Collusion with a Competitive Fringe: An Application to Vitamin C

Nicolas de Roos*
Department of Economics,
Sydney University

Preliminary Version

January 2004

Abstract

In a recent high profile case of collusion in the market for vitamin C, the cartel initially accommodated a fringe competitor before ultimately collapsing under the competitive burden. In this paper, the cartel's decision of when to dissolve is endogenised within a dynamic model of collusion. Demand estimates and cost information from the vitamin C market are used to calibrate the model. Results are intimately linked to the dynamic nature of the model. A cartel is found to persist only while fringe competitors remain small; fringe competitors invest heavily while a cartel is operating; entry deterrence is mitigated if cartel members are able to accommodate entry; and firms of an intermediate size are the most likely to accommodate entry.

1 Introduction

A cartel could respond in many ways to the entrance of a competitor. In the market for vitamin C, a cartel decided to accommodate the emergence of a collection of Chinese firms by maintaining cartel policies until the Chinese firms had captured a substantial fraction of the market.[1]

*Author's contact information: e-mail: n.deroos@econ.usyd.edu.au; phone: +61 2 9351 7079

[1]The experience in the vitamin C market shares similarities with the bromine cartel of the late nineteenth century analysed by Levenstein (1995). In this market, a cartel was maintained despite the existence of a lower cost entrant for a brief period before the entrant joined the cartel.

Most models of collusion in the presence of a competitive fringe apply a static or repeated games framework with symmetric firms. Such an approach is not capable of explaining this type of behaviour. This paper develops a dynamic model of collusion that endogenises cartel behaviour in the presence of a competitive fringe. It seeks to address questions about when a cartel will decide to break up when faced with entry, and when it will try to maintain cartel policies. The model is also applied to the vitamin C market, where it provides one rationale for the experiences there. However, it might also provide guidance for any market in which entry and collusion are possibilities and information asymmetries are present.

Most existing models of collusion do not consider the issues of entry and competition from outside the cartel.[2] The literature on collusion in the presence of a competitive fringe tends to use a static or supergame-theoretic framework with symmetric firms. For example, in a symmetric Cournot model with linear cost and demand, Selten (1973) finds that the stability of a cartel is very sensitive to the number of firms in the market.[3] A cartel can be sustained with four or fewer firms, but is unlikely to prevail with six or more. Shaffer (1995) examines a quantity setting cartel that can act as a Stackelberg leader to, or act simultaneously with, a competitive fringe that adopts either Cournot or Bertrand behaviour, finding the existence of a stable cartel. Posada (2001) uses both a static and repeated games framework to analyse a Stackelberg leadership cartel with a competitive fringe operating in a differentiated products market. It is found that cartel stability depends on the degree of product differentiation as well as the cartel size and industry size. In contrast to symmetric, homogeneous product models, a free rider problem, where all firms wish to be fringe competitors, does not necessarily emerge.

However, a cartel's problem when faced with a competitive fringe is typically dynamic. A fringe firm left unchecked can quickly become a formidable competitor. In addition, by imposing symmetry, no explanation can be offered for the existence of a small fringe firm. In symmetric models involving homogeneous products, the fringe firm will typically free ride on the cartel and produce more and receive greater profits than cartel members.

To address these concerns, the model developed below is inherently dynamic, allowing asymmetric firms to make decisions about investment, entry, exit, and collusion. The model incorporates fringe competition into the model developed in de Roos (2004). It shares many of the modelling features of the dynamic collusion model of Fershtman and Pakes (2000), but differs in a few important respects. First, the possibility of fringe competition is included. Second, the nature of the collusive agreement differs substantially. Third, the nature of punishment is quite different. The implications of the latter two are discussed in de Roos (2004), and are not

[2] A brief discussion of this literature is contained in de Roos (2004).

[3] For a more extensive list of references for static and repeated game models of cartels with fringe competition, see Posada (2001).

taken up here.

In the model developed, firms engage in repeated quantity competition in a market for a homogeneous product, subject to capacity constraints.[4] Firm capacities act as state variables of the model. Each period, firms choose investment spending aimed at increasing capacity and thereby increasing future profitability. Firms are free to exit at any time and make decisions about collusion and punishment. A single firm may choose to enter the market each period if market conditions are favourable. Should entry occur while a cartel is operating, the entrant becomes a fringe competitor alongside the cartel. The cartel then has to decide whether, and if so, when, to dissolve in the face of this competition.

Several major assumptions are made in order to make the model tractable. In order to examine the implications of a competitive fringe, it is assumed that a firm entering a market in which a cartel operates will compete with the cartel until the cartel wishes to dissolve.[5] The timing of the cartel's decision to dissolve is endogenous. Also by assumption, the competitive fringe is represented by a single firm. Given the computational burden of the solution algorithm, a fringe of many firms cannot be handled. However, the cartel's decision problem of how to handle a competitive fringe is clearly illustrated with a single fringe firm. It is further assumed that cartel members cannot become fringe competitors. This assumption carries some bite. One consequence is that the cartel can be disrupted by a single capacity-constrained cartel member who is almost indifferent to collusion. Ideally, the composition of the cartel would be completely endogenously specified. However, this gives rise to a multiplicity of equilibria.[6]

The market share rule specifies that throughout the life of a cartel, cartel members maintain the market shares they enjoyed at the time collusion began. The rationale for this assumption is discussed in some detail in de Roos (2004).[7] Several empirical examples appear to bear out this modelling assumption. According to industry experts, explicit market share rules of this nature applied in the markets for lysine and vitamins C, E, A, and B2 (Riboflavin). In addition, the realised market shares in most of the vitamin markets covered by the vitamin conspiracy are consistent with this rule of thumb.

An important feature of the cartels operating in the vitamin markets was the extent and nature of multi-market contact amongst participating firms. This paper does not seek to address this issue. The paper uses the market for human consumable vitamin C as an empirical example.

[4]We could think of capacity in a broad sense as any factor which limits sales in the short run.

[5]In a variant of the model discussed briefly below, the entrant is also allowed to join the cartel directly at any time, should the cartel accept it.

[6]Even in the absence of a dynamic model, this issue is not easy to deal with. See, for example, Bernheim, Peleg and Whinston (1987).

[7]The additional assumptions that firms face convex adjustment costs in capacity expansion and that cartel members collude on output but not investment are also discussed there.

As discussed below, the principal means by which multi-market contact was used to enforce the vitamin cartels was probably through access to required intermediate and final products.[8] In the human consumable vitamin C market, these enforcement mechanisms were not available.

The paper also makes no attempt to distinguish between a cartel's likely response to a single entrant and repeated entry. Along the lines of Kreps and Wilson (1982) and Milgrom and Roberts (1982), a cartel may be more likely to respond aggressively to an entrant if it expected future entry, provided there is some uncertainty about cartel payoffs. In the current model, repeated entry is possible, but the incidence of entry is fairly low (see the results of section 5), so this assumption appears to have only a minor affect.

The rest of the paper is structured as follows. In section 2, the experience with collusion in the vitamin C market is briefly discussed. This serves the dual purposes of placing the modelling framework in context and providing a real application. In section 3, the model is outlined. Section 4 discusses the specification of demand and cost parameters for the model. Demand parameters are obtained by estimating a market demand equation. Cost parameters are obtained by comparing two sources. First, firm own-price elasticities are inverted to yield implied marginal costs, conditional on a behavioural assumption. These are compared to estimates from industry sources. Results are presented in section 5. The type of firm behaviour is characterised for different types of market structure and compared to the type of behaviour predicted by similar models that omit fringe competition. The model is then simulated to generate a set of industry and welfare statistics. Section 6 concludes.

2 Collusion in the Market for Vitamin C

The market for vitamin C provides an engaging example of collusion to study for several reasons. First, following confessions by the major participating firms, we know collusion took place at a time when collusion was illegal. Second, the cartel operated in tandem with a fringe competitor, allowing us to examine the dynamic problem faced by incumbent firms of when to dissolve a cartel when there is a competitive fringe. Third, the major firms colluded in many vitamins markets simultaneously. A rigorous consideration of the implications of multi-market contact for cartel operations is beyond the scope of the current paper.

The "vitamins cartel" encompassed over a dozen vitamins firms, affected a total of more than a dozen vitamin products in many countries, and operated from at least as early as January 1990 until February 1999.[9] The three main firms, Hoffmann La Roche, BASF, and Rhone Poulenc,

[8] Bernheim and Whinston (1990) take up the issue of multi-market contact for collusive industries. However, the issue of access to intermediate and final products is not considered.

[9] This information can be obtained from various Department of Justice press releases. Data and sources are

operated in almost every market and accounted for the majority of sales. Many other firms operated in several markets, and several firms were single market operators. Participants in the cartel have pled guilty to collusion in the vitamin C market over the period January 1991 until the Fall of 1995.[10] In many of the vitamins markets covered by the cartel, including the market for vitamin C, the market shares of the participants were fixed at the beginning of the cartel based on prevailing market shares.

