipment suppliers to the cement industry; representatives for factor suppliers to the cement industry; Confederation of National Industry (CNI); Brazilian Institute for Geography and Statistics (IGBE); officials of regulatory agencies; officials of government ministries; investment bank analysts; international traders in cement; academics.

B.1.1 Sources and treatment of data

Cement consumption by state (i.e. demand by local market) Monthly series by state, in 1000 tonnes of cement, are obtained from the annual reports (and other reports) of the Brazilian cement industry's trade association, the National Syndicate of the Cement Industry (SNIC). This body has played a leading role in the history of the Brazilian cement industry and represents almost the entirety of the set of producers[1]. To compile consumption figures for a given state, SNIC aggregates reported shipments by its members to that state. Thus I observe shipments by cement producers to buyers broken out by destination state. Four possible sources of distortion, each deemed to be small, are: (i) Consumption figures do not include shipments by non-members to the association (namely Mizú and Davi: see footnote 1). The distortion is small given the limited capacity, limited geographic scope and recent entry dates attached to these non-members; (ii) Consumption figures do not include imports. Again the distortion is small in view of the limited penetration of imports (see Section 2.3.3); (iii) Consumption figures do not account for any cross-state shipping *at the reseller level* (i.e. shipments by resellers in state l shipping across to buyers further downstream located in state $n \neq l$). In compiling consumption by state, a shipment by a cement producer to a buyer located in a given state counts towards consumption in that state. This distortion is considered small in that the high cost of transporting cement and the fact that the industry takes into account the possibility of trade arbitrage when setting commercial terms make the scope for cross-state shipping by resellers limited. Further, the bulkiness, fast turn and short shelf life of cement leads producers to reach far "down the trade", via direct-from-plant deliveries and own distribution terminals: in spatial terms, reselling is largely a local business. In any case, shipments by resellers into a state ought to approximately cancel shipments by resellers out of that state; and (iv) Variation in inventories downstream are not accounted

[1]Up until 2003 only two recent entrants were not members of the trade association: (i) Cimento Mizú, set up in 1998 by a large independent ready-mixed concrete firm, Polimix, and (ii) Cimento Davi, set up in 2001. Both concerns consist of relatively small-scale grinding operations (respectively 0.7 mtpa and 0.4 mtpa to date), importing clinker from as far as Asia and producing slag cement (both are located in close proximity to steel producers, from whom they purchase blast furnace slag, based respectively in the states of Espírito Santo, *ES*, and Minas Gerais, *MG*.) See footnote 16 of Chapter 3.

for. Again the distortion is small given that the characteristics of cement (e.g. short shelf life) means inventory levels and their time-series variation is limited.

Flow of cement from plants to states Annual shipments of cement from each plant to each state is obtained from SNIC, from the same database from which the monthly consumption series by state are extracted. Thus I observe, for each year T, an $I \times L$ shipment matrix with element q_{ilT} denoting the shipments from plant i to state (local market) l in year T. To obtain the flow of cement from plants to states on a monthly basis, I assume that the distribution of shipments to market l across sourcing plants is invariant over the 12 months in each year. Thus I take plant i's shipments to market l in month $t \in T$ to be $q_{ilt} = \frac{q_{ilT}}{\sum_i q_{ilT}} q_{lt}$ where q_{lt} denotes the consumption in market (state) l in month t, as detailed above.

Cement prices by state Current retail cement prices in units of local currency for the standard 50 kg bag are provided by the Brazilian office of national statistics, the Brazilian Institute for Geography and Statistics (IGBE). This office is one of Brazil's two main providers of economic statistics, charged with carrying out population censuses, compiling the national accounts and publishing price indices. In effect, the cement price series I use is collected to compute the latter. Monthly series by state are available on the median price for a sample of retail stores (commonly referred to as resellers) located in each state. (Producer prices are not observed; these are backed out from retail prices as explained below.) Owing to the high levels of inflation prevailing in the first *one-quarter* of the time period I consider[2], particular attention has been paid to the conversion of current cement prices to constant prices (in December 1999 terms). While this is done using a General Price Index (GPI), I also convert cement prices using other (economy-wide) price indices, such as a Consumer Price Index (CPI) or a Wholesale Price Index (WPI), to check the robustness of the estimation results (these are, respectively, the "IGP-DI", the "IPC-br" and the "IPA-DI", all published by the Fundação Getúlio Vargas). Further, where possible, I compare the constant price series I calculate for each state

[2]That is, the pre-stabilisation phase, or the 42 monthly observations between January 1991 and June 1994, out of a total of 156 observations (up to December 2003).

with reports on cement prices to be found in trade publications or the press. For example, the constant cement price series I calculate indicate a sharp increase in real terms in 1992; this is confirmed by aggregate real cement price indices and accounts published in trade reports at the time. One must also point out that despite the high level of inflation in the first half of the 1990s, the economic environment was far from chaotic; economic agents had learned how to cope with a chronic and fast-changing price level, and to anticipate it reasonably well in the short term. It is thus possible for the researcher to filter (upward and downward) variation in real prices from the much larger (upward) variation in nominal prices in the pre-stabilisation phase of the time period I cover.

Exogenous demand variables Several alternative series of economic activity, either in the construction and building sector or aggregated across sectors of the economy, are available as proxies for the exogenous demand for cement. The favoured series, issued by the Brazilian office of statistics (IBGE), reports the real index of activity in the construction sector for each of the 27 states, on an annual basis. Importantly, this series follows from a *volume* decomposition of Value Added in the construction sector (from the National Accounts) and should thus be a good proxy for exogenous demand. I blow up the index series for each state using the relative size of the construction sector between states, also obtained from the National Accounts; these can then be compared cross-sectionally (i.e. across states). Alternative quarterly series are available, which I use in checking the robustness of my estimation results.

Taking states to represent local markets In terms of geography, the most disaggregated level at which demand-side data are observed is the state level. Data availability thus require that I take each state to represent a market (i.e. I do not observe data at the "SMSA" or city level)[3]. However, I believe that this is sound, as

[3]Notice that this is done only for the purpose of estimating demand. The availability of plant-to-state shipment data allows me to circumvent the tricky task of defining (geographic) markets in my treatment of supply: for any given plant, the geographic area which is relevant to its supply decision (i.e. that plant's "market") is endogenous – depending on demand, cost and conduct – and these plant-specific areas overlap across plants.

follows. As the econometrician, I observe the flow of cement from each plant directly to buyers (i.e. to resellers and to consumers) in any given state. As explained above, the scope for cross-state shipping by resellers is limited. Thus price and consumption observations – obtained from different sources, as explained above – can reasonably be paired. A potential distortion arises from the possibility that the larger states may contain more than one local market with heterogeneous demand conditions, and that therefore identifying states with markets may hide important variation at a more local level where agents interact. To my defence, however, I discard the largest northern states, albeit for another reason as I explain in Section 2.4.2. Further, in the data I observe that cement *can* travel over significantly large distances from plants located in a state to buyers located in another state. Several field interviews and price data obtained directly from a sample of producers suggest that the spatial variation of prices within a state is minimal: prices are mostly uniform within a state. Of note, several studies in the literature have taken US states (of sizes similar to and as diverse as their Brazilian counterparts) to represent markets, such as Sutton (1998), FTC (1966) and Newmark (1998).