Many cartels are plagued by enforcement problems. In the vitamins cartel, a variety of enforcement mechanisms may have been available. If cartel members maintain excess production capacity and stipulate inter-firm sales requirements if sales quotas are not met, then deviation can be discouraged. International trade records can be used to verify sales claims. Another potential enforcement tool is the restriction of access to vital resources. In the vitamins cartel, member firms were dependent on each other for crucial intermediate inputs. This may have played an important role in the enforcement of the cartel. Cartel members also confessed to collusion in the lucrative "premix" market. Premix is essentially a combination of many vitamins in a prepared mixture primarily for the use of feed manufacturers. By controlling access to vitamins products, cartel members could potentially foreclose access to this valuable market. If these factors played a large role in enforcing the cartel, they may have also presented formidable barriers to entry for the vitamins markets. This might explain the extraordinary success of the vitamins cartel. There is also some circumstantial evidence of firms actively creating interdependence to strengthen the cartel, for example, by halting production of a required commodity for premix to create the need to buy from another firm.[11]

For the purposes of discussion, define the market for "vitamin C" as the wholesale market for human consumable ascorbic acid. This includes several physical forms of ascorbic acid which are chemically identical. The major forms included are ascorbic acid in powder or granular form, used to add nutritional content to prepared food and beverages; direct compressible ascorbic acid for use in vitamin C and multivitamin tablet preparations; and sodium ascorbate, used for flour enrichment. Production of different forms of ascorbic acid from the raw product is relatively cheap and hence there is no reason a firm could not supply all forms of the product. Because the products of competing firms are chemically identical, as an approximation, we could think of vitamin C as a homogeneous good. The purchasers of vitamin C are relatively unconcentrated. Transacted prices are not publicly posted and large buyers will typically enter

---

discussed in the Data section below. The bulk of the information discussed below was collected from industry sources.

[10] This does not necessarily preclude the existence of prior collusion.

[11] Genesove and Mullen (1999) note other examples of non-price collusion for the Sugar Institute cartel of the early twentieth century.

into annual price contracts.

Access to intermediate products is not a limiting factor for production of vitamin C. In addition, most of the demand for human consumable vitamin C is not for use in premix products. Hence, entry barriers for the vitamin C market might be less daunting than in other vitamins markets. The production technology also probably constitutes only a minor barrier to entry. The production process is technically simpler than in many of the other vitamins markets. The first step in the production process involves conversion of glucose or dextrose into sorbitol through hydrogenation. The production process employed by the majority of firms in the 1990s entails a combination of synthetic chemical steps and fermentation steps to convert sorbitol into ascorbic acid. There are some economies of scale in the synthetic steps of the process. In 1988, Chinese firms employed a process that replaces many of the synthetic chemical steps with a single fermentation step. This process reduces the cost, time, and scale required for production. In the late 1990s, other firms began shifting some production towards this process. As more firms switch processes, we might expect economies of scale to diminish in importance. The absence of powerful entry barriers could provide one reason the vitamin C cartel was less stable than the vitamins cartel as a whole.

Prior to the plea period, the market was dominated by two major firms; the Swiss firm Hoffman La Roche, and Takeda based in Japan. A German firm, BASF, and a U.S. firm, EMI Industries, also maintained small market shares. In this period, a group of Chinese firms held a small market share, and were selling vitamin C at a discount to their competitors. We shall think of the Chinese firms collectively as a single firm operating outside the cartel.[12]

Upon the formation of the cartel, the prices of the major cartel members rose steadily, with an increase of more than a third in the U.S. market (based on a simple quantity-weighted average) from a trough in December 1990 to a plateau beginning in July 1993. Cartel prices were then remarkably stable until September 1995. Throughout this period, cartel prices were consistently undercut by the Chinese firms. The Chinese market share grew steadily over this period, rising to over a third of world output by September 1995. Following the dissolution of the cartel in the market for vitamin C, prices fell rapidly, dropping below the levels prevailing before the cartel began operations within a year. Prices continued to fall to be less than half of the cartel high by December 1999.

To summarise, the vitamin C cartel was faced with a competitive fringe in the form of the Chinese firms. The cartel decided to accommodate the fringe while it was small, and then abandon the cartel and compete vigorously once the fringe competitor had grown. The model below embeds this fringe competition theme within a dynamic model of collusion.

[12] Industry sources suggest that the Chinese firms were State controlled.

3 The Model

The model of de Roos (2004) is augmented to allow the presence of a fringe competitor operating outside a cartel. Figure 1 depicts the order of decision making in each period. Four possible competitive regimes govern the profits accruing to each firm in each period. In the non-cooperative and punishment regimes, profits are determined by the Nash equilibrium in quantities to a one-shot capacity constrained game. In the collusive regime, profits are determined by joint profit maximisation subject to capacity constraints and a market share constraint. The market share constraint provides that throughout the life of a cartel, participating firms maintain the share of industry capacity they enjoyed at the onset of the cartel. In the fringe regime, the cartel prevails despite the presence of a fringe competitor. Cartel and fringe outputs are determined by the Nash equilibrium in quantities to a two-firm one-shot capacity constrained game.[13] Individual cartel member output is then determined by the market share constraint.

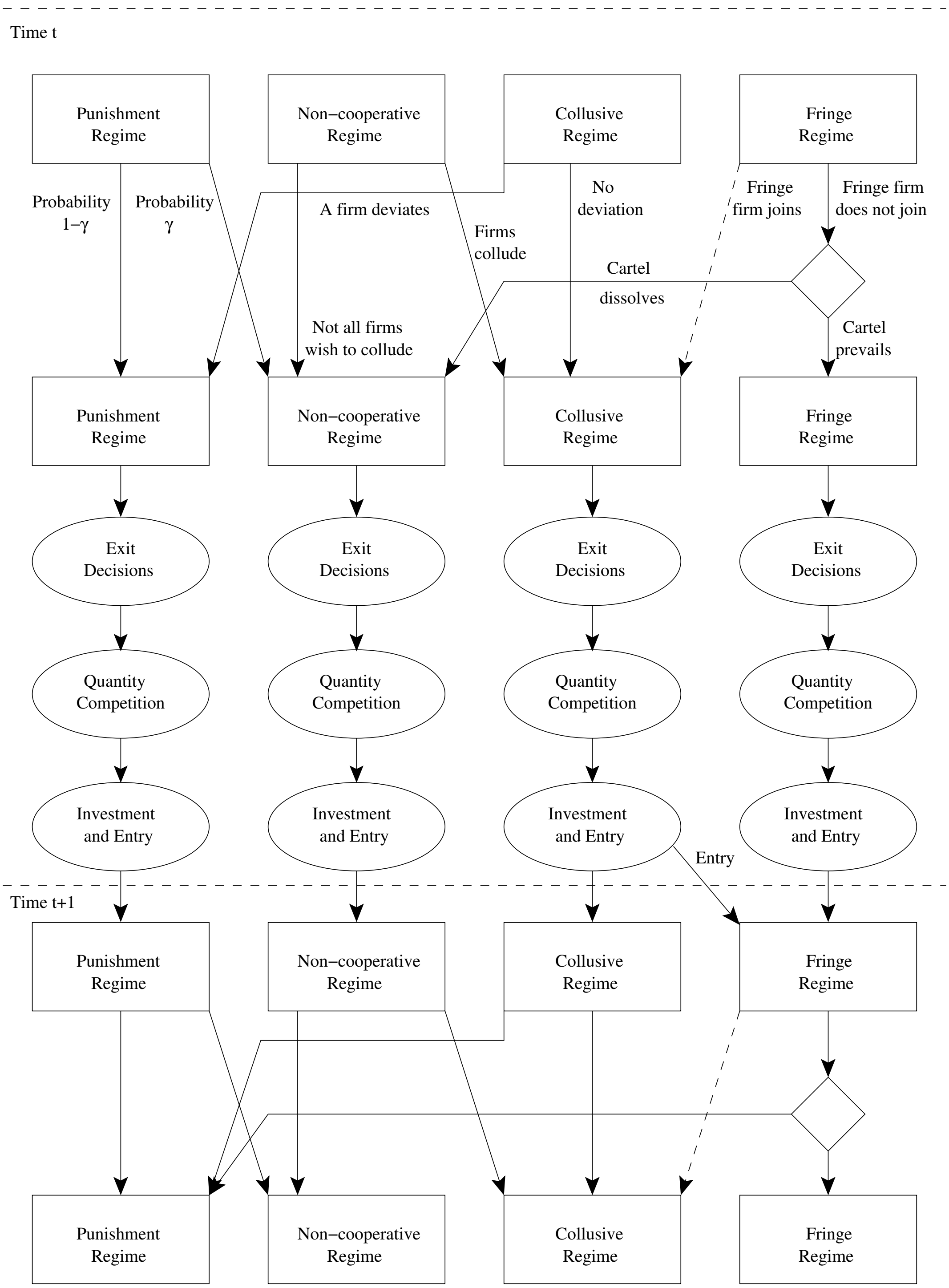
If the incumbent firms achieve a consensus on collusion, play switches from the non-cooperative to the collusive regime. If a firm enters while a cartel is operating, play switches to the fringe regime. Starting from the fringe regime, the incumbent firms decide whether to dissolve the cartel or continue colluding. A broken line is used to indicate a variant of the model in which the fringe firm has the opportunity to join the cartel directly.[14]

If any of the incumbent firms wish to dissolve the cartel while the fringe regime is operating, play switches to the non-cooperative regime. A consensus to maintain the cartel ensures the fringe regime is maintained. Notice that there is no punishment imposed on firms dissolving the cartel when faced with a competitive fringe. By contrast, if any firm wishes to dissolve the cartel while operating in the collusive regime, they can do so, but this will invoke the punishment regime. While in the punishment regime, there is no possibility of collusion. With fixed probability, γ, firms negotiate their way back to the non-cooperative regime. The rationale for this disparity between the fringe and collusive regimes is the following. While in the fringe regime, competition from the fringe firm represents a substantial present or future threat to the cartel. Dissolving the cartel could plausibly reflect a defense against such competition rather

[13] Cartel output is constrained by aggregate cartel capacity.