Factor prices Factor prices are either observed in the form of current prices, in which case they are converted to constant prices as explained above, or already reported in the form of constant prices. Though alternative series proxying each factor price are available – which I use as alternative instruments in the demand estimation or to check the robustness of the supply-side estimation – the main series are:

- Fuel oil: country-wide delivered prices from refineries in units of local currency per kg (excluding sales taxes) are obtained, on a monthly basis, from the oil industry regulator, the National Agency for Oil. I add sales taxes to these prices according to legislation. (Owing to policy in the oil sector, price variation across regions during the time period of the study, has been minimal.)

- Diesel oil: country-wide delivered prices from refineries in units of local currency per litre are obtained, on a monthly basis, from the oil industry regulator, the National Agency for Oil.

- Coal: FOB prices of coal in local currency units per tonne are obtained, on an annual basis, averaged across mining firms, from the Ministry for Mining and Energy. Price lists are also obtained from a sample of mining firms. Of note, coal mines are located in the South of the country; freight to cement plants employing coal as kiln fuel (largely located in the South) is added accordingly (see comments on freight cost below).

- Electricity: state-level delivered prices to (high-voltage) industrial consumers in local currency units per MWh are obtained, on a monthly basis, from the electricity industry regulator, the National Agency for Electrical Energy.

- Labour: manufacturing-industry real wage indices in the 12 states with the largest industrial output, in addition to a country-wide index, are obtained, on a monthly basis, from the Confederation of National Industry.

Plant characteristics Plant characteristics such as ownership, capacity (i.e. kiln pyroprocessing capacity and grinding capacity), number of kilns, age, technology (i.e. the type of equipment and process, whether dry or wet, whether a preheater is employed) and the fuel mix employed by kilns (largely either fuel oil, coal, or more recently pet coke or natural gas) is available from the Brazilian trade association[4] and from different editions of the World Cement Directory, published by the European Cement Association (Cembureau) every three years, compiling information on cement producers across the world. Data is complemented by or confirmed against information from (i) industry publications, (ii) investment banking reports, (iii) the press, (iv) companies' websites, (v) academic publications, and/or (vi) field interviews (see below). Of note, capacity and technology data may contain significant measurement error. The shortest distance by road from

[4]Plenty of other information is available, such as stock levels by plant, or the form shipments from each plant take, in terms of packaging (in bags or in bulk) or in terms of the means of transportation (by road, rail or water). Aggregating across all plants, between 1997 and 1999 81% of shipments were in bags, and 91% of shipments were by road. The breakdown of shipments among different buyer channels is also available, with resellers accounting for 76% and ready-mixed concrete firms accounting for 11%, in this same period.

each plant to the main metropolitan areas in each state is available from the Ministry of Transport.

Computing plant marginal cost: an upper bound Using the fixed-coefficient nature of cement production technology, I can directly calculate marginal cost from observed factor prices, the observed plant characteristics and engineering estimates of the fixed coefficients. I employ the term "calculate" rather than "estimate" since obtaining marginal cost does not involve statistical inference; however, calculated marginal costs are indeed estimates – in fact they are estimated upper bounds to the true marginal costs – in the sense that there inevitably are unobserved plant characteristics, as I explain below. In view of the fixed-coefficient technology and my understanding of the industry, I model plant marginal cost as flat in quantity up to capacity. To the extent that marginal cost varies across kilns within the same plant complex, this will be an approximation to the true plant marginal cost which would then be a step function in quantity. (For example, if a plant consists of two kilns, labelled 1 and 2 in order of most efficient first, with kiln marginal costs denoted c_1 and $c_2 > c_1$, denoting other marginal cost by c, and denoting kiln capacities by K_1 and K_2, then marginal cost would be $MC(q) = c + c_1$ if $q \leq K_1$ or $c + c_2$ if $K_1 \leq q \leq K_2$.) Clearly, this will be of relevance only if plant capacity utilisation varies sufficiently over time that the *marginal kiln* in operation differs (e.g. a less efficient kiln is fired up and shut down for months at a time according to demand). I thus mitigate any distortion stemming from my approximation of plant marginal cost as being flat in quantity by taking capacity utilisation (see below) and the characteristics of the marginal kiln into account when computing marginal cost. (It should also be noted that in recent decades the trend for cement plants has been to favour large single-kiln production lines as against multiple small lines, in view of the economies of scale.) It is also worth clarifying the way a kiln works. A kiln, when in operation, must run at close to full capacity; it cannot be operated at any given moment at, say, 50% capacity. Further, firing up a kiln is costly so when in operation a kiln typically runs for at least several days or weeks. As for plant marginal costs, it is clear that these will vary across plants according to the technology, capacity and age of the equipment and the fuel employed by the kiln; these are accounted for to the best of my knowledge as I explain below[5]. Plant marginal costs fall into four main categories – kiln fuel, electricity, mineral extraction royalties, and labour/packaging/other costs – as follows:

1. Kiln fuel: This is the main component of plant marginal cost. Based on engineering estimates, the heat content required to produce 1 kg of clinker using the dry process (see Section 2.3.1) will typically fall in the 650 - 850 kcal range (e.g. see World Cement, January 2000 issue). (The wet process consumes over double this.) A kiln's (thermal) energy efficiency will depend on the capacity, technology (including the specifications of preheating, cooling and waste heat recovery systems used) and age of the kiln. (The kiln's brick lining has to be changed periodically, and the time since the last relining will also impact the energy efficiency of the kiln.) Interview-based evidence, however, indicates that the energy efficiency of kilns in operation in the Brazilian cement industry is (i) high relative to its global peers (including the US industry), with producers having shifted to the dry process chiefly over the 1980s, (ii) has continued to improve over the 1990s, and (iii) presents low variation across producers (with perhaps two exceptions, both with lower productivity). Based on observed plant characteristics and interviews, I classify the energy efficiency of each plant as "above average", "average" and "below average", assuming energy contents of 690, 730 and 800 kcal/kg of clinker respectively. (For example, with respect to kiln capacity, a kiln with capacity in excess of 1 mtpa will require a heat content approximately 6% below that of a kiln with capacity of 0.25 mtpa, controlling for other characteristics.)[6] To arrive at the marginal cost relating to kiln fuel, I consider two types of fuel – fuel oil and coal – for which I observe prices, as explained above. I then use the observed fuel mix for each plant and the properties of each fuel to obtain fuel cost. (For example, an average-efficiency kiln burning a certain

[5]To this end I have met with engineers working in cement plants or working for the technical arm of the cement producers' trade association (the Brazilian Association for Portland Cement, ABCP), as well as meeting executives of equipment suppliers to the cement industry.