[14] More precisely, in this variation, a two stage game is played. In the first stage, the fringe firm decides whether or not to join the cartel and the cartel members decide whether or not to accept the fringe firm. If the fringe firm joins, play switches to the collusive regime. If the fringe firm does not join, then in the second stage the incumbent firms decide whether to continue their collusive venture. The two stage game is not emphasised for two main reasons. First, the cartel will tend to break up before the fringe firm joins the cartel, making the first stage largely redundant. Second, convergence problems were encountered in the calculation of a solution to the two stage game problem.

Figure 1: Sequence of Events



than a systematic disagreement amongst cartel members.[15]

The following sequence of play occurs each period. At the beginning of the period, decisions about which competitive regime to follow, described above, are made first. Incumbent firms then decide whether to exit the market. Quantity competition then takes place subject to the prevailing competitive regime and incumbent exit decisions. Finally, incumbents decide how much to invest and potential entrants decide whether to enter. Entrants do not produce in their first period of operation.

3.1 The State Space

The state of the world is described by the prevailing competitive regime, the capacities of the firms, and the market shares of colluding firms. Define $\Omega \equiv \left\{\tau, \tau g, \tau g^2, \ldots, \tau g^{\bar{k}}\right\}$ as a discrete set of feasible capacities that can be mapped onto the set of positive integers, where $g > 1$, $\tau > 0$, and $\bar{k}$ is in the set of positive integers. Minimum and maximum feasible capacities are given by τ and $\tau g^{\bar{k}}$, respectively. Define $\omega_t = \{\omega_{i,t}\}_{i=1}^{n_t}$ to be the vector of capacities of the incumbent firms, where $\omega_{i,t} \in \Omega$, i indexes firms, t indexes time periods, and n_t is the number of active firms in period t. Similarly, define $\mu_t = \{\mu_{i,t}\}_{i=1}^{\bar{n}_t}$ to be the vector of capacities at the time the last collusive agreement was negotiated, where $\mu_{i,t} \in \Omega$ and $\bar{n}_t$ is the number of cartel members. Let N^C represent the set of firms in a cartel, and an f subscript denote a fringe competitor. Then, the state space, conditional on being in each of the competitive regimes can be written as follows:

$$
\begin{aligned}
S^N &= S^P \equiv \left\{\omega_t \middle| \text{ for } j > i, \omega_{i,t} \geq \omega_{j,t}\right\} \\
S^C &\equiv \left\{\omega_t, \mu_t \middle| \text{ for } j > i, \omega_{i,t} \geq \omega_{j,t} \text{ and if } \omega_{i,t} = \omega_{j,t}, \text{ then } \mu_{i,t} \geq \mu_{j,t}\right\} \qquad (1) \\
S^F &\equiv \Big\{\omega_{f,t}, \omega_{-f,t}, \mu_t \Big| \text{ for } j, i \in N^C \text{ and } j > i, \\
&\quad \omega_{i,t} \geq \omega_{j,t} \text{ and if } \omega_{i,t} = \omega_{j,t}, \text{ then } \mu_{i,t} \geq \mu_{j,t}\Big\},
\end{aligned}
$$

where the superscripts N,P,C, and F denote the non-cooperative, punishment, collusive, and fringe regimes, respectively, and a negative subscript denotes omission of that element.[16] The non-cooperative and punishment regimes are fully described by the vector of incumbent capacities. In the collusive regime, a firm's share of profits is determined by its share of industry

[15] A variant of the model was also considered in which dissolving the cartel in the presence of fringe competition led to the punishment regime rather than the non-cooperative regime. Results are broadly similar with this variation. The primary difference is that firms are less likely to collude when there are only two firms in the market because there is a stronger desire to deter entry.

[16] In the collusive and fringe regimes, we can further restrict attention to states in which $min_i \{\mu_{i,t}\} = \tau$ because all other μ_t vectors share the common factor g.

capacity at the onset of collusion, so the vector of initial capacities is also important. In the fringe regime, the identity of the fringe competitor becomes relevant.

3.2 Profit Functions

Profits are determined by quantity competition for a homogeneous product subject to capacity constraints and, when a cartel is operating, market share constraints. Because state variables are independent of current production decisions, the production game can be treated in isolation from all other firm decisions. The inverse demand function is given by $P(Q_t) = a - bQ_t$, where Q_t is market output, and a and b are demand parameters to be estimated in Section 4. Firms are assumed to face a constant marginal cost of production of mc.

In the non-cooperative and punishment regimes, the profit vector, $\pi^N(\omega_t) = \left\{\pi_i^N(\omega_t)\right\}_{i=1}^{n_t}$, is determined by the unique solution to a one-shot capacity-constrained quantity game. The collusive profit vector, $\pi^C(\omega_t, \mu_t) = \left\{\pi_i^C(\omega_t, \mu_t)\right\}_{i=1}^{n_t}$, is determined by joint profit-maximisation subject to capacity and market share constraints. Details of the calculation of these profit functions can be found in de Roos (2004). Fringe profits, $\pi^F\left(\omega_{f,t}, \omega_{-f,t}, \mu_t\right) = \left\{\pi_i^F\left(\omega_{f,t}, \omega_{-f,t}, \mu_t\right)\right\}_{i=1}^{n_t}$, are determined as follows. A one-shot capacity constrained duopoly game between the fringe firm and the cartel determines the output of the fringe firm and the cartel as a whole. Outputs of individual cartel members are then determined by the market share constraint.

3.3 Investment, Entry, and Exit

The probability of successful investment is an increasing, concave function of investment spending. Define $\eta_{i,t} \in \{1, g\}$ and $\nu_t \in \{1, g\}$ to be the outcome of the firm's investment process and the outcome of some exogenous process capturing developments in the industry, respectively. Then, the transition of firm i's capacity is governed by

$$\omega_{i,t+1} = \omega_{i,t}\frac{\eta_{i,t+1}}{\nu_{t+1}} \tag{2}$$

$$\eta_{i,t+1} = \begin{cases} g & \text{with probability } \frac{\alpha x_{i,t}}{1+\alpha x_{i,t}}, \\ 1 & \text{with probability } \frac{1}{1+\alpha x_{i,t}} \end{cases} \tag{3}$$

$$\nu_{t+1} = \begin{cases} g & \text{with probability } \delta, \\ 1 & \text{with probability } 1-\delta. \end{cases} \tag{4}$$

where the constant $\alpha > 0$ determines the effectiveness of investment, and $x_{i,t}$ is firm i's investment expenditure in period t.

Entry decisions are made concurrent with investment decisions. Each period, a single potential entrant observes an entry cost draw, x_e, from a uniform distribution $U\left(x_e^{min}, x_e^{max}\right)$ and then decides whether to enter by comparing the expected discounted value of entering with capacity ω_e with the cost of entry, x_e. Should the firm decide to enter, it begins production in the following period.

Incumbent firms are free to exit in every period before they engage in quantity competition. An exiting firm receives a pay-off of ϕ and takes no further part in the game. A firm will therefore exit if the expected discounted value of remaining in the market is less than ϕ.