[6]In addition to the above, I use other sources of information such as a report compiled by the industry's trade association in 1993 on plant productivity with respect to energy inputs (stating the amount of hydrocarbon equivalent burned by tonne of clinker produced by plant). Note further that I neglect time-series variation in plant-specific energy efficiency given that this has been low over the time period of the study.

grade of fuel oil – with "inferior calorific power" of 9750 kcal/kg of fuel oil – will require $730/9750 \times 1000 = 75$ kg of fuel oil per tonne of clinker.)

2. Electricity: While thermal energy is required to produce clinker in the kiln, electricity is used mostly for grinding raw material, solid kiln fuel (such as coal) and clinker – a process known as comminution – and to a lesser extent to operate conveyor belts and packaging lines. Considerations here are similar to those made for kiln fuel. For example, in terms of technology, the more modern vertical roller mills tend to consume less power than the ball mill system. Again based on engineering estimates, the total plant electricity content required to produce 1 tonne of cement typically falls in the region of 90 - 105 kWh.

3. Mineral extraction royalties: The marginal cost component arising from the extraction of raw material (limestone and clay), from a quarry usually located within the plant complex, follows from legislation. The "Financial Compensation for the Extraction of Mineral Resources" (CFEM) requires that the cement producer collect 2% of its revenues from the sale of cement, net of sales taxes and freight, in the way of compensation to the government (see below for producer prices and sales taxes). Exceptions to this requirement, where negotiated between producers and the government, are not observed (see comment below on unobservables).

4. Labour/packaging/other costs: As mentioned in Section 2.3.1, labour essentially performs a supervisory role. One may argue that a certain proportion of a plant's labour cost is fixed. The variable proportion of labour would correspond to quarrying personnel and possibly workers involved with the packaging and distribution centre operations. I refrain from discussion, given the relatively low cost of these plant operations and the fact any bias in the direction of overstating marginal cost reinforces the results of this study. Packaging costs will vary according to the proportion of a plant's production that is shipped in bags (largely in the form of the standard 50 kg bag) as opposed to bulk shipments; recall that I observe this proportion. As such, based on information at hand, I take this component of marginal cost to

amount to around 5% of net producer price.

A final comment relates to unobservables. Despite a researcher calculating marginal cost to the best of his ability from observables, there will always be an unobserved (to the researcher) component to productivity across plants. In view of this, in computing marginal cost I choose to "err on the side of conservatism", in this case by overstating marginal cost. I claim the calculated marginal cost is thus an upper bound to the true marginal cost, thus leading to an understated price-cost margin. The reason behind this choice is that, when I turn to the testing of conduct in Section 2.4.4, such a bias reinforces the results of this study. There are several potential sources of bias (in the direction of overstating marginal cost). One source of unobserved plant heterogeneity may be the quantity discount enjoyed by Votorantim, the largest producer on a nationwide basis, in acquiring fuel oil, electricity or trucking services vis-à-vis the smaller producers. To the extent that producers manage to acquire factors on different terms (controlling for location), the marginal cost I compute may be overstated for the firms with greater bargaining power over suppliers. Another possibility is that producers substitute away from traditional kiln fuels such as fuel oil and coal in favour of cheaper alternatives[7]. Yet another bias stems from the formulation I use to compute the marginal cost of cement (see footnote 28 of Chapter 2). For simplicity I take 1 kg of cement to correspond to 1 kg of clinker. Now clinker is the most expensive input to the grinding process and even "pure" cement (referred to commonly as type I cement, or simply ordinary cement) is comprised of 96% clinker and 4% gypsum by weight. To the extent that different formulations of cement are produced, with a lower proportion of clinker (and a higher content of lower-cost additives, such as slag, pozzolane and/or filler), the bias in the direction of overstating marginal cost will be higher. For example, composite (type II) cement, with a clinker content in the region of 70 - 80%, accounted for 78% of the Brazilian industry's total production

[7]Indeed, since 2000 the use of pet coke (imported by some producers themselves from the Mexican Gulf) and natural gas (to the extent that a plant is located in proximity to a pipeline) is on the rise. A clinker kiln will in principle burn any material with a sufficiently high energy content, such as used rubber tyres, solvents and hazardous waste materials. The equipment supplier FLSmidth speculates that in the long term the cost of kiln fuel could fall to zero, or even turn negative, with cement producers being paid to dispose of waste materials.

between 1999 and 2001. Finally, note that *fixed cost* heterogeneities across plants, stemming for example from decreasing unit capital costs as a function of capacity (i.e. economies of scale), are not relevant for the computation of marginal cost and the resulting price-cost margins.

Computing plant-to-market freight cost This is the first component to ex-plant marginal cost that I consider (the other two are the reseller mark-up and producers' sales taxes). In the cement industry, as seen in Section 2.3.1, freight is a large component of cost. The vast majority of shipments from producers to buyers take place by road and are provided for by the producers. I do not observe the exact freight rates paid by cement producers. But fortunately I do observe a good proxy for the freight of cement. The transportation of agricultural goods such as soyabean and maize is reportedly a close substitute to the supply of cement freight, in view of product and market characteristics (Soares and Caixeta Filho 1996)[8]. I use a database containing approximately 30,000 observations on freight prices for some agricultural goods collected over the period 1997 to 2003 for thousands of different routes across Brazil[9]. Figure 5-1 summarises the results of some auxiliary reduced-form regressions. These should be seen as hedonic regressions with the purpose of predicting the price of freight. Given that I do not observe quantities demanded and supplied in the market for freight, I cannot estimate a structural model of the market for freight. (Nor do I think this is necessary in view of my objective, which is to predict the freight cost of cement from plant i to local market l based on observed data.) Freight prices (once converted to constant prices as explained

[8]For example, an interview with a cement industry executive revealed that during the soyabean harvesting season (March through May) the producer he works for encourages large resellers to themselves pick orders up at the plant, for fear of relying too heavily on the scarce supply of outside truckers observed during these months. This further suggests that freight of cement and freight of soyabean are close substitutes, and therefore that their prices should be similar. Most cement producers outsource trucking services, mostly to independent truckers who are registered in their databases and simply turn up at the door and are hired on the spot (or are hired through cooperatives or middlemen). According to this executive, the cement industry is the top industrial contractor of trucking services in the country.

[9]I am indebted to Professor José Vicente Caixeta Filho of ESALQ, at the University of São Paulo, for providing an extract of the SIFRECA freight database. Data pertaining to soyabean, maize and (the mineral) limestone was kindly made available.

earlier[10]) are regressed on exogenous variables such as the distance of the route, the squared distance, a shipment-to-port dummy (to capture exports), transportation-mode dummies (by water or by rail, as opposed to by road), seasonal dummies or monthly dummies (to capture the harvesting cycle), the price of diesel oil (the main cost component for freight), a packaging dummy (shipment of bagged produce as opposed to bulk) and product-type dummies (e.g. powdered soyabean), in addition to interaction variables. It is clear from the R^2 of the OLS regressions that the fit is very high; the heteroskedasticity-robust standard errors are low. Freight prices (in R$ per tonne) are increasing in distance (and concave, though slightly so over the relevant range). Consider the results for specification (II). At the sample means of the variables (735 km for distance and R$ 0.422 per litre for the price of diesel oil), the predicted price of freight for a tonne of soyabean shipped in bulk by road to a destination other than a port and in the month of April amounts to $3.358 + 0.0405 \times 735 - 5.44 \times 10^{-7} \times 735^2 + 6.519 \times 0.422 = 3.358 + 29.768 - 0.294 + 2.751 =$ R$ 35.56 (with a standard error of R$ 0.21). Shipping to a port (possibly as a result of longer waiting times to unload) adds $1.813 + 0.00041 \times 735 =$ R$ 2.12 (s.e. R$ 0.15), and when this shipping to a port takes place during the harvest season freight prices are predicted to increase by a further R$ 2.30 (s.e. R$ 0.32). Shipping by waterway costs $14.269 + 0.00498 \times 735 =$ R$ 17.93 (s.e. R$ 0.25) less than by road, while shipping by railway costs $2.349 + 0.01538 \times 735 =$ R$ 13.66 (s.e. R$ 0.28) less than by road. Shipping in bags as opposed to in bulk raises the price of freight by R$ 0.25 though this estimate is not significantly different from zero. Compared to April, the peak month of the harvesting season, shipments in any other month of the year are cheaper (all coefficients on monthly dummies and their interactions with distance are negative). Shipments in January, the month in which prices are lowest, are R$ 4.87 (s.e. R$ 0.25) lower compared to April. Note that the variation in diesel oil prices over the period is R$ 0.38, accounting thus for a R$ 2.49 (s.e. 0.17) variation in freight prices (this is admittedly low, owing possibly to correlation between diesel oil prices and other variables). I choose to predict the plant-to-market freight cost for cement based on specification (II), on account of

[10] Freight prices, in units of local currency per tonne of produce shipped, are thus in December 1999 terms.

observables such as distance from the plant to the market, means of transport and the price of diesel oil.

	(I)		(II)		(III)	
	coef	s.e.	coef	s.e.	coef	s.e.
No. obs.	27974		27974		30367	
R^2	0.894		0.899		0.904	
Intercept	1.423 ***	(0.244)	3.358 ***	(0.361)	5.413 ***	(0.447)
Distance of route	0.0387 ***	(0.0005)	0.0405 ***	(0.0007)	0.0433 ***	(0.0008)
Distance of route squared	-8.12E-07 ***	(2.49E-07)	-5.44E-07 **	(2.56E-07)	-9.62E-07 ***	(2.38E-07)
Port destination dummy	2.135 ***	(0.166)	1.813 ***	(0.267)	1.720 ***	(0.238)
Water transport dummy	-17.405 ***	(0.212)	-14.269 ***	(1.094)	-11.516 ***	(1.246)
Rail transport dummy	-12.410 ***	(0.343)	-2.349 ***	(0.571)	-3.149 ***	(0.540)
Harvest season dummy	2.341 ***	(0.118)				
Port during harvest dummy	2.802 ***	(0.311)	2.295 ***	(0.318)	2.248 ***	(0.277)
Price of diesel oil	6.815 ***	(0.441)	6.519 ***	(0.443)		
Shipment in bags dummy			0.249	(0.204)	0.489 **	(0.201)
Powdered soya dummy			1.510 ***	(0.134)	1.749 ***	(0.127)
Maize dummy			-0.755 ***	(0.096)	-0.976 ***	(0.097)
Limestone dummy			-2.136 ***	(0.151)	-1.819 ***	(0.140)
Monthly dummies			Included (except April)		Included (except April)	
Year dummies					Included (except 1997)	
Distance interacted with:						
Port dummy			0.00041	(0.00031)	0.00062 **	(0.00028)
Water transport dummy			-0.00498 ***	(0.00164)	-0.00859 ***	(0.00184)
Rail transport dummy			-0.01538 ***	(0.00088)	-0.01317 ***	(0.00084)
Monthly dummies			Included (except April)		Included (except April)	
Year dummies					Included (except 1997)	

Note: Heteroskedasticity-robust standard errors

*** Significant(ly different from zero) at the 1% level; ** Significant at the 5% level; * Significant at the 10% level

Dependent variable is Freight Price in units of local currency (at December 1999 prices) per tonne of produce shipped

Figure 5-1: Auxiliary OLS regressions for plant-to-market freight cost

Backing out net producer prices from retail prices The other components to ex-plant marginal cost, apart from plant-to-consumer freight, are the reseller (retailer) mark-up and producers' sales taxes. Recall that the lion's share of the Brazilian cement business consists of producers shipping bagged cement to resellers, who then sell directly to the end user (a small-scale consumer); I only observe the prices set by these resellers, not the prices set by producers. However, I back out producer prices as follows. Based on several field interviews[11], I model the reseller