3.4 The Bellman Equations

Separate Bellman equations are defined, conditional on being in each of the four competitive regimes.[17] Notation is as follows. $\Psi \in \{P,N,C,F\}$ denotes the current operating regime. A negative subscript denotes omission of a single element, $\omega_{-i} \equiv (\omega_1, \ldots, \omega_{i-1}, \omega_{i+1}, \ldots, \omega_{n_t})$. Semi-colons separate values of the same state variable for different firms, and commas separate different state variables. Thus, $V^C\left(\omega_j;\omega_{-j},\mu_j;\mu_{-j}\right)$ describes firm j's value of operating in the collusive regime with capacity and capacity at collusion given by ω_j and μ_j, respectively, while the corresponding state values of its competitors are ω_{-j} and μ_{-j}, respectively. Profits for firm j are given by $\pi^C\left(\omega_j;\omega_{-j},\mu_j;\mu_{-j}\right) = \pi_j^C(\omega,\mu)$, $\pi^N\left(\omega_j;\omega_{-j}\right) = \pi_j^N(\omega)$, and $\pi^F\left(\omega_j;\omega_f;\omega_{-\{f,j\}},\mu_j;\mu_{-j}\right) = \pi_j^F\left(\omega_f;\omega_{-f},\mu_{-f}\right)$ for the collusive, non-cooperative, and fringe regimes, respectively. $p(.|.)$ describes the conditional state transition probabilities. Finally, indicator functions describe transition between competitive regimes. $I^C(\omega) \in \{0,1\}$ governs transition from the non-cooperative regime to the collusive regime, with

$$I^C(\omega) = 1 \Leftrightarrow V^C\left(\omega_j;\omega_{-j},\omega_j;\omega_{-j}\right) \geq V^N\left(\omega_j;\omega_{-j}\right)\ \forall j; \tag{5}$$

$I^P(\omega,\mu) \in \{0,1\}$ governs transition from the collusive regime to the punishment regime, with

$$I^P(\omega,\mu) = 1 \Leftrightarrow V^P\left(\omega_j;\omega_{-j}\right) > V^C\left(\omega_j;\omega_{-j},\mu_j;\mu_{-j}\right) \text{ for any } j; \tag{6}$$

and $I^D\left(\omega_f;\omega_{-f},\mu\right) \in \{0,1\}$ describes the cartel members' decision to abandon the cartel, with

$$I^D\left(\omega_f;\omega_{-f},\mu\right) = 1 \Leftrightarrow V^N\left(\omega_j;\omega_{-j}\right) > V^F\left(\omega_f;\omega_j;\omega_{-j},\mu_j;\mu_{-j}\right) \text{ for any } j \neq f. \tag{7}$$

[17] The computational algorithm for the model without fringe competition is described in de Roos (2004). The current algorithm is written in the C programming language and can be obtained from the author upon request. Owing to the number of discrete control variables, perfect convergence cannot be obtained for the model outlined above. To handle this problem, stochastic elements are incorporated into the collusion and punishment decisions, and the cartel dissolution decision in the fringe regime. Stochastic shocks take the form of i.i.d. firm-specific cost draws relating to these decision variables. With these assumptions, we can come arbitrarily close to convergence in both the value function and the policy functions for the parameters considered below. However, convergence is sensitive to the parameters chosen.

The Bellman equations below describe the value to firm j for each element of the state space.[18]

$$
\begin{aligned}
V^P \quad & \left(\omega_j;\omega_{-j}\right) = \max\left\{\phi, \pi^N\left(\omega_j;\omega_{-j}\right) + \max_{x\geq 0}\left[-x\right.\right. \\
& \left.\left. + \beta\sum_{\omega'}\left(\gamma V^N\left(\omega'_j;\omega'_{-j}\right) + (1-\gamma)V^P\left(\omega'_j;\omega'_{-j}\right)\right)p\left(\omega'_j|\omega_j,x\right)p\left(\omega'_{-j}|\omega,\Psi=P\right)\right]\right\} \quad (8)\\
V^N \quad & \left(\omega_j;\omega_{-j}\right) = \max\left\{\phi, \pi^N\left(\omega_j;\omega_{-j}\right) + \max_{x\geq 0}\left[-x\right.\right. \\
& +\beta\sum_{\omega'}\left(I^C\left(\omega'\right)V^C\left(\omega'_j;\omega'_{-j},\omega'_j;\omega'_{-j}\right) + \left(1-I^C\left(\omega'\right)\right)V^N\left(\omega'_j;\omega'_{-j}\right)\right) \\
& \left.\left. p\left(\omega'_j|\omega_j,x\right)p\left(\omega'_{-j}|\omega,\Psi=N\right)\right]\right\} \quad (9)\\
V^C \quad & \left(\omega_j;\omega_{-j},\mu_j;\mu_{-j}\right) = \max\left\{\phi, \pi^C\left(\omega_j;\omega_{-j},\mu_j;\mu_{-j}\right) + \max_{x\geq 0}\left[-x\right.\right. \\
& +\beta\sum_{\omega'}\left(I^P\left(\omega',\mu\right)V^P\left(\omega'_j;\omega'_{-j}\right) + \left(1-I^P\left(\omega',\mu\right)\right)V^C\left(\omega'_j;\omega'_{-j},\mu_j;\mu_{-j}\right)\right) \\
& \left.\left. p\left(\omega'_j|\omega_j,x\right)p\left(\omega'_{-j}|\omega,\mu,\Psi=C\right)\right]\right\} \quad (10)\\
V^F \quad & \left(\omega_j;\omega_f;\omega_{-\{f,j\}},\mu_j;\mu_{-j}\right) = \max\left\{\phi, \pi^F\left(\omega_j;\omega_f;\omega_{-\{f,j\}},\mu_j;\mu_{-j}\right) + \max_{x\geq 0}\left[-x\right.\right. \\
& +\beta\sum_{\omega'}\left(\left(I^D\left(\omega'_f;\omega'_{-f},\mu\right)V^N\left(\omega'_j;\omega'_{-j}\right)\right.\right. \\
& \left.\left. + \left(1-I^D\left(\omega'_f;\omega'_{-f},\mu\right)\right)V^F\left(\omega'_f;\omega'_j;\omega'_{-\{f,j\}},\mu_j;\mu_{-j}\right)\right)\right) \\
& \left.\left. p\left(\omega'_j|\omega_j,x\right)p\left(\omega'_{-j}|\omega_f,\omega_{-f},\mu,\Psi=F\right)\right]\right\}, \quad (11)
\end{aligned}
$$

In each of the regimes, incumbent firms can choose to exit the industry and receive the scrap value, ϕ, or continue operations and receive a continuation value comprising current period profits and the expected discounted value of future returns. Let us consider the fringe regime in a little more detail.[19] In the fringe regime, profits depend on the capacity of the fringe firm, the vector of capacities of cartel members, and the vector of capacities of cartel members at the time collusion was initiated. The indicator function $I^D(.)$ determines the ensuing regime. It takes a value of 1 if any of the incumbent firms wish to dissolve the cartel. The distribution of the firm's own state next period depends on investment spending in the current period, and

[18]The Bellman equations highlight the firm's investment decisions. In the model described and solved, entry into the collusive regime switches play to the fringe regime. To simplify exposition, equation 10 disregards this role of entry. Equation 11 is written from the perspective of the cartel member. An analogous expression holds for the fringe competitor.

[19]The punishment, non-cooperative, and collusive regimes are described for a related model in de Roos (2004).

the distribution of its competitors' states next period depends on the investment, exit, and entry decisions made by its rivals this period.

Entry decisions depend on the value functions defined above. Each period, one potential entrant observes an entry cost draw, x_e before deciding whether to enter. The entrant builds a plant of capacity ω_e and begins production the following period with a capacity of ω_e' with probability $p_e(\omega_e')$, where $p_e(\omega_e) = 1-\delta$ and $p_e(\omega_e/g) = \delta$. Randomness in ω_e' reflects the uncertain development of the industry during plant construction. Equations (12) to (14) describe the conditions under which entry occurs for the punishment, non-cooperative, and collusive regimes, respectively, with entry occurring if the inequalities are satisfied.[20]

$$\beta \sum_{\omega'} V^P\left(\omega_e'; \omega_{-e}'\right) p_e\left(\omega_e'\right) p\left(\omega_{-e}' | \omega, \Psi = P\right) \quad > \quad x_e \tag{12}$$

$$\beta \sum_{\omega'} V^N\left(\omega_e'; \omega_{-e}'\right) p_e\left(\omega_e'\right) p\left(\omega_{-e}' | \omega, \Psi = N\right) \quad > \quad x_e \tag{13}$$

$$\beta \sum_{\omega'} V^F\left(\omega_e'; \omega_{-e}', \mu\right) p_e\left(\omega_e'\right) p\left(\omega_{-e}' | \omega, \mu, \Psi = C\right) \quad > \quad x_e. \tag{14}$$

The entry decisions depend on the current competitive regime for two reasons. First, the value function differs by competitive regime. Second, investment decisions, and therefore state transition probabilities, depend on the competitive regime. The above entry policies combined with random entry costs give rise to probability measures describing the perceptions of the incumbent firms of the probability of entry at the time investment decisions are made. For the punishment, non-cooperative, and collusive regimes, these are given by $\chi_e^P(\omega_e; \omega_{-e})$, $\chi_e^N(\omega_e; \omega_{-e})$, and $\chi_e^C(\omega_e; \omega_{-e}, \mu)$, respectively. In the collusive regime, the probability of entry depends on the fringe regime value function because an entrant becomes a fringe competitor.

4 Estimation of Static Parameters

In the model outlined in section 3, specification of the profit function requires knowledge of the demand and cost parameters a, b, and mc. In this section, estimates of the demand parameters will be obtained and the choice of the marginal cost parameter will be discussed.

[20]Entry is not currently considered into the fringe regime because at least 3 firms are required for the fringe regime to be defined and, due to the computational burden of the state space, the model can be solved for a maximum of 3 firms.

4.1 Demand Specification

As noted in section 2, the major uses of vitamin C are to add nutritional content to prepared food and drinks and as a direct nutritional supplement in the form of vitamin C and multivitamin tablets. In both cases, the end use is for individuals supplementing their natural vitamin intake. Hence, anything that influences perceptions about dietary needs could affect the demand for vitamin C. Much scientific research is conducted into the benefits of dietary supplements, so we might expect the demand for vitamin supplements to vary over time with public perceptions of their effectiveness. In the short run, while the concentration of vitamin supplements in prepared foods is fixed, vitamin C demand should also depend on the demand for these final products.

The following log-linear and linear aggregate demand specifications were considered[21]

$$\ln Q_t = \alpha_1 - \beta_1 \ln P_t + X_t \gamma_1 + \varepsilon_{1,t} \quad (15)$$

$$Q_t = \alpha_2 - \beta_2 P_t + X_t \gamma_2 + \varepsilon_{2,t}, \quad (16)$$

where Q_t denotes market sales volume in period t, P_t is a volume-weighted market price, X_t is a vector of explanatory variables, and $\varepsilon_{i,t}$, $i = 1,2$, is the unobserved demand error. X_t contains an index of positive media citations related to vitamin C, constructed by counting the number of articles about vitamin C with a positive bent in each period and applying a smoothing function to the resultant time series. An analogous negative citations index is also included. Other explanatory variables included are the United States population, U.S. disposable personal income, and a general food index. In equation (15), natural logarithms are taken for all components of X_t, with the exception of the citations indices. Data sources are described in the next section.

Market prices and quantities are jointly determined by equations (15) and (16), and a behavioural assumption about the prevailing competitive regime. To handle the simultaneity problem, the demand equations are estimated by two stage least squares. The most dramatic shifts in prices and quantities surround the onset and the break up of the cartel. Consequently, one potential instrument for prices is the cartel period itself. This is a valid instrument because the timing of the decision to terminate the cartel relates primarily to the alarming rise in market share of the group of Chinese firms, and is unlikely to be strongly influenced by demand shocks. Owing to the existence of contracts, adjustment costs in production, concerns about cartel detection, and uncertainty about rival behaviour, price and quantity movements are smooth, rather than jumping between competitive regimes. The cartel period is interacted with a trend and a

[21] Some product differentiation exists amongst suppliers of vitamin C. Ideally, we would estimate a differentiated products model of demand. Firm-specific own- and cross-price elasticities are obtained for such a model. Details are contained in section 4.4 below. However, owing to the collinearity of firm prices, some of the elasticities are imprecisely estimated and not robust to specification changes. Accordingly, these estimates are not used in the specification of the profit function.

quadratic term to form additional instruments that capture the gradual nature of price movements. Other instruments used include bilateral exchange rates between the United States and Japan, Germany, China, and the UK (major production regions of participating firms), and the average hourly wage of chemical production workers.

4.2 Data

Firm-level aggregate quantities and quantity-weighted average prices for the U.S. market are obtained from a confidential industry source. The data covers three of the major cartel members over the period January 1991 to December 1999, accounting for the vast majority of cartel sales. Data for a fourth, smaller cartel member is only available from January 1996, and is therefore omitted. Chinese quantity and price data is obtained from the U.S. Census Bureau's Schedule B import data. Sales volumes and values are used to form an implicit price series. All prices are in $/kg. Quantities are in thousands of kg.

Turning to the demand and cost shifters, U.S. population is obtained from the Census Bureau. Population is measured in millions. Bilateral exchange rates are obtained from Datastream. The average hourly earnings of chemical production workers is obtained from the Bureau of Labor Statistics.

4.3 Estimation Results

Equations (15) and (16) were estimated by ordinary least squares (OLS) and two stage least squares (2SLS). Insignificant variables were sequentially eliminated and tested for joint significance. Columns 1 and 2 present the log-linear demand estimates. The OLS and 2SLS estimates are very similar because most of the price and quantity variation is due to the formation and cessation of the cartel rather than changes in demand conditions.

The results suggest that demand for vitamin C is very inelastic. This is inconsistent with the behaviour of a profit maximising monopolist. However, there are several potential reasons we might observe this in the presence of a cartel. The cartel was clearly constrained by competition from the Chinese firms in the first half of the sample, and dissolved in the second half of the sample. Cartel members were also concerned about detection by the anti-trust authorities, and cheating from within the cartel. The cartel was also concerned about the possibility of entry. A limit pricing story could be told if firms outside the cartel are unsure of the competitive regime operating in the market. Consequently, we could not expect the cartel to behave like a profit maximising monopolist. Of the other explanatory variables, negative media reports tend to have a negative impact on vitamin C demand, positive reports have no discernable impact, and

Table 1: Demand Results[a]

	Equation (15)		Equation (16)	
	OLS	2SLS	OLS	2SLS
Constant	-63.527	-64.447	-7297.1	-7360.0
	(8.775)	(8.935)	(1139.7)	(1161.1)
Market Price	-0.112	-0.104	-25.661	-25.048
	(0.056)	(0.058)	(8.728)	(8.962)
Positive Citations	0.010	0.009	31.216	30.633
	(0.014)	(0.014)	(22.603)	(22.757)
Negative Citations	-0.127	-0.127	-179.7	-179.3
	(0.043)	(0.044)	(69.85)	(70.12)
U.S. Population	6.254	6.326	34.695	34.914
	(0.695)	(0.707)	(4.076)	(4.150)
Observations	108		108	
R^2	0.770	0.769	0.762	0.760
$\bar{R}^2$	0.761	0.760	0.753	0.751
Dep Var Sum of Sq	21917.57		2.795e8	
Sum of Sq Errors	1.167	1.174	3.026e6	3.048e6
Autocorrelation	0.057	0.058	0.055	0.054
Durbin-Watson	1.911	1.908	1.904	1.904

[a] Figures in parentheses () are standard errors.

population growth adds to demand.[22]

The results for the linear demand specification are very similar. These are presented in columns 3 and 4. The coefficient on market price implies a market demand elasticity that varies between 0.06 and 0.33 over the sample. The interpretation of the remaining explanatory variables is unchanged.

The profit function coefficients a and b are obtained from the linear demand estimates as follows. Sample averages of non-price variables are used to determine a. That is, $a \equiv \left(\alpha_2 + \gamma_2 \left(\sum_{t=1}^{T} X_t\right)/T\right)/\beta_2 = 72.41$. The demand slope coefficient is given by $b \equiv 1/\beta_2 =$

[22] Because population follows a trend over the sample, the coefficient on population might simply reflect trend demand growth. Substitution of a trend for the population variable yields similar results. U.S. disposable personal income was omitted because an apparently spurious negative relationship was obtained, associated with a single outlier in the data.

0.0390.

4.4 Marginal Costs

Two sources are used to inform the choice of marginal cost parameter. First, an Almost Ideal Demand System (Deaton and Muellbauer (1981)) is estimated to obtain the own- and cross-price elasticities of demand for the major firms. The market demand equation, (15), is estimated as a first step. Then, the following firm-specific equations are estimated.

$$\mathcal{S}_{i,t} = \theta_{1i} + \theta_{2i} \ln Q_t + Z_t \theta_{3i} + \sum_{j=1}^{N} \lambda_{ij} \ln P_{j,t} + \varepsilon_{i,t} \tag{17}$$

where $\mathcal{S}_{i,t}$ is the expenditure market share of firm i, and Z_t is a vector of explanatory variables.[23] Estimation of equation (17) is hampered by the collinearity of firm prices. Hence, parameters are imprecisely estimated and sensitive to specification. With an assumption of Nash-Bertrand pricing, implicit marginal costs can then be obtained by applying the own-price elasticities to the price data in the post-cartel period. That is, $mc_{i,t} = \frac{P_{i,t}(1+e_i)}{e_i}$, where e_i is firm i's own-price elasticity. Plausible implicit marginal costs of two of the cartel participants are obtained. They vary between \$6 and \$4 over the post-cartel period.

Second, a rough idea of marginal costs is obtained from industry sources, broadly confirming these estimates. Accordingly, a marginal cost parameter of $mc = 5$ is used in the analysis below. Given the size of the demand intercept, a, the profit function is not overly sensitive to small changes in the marginal cost parameter.

5 Results

Discussion of the results begins with a brief description of the results of two related models which are embedded in the model outlined above. The first is a model without collusive possibilities. The second is a model with the possibility of collusion, but without fringe competition. These models are used as points of reference to compare the results of the full model. The nature of the equilibrium is then discussed. The types of environments conducive to collusion and tolerance of a fringe competitor are described. The model is then simulated and industry statistics are presented. Throughout, comparisons are made with the other two model specifications. Finally, the results are briefly related back to the experience in the market for vitamin C.

[23] Hausman (1999) describes the calculation of firm-specific own- and cross-price elasticities from equations (15) and (17).

Table 2: Parameters Used for the Base Model

Parameter	Description	Value
a	Demand intercept	72.41
b	Demand slope	0.039
mc	Marginal cost of production	5.0
α	Investment efficiency	0.00007
ω_e	Entrant's starting capacity	$\tau g^3 = 216.0$
x_e^{min}	Minimum entry cost	$35,000$
x_e^{max}	Maximum entry cost	$65,000$
ϕ	Scrap value of firm	2000
τ	Minimum capacity	125.0
g	Capacity growth factor	1.2
$\overline{k}$	Number of feasible capacity levels	16
β	Discount factor	0.98
δ	Capacity depreciation rate	0.25
γ	Renegotiation probability	0.05
N	Maximum number of firms	3

Table 2 contains a list of the parameters used in the model, together with their values. The demand coefficients, a and b, are obtained from the linear demand estimation discussed in section 4.3, and the marginal cost parameter is discussed in section 4.4. A high discount factor is chosen to reflect the monthly planning horizon of firms in the vitamin C market and the availability of monthly data. The remaining dynamic parameters are chosen to allow a concentrated industry. Low entry costs (relative to incumbent firm values) are chosen to permit entry behaviour similar to that seen in the vitamin C market. That is, entry is possible even when incumbent firms have large capacities. With the low investment efficiency parameter chosen, the process of a fringe firm raising capacity to match its incumbent competitors can be a lengthy one.[24] Due to the computational burden of a large state space, an upper limit of three firms is imposed.