[11] These interviews include cement producers' field representatives and sales executives, buyers of cement and representatives of the construction sector's trade associations across a sample of local markets. Information provided in these interviews was also consistent with a report on the

as competitive. I thus avoid the issue of double marginalisation. A reseller's cost consists largely of (i) two forms of sales tax ("PIS" and "COFINS", not to be confused with the sales taxes collected by the producer), which are proportional to the retail price (varying from 2.65 to 3.65% over the time period), and (ii) labour costs (for unloading the truck, storage handling and stocking shelves). While sales tax will be perfectly correlated with cement prices, this is not the case for labour costs. In any case, based again on the field interviews, I assume that resellers apply a fixed proportional mark-up over the producer price (namely in the region of 13%) and can then back out producer prices from observed retail prices. To the extent that (i) labour costs vary across markets, or (ii) some resellers evade taxes, or (iii) some resellers occasionally choose to price cement as a "loss leader" to lure consumers into their stores, the reseller mark-up may vary across resellers. This variation, however, should be small (and the bias, again, is in the direction of overstating marginal cost). Further, price discrimination by producers on the basis of customer size (i.e. quantity discounts) is very limited across resellers[12]. A 5-10% discount may be offered to large buyers who buy in bulk (ready-mix concrete firms and large construction firms) yet again this corresponds to a small share of the business compared to that flowing through resellers. Among the robustness checks I perform, I compare *observed* producer prices that I was fortunate to obtain from a subset of producers to the backed-out producer prices. Finally, to calculate producers' sales taxes and thus arrive at net producer prices, I consider federal and state-level tax legislation. Despite the awkwardness of Brazil's sales tax system (one needs to compute five different sales taxes, namely "ICMS-normal", "ICMS-ST", "IPI", in addition to the producer's own collection of "PIS" and "COFINS"; note further that tax rates vary according to the origin and destination of the shipment) the total sales tax collected by a producer upon selling to a buyer is conveniently proportional to the (net or gross) price set. Despite sales taxes on cement (and on

supply chain prepared by a consulting firm for the cement industry trade association (Booz Allen & Hamilton 1990).

[12]As mentioned earlier, in view of the bulkiness, fast turn and short shelf life of cement, producers reach far "down the trade", via direct-from-plant deliveries and own distribution terminals. They resort to distributors for a minor share of their business. Even a relatively small retailer (reseller) will be able to place an order directly with the producer; recall that a 15 ton (25 ton) truckload corresponds to only 300 (500) bags.

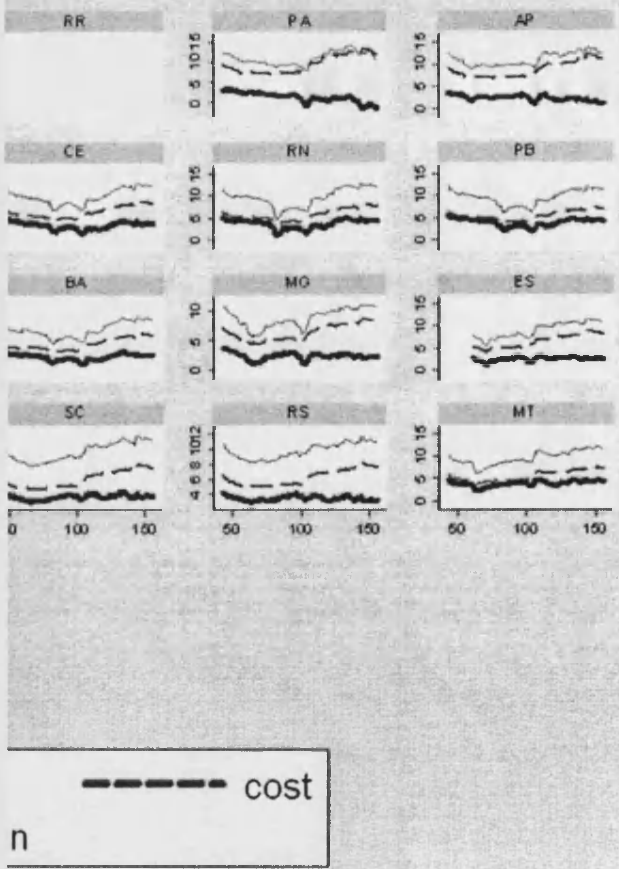
other products in general) being high (e.g. towards the end of the time period of study, sales taxes owed by a producer located in the state of São Paulo selling to a buyer located in the same state amounted to 28% of the gross producer price), sales tax evasion on the part of producers is considered to be minimal. However, to the extent that producers manage to negotiate reductions in their tax liabilities with state governments eager to attract investments – negotiations which I do not observe – the marginal cost I calculate will again be overstated (see the earlier paragraph on unobservables).

A glance at margins Figure 5-2 displays the by-market evolution of consumer prices, marginal costs and price-cost margins since July 1994 for nationwide leading firm Votorantim (Figure 5-2 is the by-state counterpart to Figure 2-11 in Section 2.3.3, where these series are aggregated across the 25 states where Votorantim is present).

Capacity utilisation As mentioned above, among other plant characteristics I observe capacity. As is usually the case with capacity figures, these are admittedly prone to considerable measurement error. Nevertheless, I proceed to analysing the evolution of plant capacity utilisation, including the three years post stabilisation of steep consumption growth (1995 to 1997). Throughout the time period, capacity outstrips production (*including* the growth years, although the slack is lower[13]), with capacity utilisation hovering around an average 65%. As discussed in Section 3.2, capacity seems to play a strategic role in the industry. Of particular interest, capacity utilisation appears to be similar across firms and plants, with firms' capacity utilisation rates being correlated over time, as firms' market shares are fairly stable, with some exceptions. This finding further supports my modelling of marginal cost as being flat in quantity (up to capacity). To the extent that

[13]Note therefore that the post-stabilisation boom in demand did not catch the cement industry unprepared in terms of capacity. Other industries with tighter capacity facing the same boom in demand saw either entry (such as imports) or an increase in prices, which clearly was not the experience of the cement industry (recall that imports were largely kept at bay and, not unrelatedly as I argue in this thesis, prices fell).

argin on Votorantim's sales by state



s producer are less than Rk 10

as capacity utilisation rises, older, smaller and energy-inefficient kilns may be put back into use, marginal costs may rise as capacity becomes tight.

B.1.2 Robustness checks on direct measures of marginal cost

In addition to checking my calculation of the reseller's mark-up on a subset of the data where I *do* observe producer prices – thus enabling these to be compared to the producer prices I back out from observed consumer prices, as explained above – I perform two other robustness checks of the constructed marginal costs. The first check is centred on the Portugal-based multinational Cimpor, which in 1997 bought its way into Brazil and in 1999 became the third largest firm in the country upon acquiring Brennand. This firm is of particular interest in that it is listed on the Lisbon stock exchange and chooses to report its financial results broken out by country of operation (and line of business). I can thus use their reported results for Brazil as a robustness check for my calculated price-cost margins. The evolution of the calculated price-cost margin, as a percentage of net producer sales (i.e. net of sales taxes), is graphed in Figure 5-3. This evolution is compared to Cimpor's reported EBITDA (earnings before income tax and depreciation allowance, also known as operating cash flow) as a percentage of net sales, over the period 1998 to 2003. (I can further check my estimates of resellers' mark-ups and sales taxes by comparing my calculated net producer sales, backed out from observed consumer prices, to their reported net sales.) The time series fit between constructed and reported figures is good. For example, I estimate Cimpor's average price-cost margins as a percentage of net producer sales rising from around 47% in 2000 to 56% in 2002. Cimpor reports a similar rise in this period, from 44% to 55%.[14] If anything, my calculated price-cost margins are slightly higher than the EBITDA figures Cimpor reports. This is to be expected, for while my cost estimates include only (constant) marginal cost, Cimpor's EBITDA figures are net of other costs such as plant overhead and sales and administrative expenses. Indeed, my

[14]Comparing operating cash flow (EBITDA) margins across the 9 countries (in Iberia, Africa and South America) where Cimpor is active, Brazil's cement operations are the most profitable: a 55.5% EBITDA margin in Brazil compared to an average 39.2% across all countries.

price-cost margins appear to be conservative (on the low side), as expected from the discussion above regarding the directional bias in the construction of marginal cost owing to unobservables (in the way of overstating marginal cost).