[24]A consequence of this specification is that firms will often be capacity constrained, even once established. A more flexible specification could allow different investment efficiencies for different types of firms.

5.1 Comparison with Other Models

A model without the possibility of collusion is used as a basis for comparison. This corresponds to the punishment regime with $\gamma = 0$. Hence, there is no possibility for transition to any form of collusion. The industry is simulated for $1,000,000$ periods, starting from an initial state vector of $(4,0,0)$. This corresponds to a monopoly with capacity of τg^4. Results are depicted in column 1 of Table 3. There are 3 firms in the market for the majority of the simulation periods. There is a high incidence of entry and exit. This results from a combination of two main factors. First, firms need to invest heavily to increase capacity. This is reflected in the high level of investment spending undertaken by firms. Second, entrants come in with small capacity, and can build capacity only slowly. Hence, many entrants will last only a relatively short period. This is reflected in the large variance in firm lifespans. The results also reveal that firms charge a high mark-up on marginal costs. Two factors contribute to this. First, the demand estimates reveal a very inelastic demand. Second, with the parameters chosen, firms are often capacity constrained.

Column 2 of Table 3 depicts results for a model with collusive possibilities, but with no fringe competition. This corresponds to the model of de Roos (200), and an analysis of the difference between the models presented in columns 1 and 2 can be found there. The most striking difference between the results presented there, and those obtained for the parameters chosen here is the greater incidence of punishment. This arises because firms are more often capacity constrained with the current set of parameters.[25] Firms will prefer to break out of the collusive regime if they are capacity constrained because profits are similar across regimes and their rivals tend to invest more heavily in the collusive regime.

5.2 The Complete Model

5.2.1 Collusion States

The states in the non-cooperative regime in which incumbent firms achieve consensus to collude is shown by the interior of the contour plot in Figure 2. The axes index the capacities of three firms. A value of $k > 0$ corresponds to a capacity of τg^k, while a value of $k = 0$ indicates the absence of a firm. That is, shaded faces of the cube describe states in which there are two active firms who wish to collude, and the interior of the cube depicts states in which there are three active firms who elect to collude.

Firms decide to collude in three types of situation.[26] First, when there are 3 firms of at

[25] The addition of stochastic punishment costs also contributes marginally.

[26] Firms collude in similar environments in the model without fringe competition (not shown). For a discussion,

Table 3: Industry Statistics for Three Model Specifications

Industry Feature	No Collusion	No Fringe Regime	Full Model
Periods with 0 firms	0.0%	0.0%	0.0%
Periods with 1 firm	0.0%	0.0%	0.0%
Periods with 2 firms	30.20%	0.58%	0.55%
Periods with 3 firms	69.80%	99.41%	99.44%
Periods with entry	2.59%	0.52%	0.51%
Periods with exit	2.60%	0.51%	0.50%
Periods in fringe regime	–	–	0.27%
Periods in collusive regime	–	36.13%	36.39%
Periods in non-cooperative regime	–	37.65%	38.67%
Periods in punishment regime	–	26.22%	24.68%
Mean market price	26.98	32.95	32.97
	(4.700)	(7.175)	(7.138)
Mean investment per incumbent	4,479.4	4,947.6	4,958.0
	(2,611.2)	(2,419.4)	(2,420.6)
Mean firm production	432.1	338.2	338.0
	(126.9)	(85.64)	(85.53)
Mean firm capacity	559.9	442.1	442.0
	(264.1)	(150.2)	(149.7)
Mean one-firm concentration ratio	0.440	0.377	0.376
	(0.075)	(0.056)	(0.056)
Mean consumer surplus	26,759	20,637	20,614
	(5,006.0)	(7,404.8)	(7,370.6)
Mean producer surplus	12,090	11,924	11,787
	(7,764.6)	(6,898.8)	(7,319.4)
Mean total surplus	38,849	32,561	32,401
	(10,658)	(9,571.1)	(9,883.3)
Mean firm value	4,379.6	35,951	85,842
	(39,166)	(86,296)	(93,738)
Mean firm lifespan	104.8	579.8	692.3
	(372.0)	(735.9)	(792.7)

Standard errors of observations are in parentheses ().

least moderate size with equal capacities, collusion becomes attractive. This is shown by the diagonal line in the interior of the cube. The reason is that the smallest firm is sufficiently patient to suffer a spell of reduced profits in order to obtain a larger share of the cartel profits in the future. Smaller firms will not collude even if they have equal capacities because their capacity constraints will bind and prevent them from enjoying the fruits of collusion. Second, two small firms may decide to collude, even if they are asymmetric. This is because they anticipate the benefits of collusive profits in the near term, profits are only slightly affected by the entry of a fringe competitor in the short term, and, should another firm enter the market, the cartel can be easily dissolved in the future without resulting in punishment. Finally, two symmetric firms of moderately large size will wish to collude. These firms can enjoy the benefits of collusion until a fringe competitor enters and becomes unbearably large. The cartel can then break up at any time without recourse to punishment. In addition, if firms are of at least moderate size, it becomes prudent for the smallest firm to wait until they have a capacity equal to their competitor before colluding. Interestingly, if there are two very large firms in the market, they are less likely to collude, even if they have equal capacities. This occurs for two principal reasons. First, entry is much more likely if firms are colluding. Second, a fringe firm invests more aggressively than a firm of equal size in the non-cooperative regime. Hence, incumbents may prefer to forego the higher short term profits accruing to the cartel with a competitive fringe, and instead behave competitively to discourage the entry and rapid expansion of a competitor.

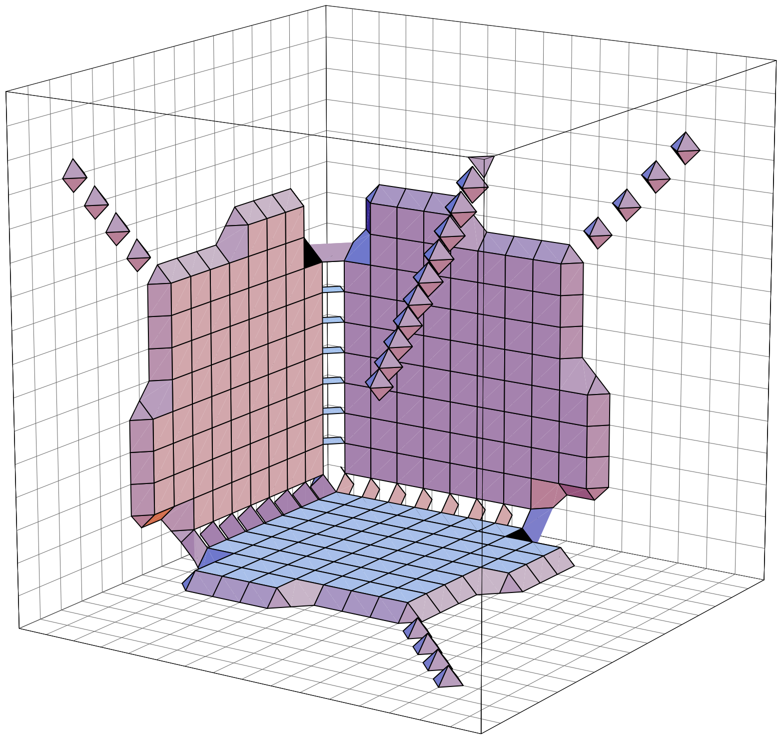
5.2.2 Dissolving the Cartel

Figure 3 depicts states in the fringe regime in which the cartel is upheld. As before, the axes index the capacities of three firms. The vertical axis corresponds to the fringe firm and the remaining axes to the cartel members. Figure 4 shows the interior of the object in Figure 3. The figures are drawn for states in which cartel members have equal collusive market shares. The cartel is unlikely to be maintained in the face of fringe competition when cartel members have unequal market shares. This is because a sufficiently patient cartel member with a low market share can break out of the cartel without instigating punishment. It can then build up capacity while the fringe firm builds up capacity, and then collude when the firms are symmetric.

A cartel with symmetric market shares will remain steadfast in the face of fringe competition while the fringe competitor is relatively small and incumbent firms are large enough to be unconstrained by capacity in the absence of collusion. Should a fringe competitor build up a large capacity, cartel profits will drop below the competitive level and cartel members will elect to dissolve the cartel. If cartel members would be constrained by capacity in the absence of

refer to de Roos (2004).