Figure 5-3: Evolution of the average price-cost margin for firm Cimpor. In constant Reais x 1000 per month (at December 1999 values) and as a percentage of net producer sales. Aggregated across all states.

The second additional robustness check is based on accounting data of the cement industry surveyed annually by the Brazilian Institute for Geography and Statistics (IBGE) as part of their Annual Industry Survey (PIA) series. Figure 5-4 depicts the average accounting gross margin (defined as producers' Net Sales minus Cost of Goods Sold) as a percentage of net sales for a sample of establishments over the 1990s; the number of establishments varies between 33 and 55 and only aggregate data is published. The accounting gross margin is high, hovering around 50%. Note that the accounting definition of Cost of Goods Sold does not include freight expenses but does includes accounting depreciation, so the accounting gross margin cannot be immediately compared to my constructed price-cost margin (which does consider freight but not depreciation). Further, I do not know the identity of the surveyed establishments. However, the magnitude of both series appears to be consistent. Of perhaps greater importance, the variation in the surveyed accounting gross margin is consistent with the observed fall in prices beginning in 1992 and the rise in prices commencing in 1997. (Notice the capital-intensive nature of the industry: on average payroll – corresponding not only to plant but also to sales and administrative employees – accounts for less than 10% of a producer's net sales.)

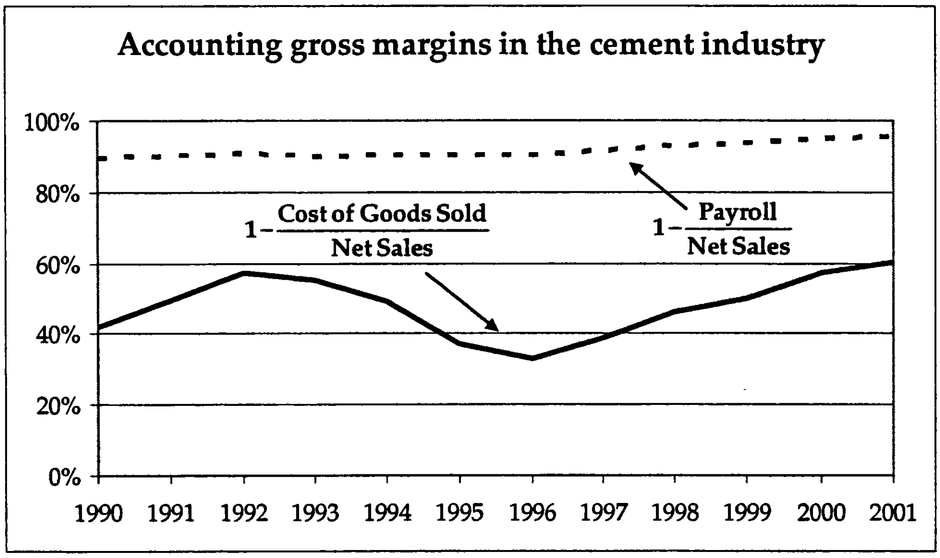


Figure 5-4: Accounting gross margins in the cement industry, from surveys conducted by the Brazilian Institute for Geography and Statistics

C. Appendix to Chapter 3

C.1 Characteristics of the cement industry and their relation to the literature on collusion

The table below, based on Ivaldi et al (2003), lists industry characteristics that are understood to facilitate collusion, summarising the extent to which they characterise the cement industry[1]. The purpose of this section is to show that tacit collusion in the Brazilian cement industry, orchestrated for instance via market division, is a concrete possibility in that the characteristics of the industry are consistent with those characteristics that the literature suggests make tacit collusion more likely.

[1] A discussion of each characteristic and how it relates to the cement industry is beyond the present scope. The purpose of the table is to convey a simple message with which most academics and analysts would agree, stated loosely as follows: cement is an archetypal example of an industry whose characteristics make it *more* collusion-prone than *less*.

Characteristics facilitating collusion	Characterise the cement industry?
Few competitors (in a given local market)	Yes
Entry barriers	Yes
Frequent interaction among firms	Yes
Market transparency (1)	Yes
Growing demand	Yes in some local markets
Mild business cycles (intensity and length)	Yes in some local markets
Low innovation	Yes
Homogeneous product (low differentiation)	Yes
Low cost asymmetry (2)	Yes in some local markets
Low capacity asymmetry	No
Evenly-distributed idle capacity	Yes in some local markets
Multimarket contact (3)	Yes
Low price elasticity of demand in equilibrium (4)	Yes
Absence of countervailing buyer power	Yes
Structural links between firms (e.g. cross-ownership)	Yes in some cases
Small competitors on the fringe	Yes
Absence of network effects or learning effects	Yes

Notes to the table: (1) Market transparency. Of note is the (potential) monitoring role played by the cement producers' trade association in compiling and sharing data on firm-level quantities, prices and/or capacity utilisation. (2) Low cost asymmetry. Scherer et al (1975) report very low economies from multiplant operation in the cement industry. (3) Multimarket contact. It is worth reflecting on the changes to market structure across time. As mentioned earlier, of the 19 producers operating in Brazil in 1991, the industry had consolidated to 12 firms by 1999. Yet this came about not in the form of increased concentration at the local market level, but largely in the form of firms making acquisitions in markets where they did not previously operate. Thus despite the changes to asset ownership, the number of firms shipping to any given market has not changed significantly. It is rather multimarket contact which has increased, in the sense that the number of local markets in which any two producers meet has gone up considerably over the decade. See Bernheim and Whinston (1990) and the discussion in

Section 3.2. (4) Low price elasticity. Ivaldi et al (2003) argue that rather than enhancing the sustainability of collusion, in the sense of increasing the range of discount factors at which a collusive equilibrium can be supported, the effect of inelastic demand may be to increase the profitability of collusion by raising the optimal collusive price. A similar point may be made regarding market division, by which firms concentrate shipments to those markets where their plants are located and reduce cross-hauling to neighbouring markets (i.e. to other firms' backyards). To the extent that under competition one may observe cross-hauling (e.g. under Cournot competition), the effect of high transport costs, as is the case for cement, may then be to raise the profitability of a collusive agreement that prescribes large shares in a given local market to those firms with plants located in that market. While the effect of inelastic demand works via demand (higher collusive prices increasing the profitability from collusion), the effect of high transport cost operates via cost (the reduction of cross-hauling increasing the profitability from collusion). See Section 3.2.