Figure 2: Collusion States

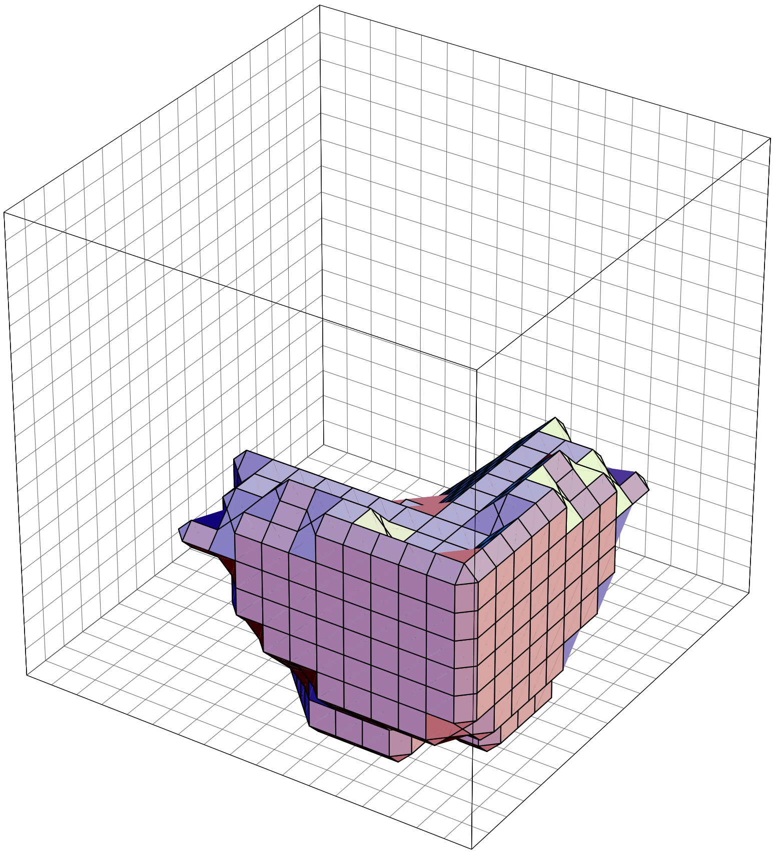


collusion, then there is also no point in maintaining the cartel. It is preferable to break up the cartel, sit out a non-cooperative spell and then collude later once the fringe firm has built up an equal capacity.

As the fringe firm grows, the range of cartel member capacity states in which the cartel will prevail first expands, and then contracts. The cartel may break down if the fringe firm is about to exit and cartel members have large and approximately equal capacities. This is best seen in Figure 4. In this situation, cartel members can collude again once they have equal market shares. The cartel may dissolve if the incumbents are very large and the fringe firm is very small, even if exit is not imminent. Breaking up the cartel dramatically discourages the fringe firm from investing in additional capacity and could result in the fringe firm exiting in the near future.

Once the fringe firm grows beyond a certain point, cartel profits get squeezed and the cartel becomes more difficult to sustain. Interestingly, the cartel is maintained for some states in which cartel profits are below the non-cooperative level, provided that at least one of the incumbents is very large. This arises from the difference in investment patterns across the fringe and non-cooperative regimes. If the fringe regime is maintained, a cartel covering all firms is likely to form sooner. This becomes less important if both cartel members are not overly large.

Figure 3: Fringe States, Exterior

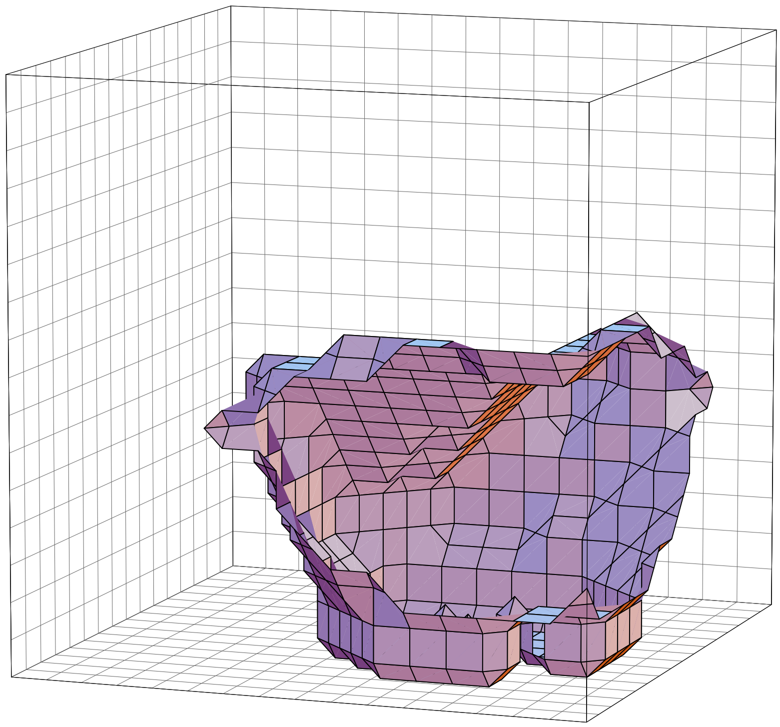


5.2.3 Punishment, Entry, and Investment

Firms will resort to punishment whenever their collusive shares are asymmetric. Even with symmetric collusive shares, the punishment regime will be invoked in several types of environment. If there are three firms in the market and at least one of the firms is very small, that firm will obtain little additional profit by sustaining collusion. However, collusion encourages the investment of rivals, so a small firm will seek to terminate a cartel. While there are only two firms in the market, collusion is difficult to sustain. If the firms are both small, again little profit is to be gained through collusion. While the incumbent firms are small, entry is guaranteed. By entering the punishment regime, incumbent firms can hamper the rapid expansion of a competitor. If the firms are both very large, they can discourage entry through the punishment regime.[27]

[27] Some readers may baulk at the credibility of a strategy that invokes punishment in order to deter entry. In order to address this, a variant of the model was entertained, in which entry breaks up the punishment regime, resulting in the non-cooperative regime. However, even the non-cooperative regime is much less preferable to the fringe firm. In addition, this makes entry deterrence less painful for the incumbent firms. The results are very similar for

Figure 4: Fringe States, Interior



On the other hand, firms of an intermediate size will continue colluding even with the prospect of entry because there are collusive profits to enjoy, and entry cannot be discouraged. In the collusive model without fringe competition, firms punish on a similar region of the state space. The exception is that with two firms in the market, firms are more likely to invoke punishment. This results from the different treatment of entry. In the model without fringe competition, entry breaks up the cartel and induces the non-cooperative regime. Thus, in this model, there is a greater desire to discourage entry by invoking the punishment regime.

With the parameters specified, entry is very attractive in each of the competitive regimes if there are less than three firms in the market. As we might expect, for a given vector of incumbent capacities, an entrant is more likely to be drawn in to the market while the collusive regime is operating. Entry is least attractive while the punishment regime is in force.

There are several patterns to investment behaviour. In all regimes, particularly small firms may undertake little in the way of investment spending if their rivals are large firms. Such firms have given up the prospect of a long life in the industry. Otherwise, investment tends to be higher for smaller firms. Investment is also higher for small firms and lower for large firms relative to a model without collusive possibilities. This is especially so for the non-cooperative regime.

this variation of the main model.

This reflects the desire for firms to speed the onset of collusion. Investment is particularly high for firms who have capacities just below those of their competitors. As we might expect, in the fringe regime, investment is comparatively high for the fringe firm and comparatively low for large cartel members. However, a fringe firm will dramatically reduce investment if it fears its continued expansion would end the cartel.

5.2.4 Industry Simulations

Column 3 of Table 3 presents statistics arising from industry simulations for the full model. Again, the industry is simulated for $1,000,000$ periods, starting in the non-cooperative regime at an initial state vector of $(4,0,0)$. With the entry cost parameters chosen, 3 firms operate for the vast majority of simulation periods. As we might expect, there tends to be a greater number of firms operating than in the model without collusive possibilities. This is also reflected in the lower average one-firm concentration ratios for the industries permitting collusion. The pattern of investment is different in the collusive industries presented in columns 2 and 3 compared to the non-collusive industry of column 1. Investment tends to be more vigorous for small firms, but low for large firms relative to an industry without collusion. Consequently, each firm tends to produce less and hold a lower capacity in a collusive industry, but record a higher average level of investment.

In the complete model, firms operate in the fringe regime for only a small fraction of periods.[28] Several factors contribute to this. First, with two firms in the market, incumbents might enter the punishment regime to deter entry or slow the advance of a future entrant. For small, capacity constrained firms, there is also no profit motive for maintaining collusion. Second, fringe firms will often enter the market when incumbent firms are small and therefore have little interest in maintaining collusion. Finally, three firms tend to survive for long stretches of time without exit and entry and a fourth firm cannot enter due to computational limitations. Hence, there is often no opportunity for a fringe firm to enter the market.

Intuitively, consumer surplus is higher and market prices are lower on average if collusion is not permitted. Producer surplus is approximately the same if we allow collusion, reflecting the greater average number of firms and greater industry-wide investment. Firms tend to be longer lasting and enjoy a higher expected discounted value of operations if collusion is possible. This is primarily due to the restriction on entry of a fourth firm.

Industry statistics look very similar for both of the models with collusion. This is largely

[28]A simulation period is only characterised as falling within the fringe regime if the fringe regime persists through all stages of play for that period. For example, if entry occurs while two firms are colluding in period t, then the cartel dissolves in period $t+1$, period t is attributed to the collusive regime, and period $t+1$ to the non-cooperative regime.

because, with the parameters chosen, the fringe regime operates for only a small fraction of periods. There is a striking difference in the average realised firm lifetime operating value. A recent entrant will anticipate a substantially greater discounted stream of profits in the model incorporating fringe competition. In the model with fringe competition, a firm entering the market while a cartel is operating will invest heavily and quickly grow while being generously accommodated by the cartel. In contrast, the model without fringe competition assumes the cartel is broken up upon entry. Thus, an entrant can anticipate a much more favourable beginning to its operating life. If further entry were possible, much of this disparity in expected firm values might disappear.

In Table 4, industry statistics are broken down by competitive regime. Columns 1 to 4 describe the industry conditional on the operation of the Fringe, Collusive, Non-cooperative, and Punishment regimes, respectively. Unsurprisingly, the collusive and fringe regimes produce the highest average prices and lowest average consumer surplus. Producer surplus is also greater in the collusive and fringe regimes, particularly the collusive regime. Because firms will tend to collude only when they are larger, the collusive regime is characterised by greater average firm capacity, but substantially lower average investment per firm compared to the other regimes. Firms will hold substantial excess capacity only in the collusive and fringe regimes. Finally, in the fringe regime, fringe firms invest substantially more than cartel members because their profits are particularly sensitive to their capacity.

5.3 The Vitamin C Market

The above model is capable of explaining the pattern of firm behaviour observed in the vitamin C market. Figures 3 and 4 show that, for a given set of parameters, the environments in which a cartel will tolerate fringe competition can be delineated. However, without significant modification, it is unable to match the details of the vitamin C market. Many stylised assumptions are made and computational limitations imposed that strongly restrict the predictive power of the model. First, a maximum of 3 firms can operate in the market due to the computational burden of the algorithm. Second, firms are assumed to engage in quantity competition for a homogeneous product whereas firms in the market for vitamin C set prices and enjoyed a minor degree of product differentiation. Third, all firms are assumed to possess a low investment efficiency. In practice, we might expect the investment efficiency of an incumbent firm to be higher than that of a recent entrant, particularly if we think of capacity in a general sense as incorporating all things that impact on the short term ability of a firm to produce and sell its product. A consequence of this assumption is that firms are often capacity constrained in simulations. Fourth, we might expect the rate of growth of a fringe competitor to depend on the relationship between

Table 4: Industry Statistics for Different Regimes

Industry Feature	Fringe	Collusive	Non-cooperative	Punishment
Periods with 0 firms	0.0%	0.0%	0.0%	0.0%
Periods with 1 firm	0.0%	0.0%	0.02%	0.0%
Periods with 2 firms	0.0%	1.07%	0.05%	0.57%
Periods with 3 firms	100.0%	98.93%	99.93%	99.42%
Periods with entry	–	1.03%	0.06%	0.48%
Periods with exit	0.78%	1.03%	0.06%	0.44%
Mean investment per incumbent	5,640.2	3,645.4	5,822.6	5,526.6
	(2,674.0)	(1,721.0)	(2,623.1)	(2,120.0)
Mean fringe firm investment	6,873.4	–	–	–
	(1,746.6)			
Mean cartel firm investment	5,023.5	–	–	–
	(2,838.9)			
Mean firm production	306.1	288.5	363.2	372.1
	(87.31)	(10.56)	(100.0)	(87.96)
Mean firm capacity	422.8	529.0	387.3	400.0
	(228.9)	(134.3)	(136.5)	(129.1)
Mean market price	36.63	38.80	29.96	29.00
	(4.768)	(1.124)	(7.176)	(6.628)
Mean one-firm concentration	0.409	0.336	0.404	0.391
	(0.052)	(0.027)	(0.056)	(0.050)
Mean consumer surplus	16,717	14,511	23,776	24,747
	(3,937.5)	(755.1)	(7,325.9)	(6,895.5)
Mean producer surplus	13,456	17,730	8,378.2	8,832.4
	(7,169.4)	(7,034.3)	(3,641.4)	(4,483.1)
Mean total surplus	30,173	32,241	32,155	33,579
	(10,856)	(7,599.4)	(10,470)	(10,453)

Standard errors of observations are in parentheses ().

fringe and cartel prices, but this feature is not incorporated in the model.

As a result of these limitations, the simulated industry statistics differ from observations in the vitamin C market in several respects. First, the behaviour of prices differs. The market prices generated are substantially higher than those observed in the vitamin C market. The assumption of quantity competition together with the exogenous limitation on the maximum permissable number of firms and the low investment efficiency all contribute to this result. Prices also jump when regimes switch. In the vitamin C market, price movements are smoother because of contracts and efforts by cartel members to avoid detection by anti-trust authorities. Second, persistent asymmetries in firm sizes are not explained. Such asymmetries could be accounted for by incorporating asymmetries in costs or the effectiveness of investment. Third, the fringe regime is only rarely observed. This results because the incumbent firms are typically too small to find maintaining a cartel attractive in the face of fringe competition. If investment is made more effective, incumbents will tend to be larger, but fringe firms will grow rapidly. Consequently, a long lasting fringe regime is unlikely. This feature could be incorporated fairly simply into the model by specifying a different investment effectiveness for incumbents and recent entrants. An exogenous random transition, of the same nature as the transition from the punishment regime to the non-cooperative regime, could capture this with little additional computational burden.

An interesting test of the model is its ability to explain the difference in experiences in the lysine market, studied in de Roos (2004), and the vitamin C market. In the lysine market, the entry of Archer Daniels Midland (ADM) triggered an immediate, severe price war. The price war was only resolved and collusion initiated after ADM had built up a market share comparable with its competitors. By contrast, in the vitamin C market, fringe competition from the Chinese firms was tolerated for at least four years before the cartel was dissolved. One possible explanation consistent with the current model is that in the lysine market, the investment process was much more effective, particularly for the entrant. That is, ADM was much more adept at building up its distribution, and marketing network, and expanding its physical production capacity in the lysine market than were the Chinese firms in the market for vitamin C.

6 Conclusion

Most of the literature on collusion with a competitive fringe has applied a static or repeated games framework. By framing the cartel's problem within an explicitly dynamic model, this paper obtains several intuitive results. First, a cartel will tend to tolerate a fringe competitor only while it remains comparatively small. Second, a fringe competitor will tend to invest heavily while a cartel is operating. Third, entry deterrence is mitigated if cartel members have the option of maintaining the cartel following the entry of a competitor. Fourth, firms of an intermediate size are the most likely to accommodate entry. The paper is strongly motivated by events in the market for vitamin C where competition from a fringe firm eventually forced a cartel to break down. The qualitative features of the vitamin C market can be simply explained by the model, and the model could provide guidance for firm behaviour in any market in which collusion and entry are possible features and firms face some information asymmetries.

References

[1] Bernheim, B. Douglas, and Michael D. Whinston (1990), "Multimarket Contact and Collusive Behavior", *RAND Journal of Economics*, 21(1), pp. 1-26.

[2] Bernheim, B. Douglas; Peleg, Bezalel; and Michael D. Whinston (1987), "Coalition-Proof Nash Equilibria: Concepts", *Journal of Economic Theory*, 42(1), pp. 1-12.

[3] de Roos, Nicolas (2004), "A Model of Collusion Timing", *International Journal of Industrial Organization*, 22(3), pp. 351-387.

[4] Deaton, A. and J. Muellbauer (1981), "An Almost Ideal Demand System", *American Economic Review*, 70, pp. 312-326.

[5] Fershtman, C. and A. Pakes (2000), "A Dynamic Game with Collusion and Price Wars", *RAND Journal of Economics*, 31(2), pp. 207-236.

[6] Genesove, David and Wallace Mullin (2000), "Rules, Communication and Collusion: Narrative Evidence from the Sugar Institute Case", NBER Working Paper No. W8145.

[7] Hausman, Jerry A. (1999), "The Competitive Effects of a New Product Introduction: A Case Study", MIT Working Paper.

[8] Kreps, D. and R. Wilson (1982), "Reputation and Imperfect Information", *Journal of Economic Theory*, 27, pp. 253-279.

[9] Levenstein, Margaret (1995), "Mass Production Conquers the Pool: Firm Organization and the Nature of Competition in the Nineteenth Century", *The Journal of Economic History*, 55(3), pp. 575-611.

[10] Milgrom, P. and J. Roberts (1982), "Predation, Reputation and Entry Deterrence", *Journal of Economic Theory*, 27, pp. 280-312.

[11] Posada, Pedro (2001), "Leadership Cartels in Industries with Differentiated Products", University of Warwick Economic Research Paper No. 590.

[12] Selten, Reinhard (1973), "A Simple Model of Imperfect Competition, Where 4 Are Few and 6 Are Many", *International Journal of Game Theory*, 2, pp. 141-201.

[13] Shaffer, Sherrill (1995), "Stable Cartels with a Cournot Fringe", *Southern Economic Journal*, 61, pp. 744-754.