Bibliography

[1] Anderson, J. E. and E. van Wincoop (2003). Gravity with gravitas: A solution to the border puzzle. *American Economic Review 93*(1), 170-192

[2] Barkoulas, J., C. Baum and A. Chakraborty (2001). Waves and persistence in merger and acquisition activity. *Economic Journal 70*, 237-243

[3] Baum, C. F., M. E. Schaffer and S. Stillman (2003). Instrumental variables and GMM: Estimation and testing. Department of Economics Working Paper No. 545, Boston College

[4] Bernheim, B. D. and M. D. Whinston (1990). Multimarket contact and collusive behavior. *Rand Journal of Economics 21*, 1-26

[5] Bittlingmayer, G. (1985). Did antitrust policy cause the Great Merger Wave? *Journal of Law and Economics 28*, 77-118

[6] Booz Allen & Hamilton (1990). A cadeia de suprimentos da indústria de cimento. *Relatório preparado para o Sindicato Nacional da Indústria do Cimento.* São Paulo, Brazil.

[7] Brander, J. A. and A. Zhang (1990). Market conduct in the airline industry: An empirical investigation. *Rand Journal of Economics 21*, 567-583

[8] Bresnahan, T. F. (1982). The oligopoly solution concept is identified. *Economics Letters 10*, 87-92

[9] Bresnahan, T. F. (1987). Competition and collusion in the American automobile industry: The 1955 price war. *Journal of Industrial Economics 35*, 457-482

[10] Bresnahan, T. F. (1989). Empirical studies of industries with market power. In R. Schmalensee and R. D. Willig (Eds.), *Handbook of Industrial Organization, Volume II*, pp. 1011-1057. New York: North-Holland.

[11] Carlsson, B. (2001). Prices, competition, and industry structure: The cement industry in the United States, Japan, and Sweden, 1980-1997. Mimeo, Case Western Reserve University

[12] Carlton, D. W. (forthcoming). Using economics to improve antitrust policy. 2003 Milton Handler Lecture. *Columbia Business Law Review*

[13] Caves, R. (1991). Corporate mergers in international economic integration. In A. Giovannini and C. Mayer (Eds.), *European Financial Integration*, pp. 136-160. Cambridge: Cambridge University Press.

[14] Cembureau (several editions). *World Cement Directory.* Brussels, Belgium: The European Cement Association.

[15] Clark, K. B. (1980). Unionization and productivity: Micro-economic evidence. *Quarterly Journal of Economics 95*(4), 613-639

[16] Corts, K. S. (1999). Conduct parameters and the measurement of market power. *Journal of Econometrics 88*, 227-250

[17] Deneckere, R. and C. Davidson (1985). Incentives to form coalitions with Bertrand competition. *Rand Journal of Economics 16*, 473-486

[18] Dixit, A. (1980). The role of investment in entry deterrence. *Economic Journal 90*, 95-106

[19] Dumez, H. and A. Jeunemaître (2000). *Understanding and Regulating the Market at a Time of Globalization: The Case of the Cement Industry.* London: Macmillan Press.

[20] Eastman, H. C. and S. Stykolt (1966). *The Tariff and Competition in Canada.* Toronto: University of Toronto Press.

[21] Evenett, S. J. and W. Keller (2002). On theories explaining the success of the gravity equation. *Journal of Political Economy* 281-316

[22] Fauli-Oller, R. (2000). Takeover waves. *Journal of Economics & Management Strategy 9*(2), 189-210

[23] Federal Trade Commission (1966). *Economic Report on Mergers and Vertical Integration in the Cement Industry.* Staff Report.

[24] Feenstra, R. C. (2002). Border effects and the gravity equation: Consistent methods for estimation. *Scottish Journal of Political Economy 49*(5), 491-506

[25] Gasmi F., J. J. Laffont and Q. H. Vuong (1990). A structural approach to empirical analysis of collusive behavior. *European Economic Review 34*, 513-523

[26] Gasmi F., J. J. Laffont and Q. H. Vuong (1992). Econometric analysis of collusive behaviour in a soft-drink market. *Journal of Economics and Management Strategy 1*, 277-311

[27] Genesove, D. and W. P. Mullin (1998). Testing static oligopoly models: conduct and cost in the sugar industry, 1890-1914. *Rand Journal of Economics 29*, 355-377

[28] Ghemawat, P. and C. Thomas (2004). Identifying the sources of sustained performance differences: A study of the international cement industry. Mimeo, Harvard Business School

[29] Gilbert, R. and X. Vives (1986). Entry deterrence and the free rider problem. *Review of Economic Studies 53*, 71-83

[30] Golbe, D. and L. White (1993). Catch a wave: The time series behaviour of mergers. *Review of Economics and Statistics 75*(3), 493-499

[31] Gollop, F. and M. Roberts (1979). Firm interdependence in oligopolistic markets. *Journal of Econometrics 10*, 313-331

[32] Graham, E. (1985). Intra-industry direct foreign investment, market structure, firm rivalry and technological performance. In A. Erdilek (Ed.), *Multinationals as Mutual Invaders: Intra-industry Direct Foreign Investment*, pp. 67-88. London: Croom Helm.

[33] Greenhut, J. G. and M. L. Greenhut (1975). Spatial price discrimination, competition and locational effects. *Economica 42*, 401-419

[34] Harris, R. (1984). Applied general equilibrium analysis of small open economies with scale economies and imperfect competition. *American Economic Review 74*, 1016-1032

[35] Head, K. (2003). Gravity for Beginners. Mimeo, Sauder School of Business, University of British Columbia

[36] International Cement Review (several editions). Dorking, UK: Tradeship Publications.

[37] Ivaldi, M., B. Jullien, P. Rey, P. Seabright, J. Tirole (2003). The economics of tacit collusion. *Final Report for DG Competition, European Commission.* IDEI, Toulouse

[38] Jans, I. and D. I. Rosenbaum (1996). Multimarket contact and pricing: Evidence from the U.S. cement industry. *International Journal of Industrial Organization 15*, 391-412

[39] Johnson, R. N. and A. Parkman (1983). Spatial monopoly, non-zero profits and entry deterrence: The case of cement. *Review of Economics and Statistics 65*(3), 431-439

[40] Kamien, M. and I. Zang (1990). The limits of monopolization through acquisition. *Quarterly Journal of Economics 105*, 465-499

[41] Kamien, M. and I. Zang (1993). Monopolization by sequential acquisition. *Journal of Law, Economics & Organization 9*, 205-229

[42] Knickerbocker, F. (1973). *Oligopolistic Reaction and Multinational Enterprise.* Boston: Harvard University.

[43] Lau, L. J. (1982). On identifying the degree of competitiveness from industry price and output data. *Economics Letters 10*, 93-99

[44] Levy, D. and J. Reitzes (1992). Anticompetitive effects of mergers in markets with localised competition. *Journal of Law, Economics & Organization 8*, 427-440

[45] Loungani, P., A. Mody and A. Razin (2002). The global disconnect: The role of transactional distance and scale economies in gravity equations. *Scottish Journal of Political Economy 49*(5), 526-543

[46] Matsushima, N. (2001). Horizontal mergers and merger waves in a location model. *Australian Economic Papers 40*(3), 263-286

[47] McBride, M. E. (1983). Spatial competition and vertical integration: Cement and concrete revisited. *American Economic Review 73*, 1011-1022

[48] Motta, M. (1992). Sunk costs and trade liberalisation. *Economic Journal 102*(412), 578-587

[49] Motta, M. (1994). International trade and investments in a vertically differentiated industry. *International Journal of Industrial Organization 12*, 179-196

[50] Motta, M. and H. Vasconcelos (2003). Efficiency gains and myopic Antitrust Authority in a dynamic merger game. CEPR Discussion Paper 4175.

[51] Mueller, Dennis (1989). Mergers: Causes, Effects and Policies. *International Journal of Industrial Organization 7* (special issue on mergers), 1-10.

[52] Newmark, C. M. (1998). Price and seller concentration in cement: effective oligopoly or misspecified transportation cost? *Economics Letters 60*, 243-250

[53] Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. *Econometrica 69*, 307-342

[54] Nilssen, T. and L. Sørgard (1998). Sequential horizontal mergers. *European Economic Review 42*, 1683-1702

[55] Norman, G. (1979). Economies of scale in the cement industry. *Journal of Industrial Economics 27*, 317-37

[56] Pagan, A. R. and D. Hall (1983). Diagnostic tests as residual analysis. *Econometric Reviews 2*, 159-218

[57] Parker, P. M. and L. H. Röller (1997). Collusive conduct in duopolies: multimarket contact and cross-ownership in the mobile telephone industry. *Rand Journal of Economics 28*, 304-322

[58] Perry, M. and R. Porter (1985). Oligopoly and the incentive for horizontal merger. *American Economic Review 75*, 219-227

[59] Porter, R. H. (1983). A study of cartel stability: The Joint Executive Committee, 1880-1886. *Bell Journal of Economics 8*, 565-576

[60] Reiss, P. C. and F. Wolak (2002). Structural econometric modeling: Rationales and examples from industrial organization. *Handbook of Econometrics, Volume 6.* Manuscript of 2002.

[61] Redding, S. and A. J. Venables (2004). Economic geography and international inequality. *Journal of International Economics 62*(1), 53-82.

[62] Roberts, M. (1983). Testing oligopolistic behavior: An application of the variable profit function. *International Journal of Industrial Organization 2*, 367-383

[63] Röller, L. H. and F. Steen (2002). On the workings of a cartel: Evidence from the Norwegian cement industry from 1955-1968. Mimeo, WZB and Norwegian School of Economics

[64] Rosenbaum, D. I. (1989). The impact of market structure on technological adoption in the portland cement industry. *Quarterly Review of Economics and Business 3*, 102-110

[65] Rotemberg, J. and G. Saloner (1986). A supergame theoretic model of price wars during booms. *American Economic Review 76*, 390-407

[66] Salant, S., S. Switzer and R. Reynolds (1983). Losses from horizontal merger: The effects of an exogenous change in industry structure on Cournot-Nash equilibrium. *Quarterly Journal of Economics 98*, 185-199

[67] Scherer, F. M., A. Beckenstein, E. Koufer, and R. D. Murphy (1975). *The Economics of Multi-Plant Operation: An International Comparison Study.* Cambridge, MA: Harvard University Press.

[68] Shaked, A. and J. Sutton (1983). Natural oligopolies. *Econometrica 51*, 1469-1484

[69] Slade, M. E. (2004). Market power and joint dominance in UK brewing. *Journal of Industrial Economics 52*, 133-163

[70] SNIC, Sindicato Nacional da Indústria do Cimento (1991 to 2002). *Annual Reports.* Rio de Janeiro, Brazil.

[71] Soares, M. G. and J. V. Caixeta Filho (1996). Características do mercado de frete para cargas agrícolas. *Preços Agrícolas 121*, 21-25, Piracicaba, Brazil

[72] Staiger, D. and J. H. Stock (1997). Instrumental variables regression with weak instruments. *Econometrica 65*, 557-586

[73] Stigler, G. (1950). Monopoly and oligopoly by merger. *American Economic Review 40*, Papers and Proceedings, 23-34

[74] Suslow, V. (1986). Estimating monopoly behavior with competitive recycling: An application to Alcoa. *Rand Journal of Economics 17*, 389-403

[75] Sutton, J. (1991). *Sunk Costs and Market Structure: Price Competition, Advertising and the Evolution of Concentration.* Cambridge, MA: The MIT Press.

[76] Sutton, J. (1998). *Technology and Market Structure: Theory and History.* Cambridge, MA: The MIT Press.

[77] Syverson, C. (2004). Market structure and productivity: A concrete example. *Journal of Political Economy 112*, 1181-1222

[78] Thisse, J. F. and X. Vives (1988). On the strategic choice of spatial price policy. *American Economic Review 78*, 122-137

[79] Town, R. (1992). Merger waves and the structure of merger and acquisition time-series. *Journal of Applied Econometrics 7* (supplement), 83-100

[80] UNCTAD (2000). Cross-border mergers and acquisitions and development. *World Investment Report.* New York and Geneva: United Nations

[81] Van Wegberg, M. (1994). Why do mergers occur in waves? Business cycles, bandwagons and the merger for market power paradox. METEOR Discussion Paper.

[82] Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data.* Cambridge, MA: The MIT Press.

[83] World Cement (several editions). Farnham, UK: Palladiam Publications Ltd.

[84] Zaghen, F. (1996). Brazil cement sector: Survival of the fittest. *Indosuez Capital Emerging Markets Equity Report.* São Paulo, Brazil.