IDENTIFICATION AND ESTIMATION OF FIRMS' MARGINAL COST FUNCTIONS WITH INCOMPLETE KNOWLEDGE OF STRATEGIC BEHAVIOR

Adam Rosen

THE INSTITUTE FOR FISCAL STUDIES
DEPARTMENT OF ECONOMICS, UCL
cemmap working paper CWP03/07

Identification and Estimation of Firms' Marginal Cost Functions With Incomplete Knowledge of Strategic Behavior

Adam Rosen[*][†]
University College London
Centre for Microdata Methods and Practice
Institute for Fiscal Studies

First Version: September 2005
This Version: December 2006

Abstract

In this paper I develop a new approach for identification and estimation of the parameters of an oligopoly model, without relying on a potentially unverifiable equilibrium assumption. Rather, I consider inference on model parameters when the researcher does not know precisely what decision rule firms use, but is willing to consider a set of possibilities. In contrast to traditional approaches in the literature, the proposed methodology allows firm behavior to vary flexibly across observations, in a manner consistent with many Nash Equilibria. I derive identification results for both homogeneous product and differentiated product markets. Due to the flexibility afforded to firm behavior, the parameters of firms' marginal cost functions may only be set identified rather than point identified. The restrictions of the model are, however, still informative. I find that the size of the identified set for marginal cost parameters depends on the elasticity of market demand, the set of decision rules considered, and the functional form assumptions imposed. I formulate how to compute consistent set estimates for marginal cost parameters and demonstrate the proposed methodology with price and quantity data on the Joint Executive Committee, a 19th century railway cartel. To perform statistical inference I implement the methodology of Rosen (2005) to construct asymptotically valid confidence regions for the partially identified marginal cost parameters. The application illustrates how

[*]I thank Allan Collard-Wexler, Joel Horowitz, Nir Jaimovich, Mike Mazzeo, Bill Rogerson, Itai Sher, Andrew Sweeting, Mike Whinston, and Fan Zhang for comments and suggestions on this paper. I am especially grateful to Chuck Manski, Aviv Nevo, Rob Porter, and Elie Tamer for continued discussion and feedback. Furthermore, I thank Rob Porter for supplying me with data on the Joint Executive Committee. I have benefitted from the comments of seminar participants at LSE, Sloan, UCL, UC Irvine, the 2005 Econometric Society World Congress and the 2006 Southern Economic Association Meetings. Financial support from the Robert Eisner Memorial Fellowship and the Center for the Study of Industrial Organization are gratefully acknowledged. All errors are my own.

[†]Mailing address: Department of Economics, University College London, Gower Street, London WC1E 6BT, United Kingdom. Email: adam.rosen@ucl.ac.uk.

the precision of estimated marginal costs depends on the elasticity of market demand as well as the extent to which firm behavior is allowed to vary.

1 Introduction

Structural models of oligopoly markets consist of three components: (i) consumers' demand, (ii) firms' cost or production functions, and (iii) an equilibrium assumption. Reliable data on marginal costs are often not available, so marginal costs must be inferred from data on prices and quantities. When combined with estimates for demand parameters, either estimated in a first stage or jointly with the imposed equilibrium conditions, the equilibrium assumption serves to pin down price-cost margins that can be used to back out marginal costs. Since Rosse (1970) this has been standard practice in empirical work in industrial organization. It is well recognized, however, that the implied marginal costs, and in turn the parameters of firms' marginal cost functions, may then be quite sensitive to the researcher's choice of an equilibrium assumption.

In this paper, I relax the equilibrium assumption, recognizing that it may not be known to the researcher which equilibrium firms play, or even if they play equilibrium at all. In lieu of an equilibrium assumption, I impose weak restrictions on firms' decision rules that are consistent with many standard equilibrium assumptions, and I examine the implications of these restrictions for identification and estimation of model parameters. This methodology allows empirical researchers to proceed with estimation when they are uncertain of the precise mechanism by which firms make their price or quantity decisions, but are willing to impose a more conservative restriction on their behavior. In particular, I investigate the consequence of imposing restrictions directly on firms' price-cost margins that are consistent with many of the equilibrium assumptions that have been used in the literature, and many sophisticated models of firm behavior.

Even if the researcher is willing to maintain that the firms play an equilibrium, there may be many different plausible equilibria to choose from. In fact, when there is dynamic interaction among firms over time, microeconomic theory reveals that a multiplicity of equilibria is the norm rather than the exception.[1] Rather than assume that observations correspond to a narrow subset of these equilibria in order to achieve point identification, the methodology of this paper allows the researcher to instead base inference on a set of feasible market outcomes consistent with a large class of Nash equilibria. The proposed methodology is flexible so that the exact set of equilibria or decision rules considered can be adapted to those thought most salient for the market of interest. Estimation of structural parameters is then based on the assumption that observed firm behavior belongs to this set, rather than a possibly unverifiable equilibrium selection assumption. The set of feasible firm decision rules considered can be thought of as implicitly mapping a set of equilibrium selection rules to implied values for model parameters.

[1] See, for example, the discussion in chapter 12, pp. 404-405 of Mas-Colell, Whinston, and Green (1995).

The cost of the flexibility afforded to firm behavior is that the model is generally incomplete. As a result, parameters of firms' marginal cost functions are, except in special cases, only set identified rather than point identified. Even though the marginal cost parameters are not point identified, the data will still be informative when the assumptions of the model are imposed. Exactly how informative depends on the stringency of the imposed behavioral assumptions, functional form assumptions, and in particular on the elasticity of market demand. This is borne out in the identification results of section 4, and illustrated in the empirical application of section 7, which uses a well-known dataset that includes prices and shipment quantities for a late 19th century railway cartel, the Joint Executive Committee.

Despite a possible lack of point identification for marginal cost parameters, this approach has many advantages. First, it tells the researcher what can be inferred from relatively weak restrictions on firm behavior. A clear benefit of this approach is that it allows the researcher to determine the extent to which a more restrictive model achieves identification through its assumptions on equilibrium behavior. This then enables the empirical researcher to determine how robust their estimates are to various equilibrium assumptions. In some cases, it may not be clear how firms set prices or quantities, and different researchers may disagree as to how firms make their strategic decisions or which equilibrium they play. By imposing conservative assumptions on firm behavior that embed different decision rules, researchers who disagree on the precise mechanism by which price cost margins are actually determined can deduce to what extent their disagreement may effect implied parameter values.

The paper proceeds as follows. In the remainder of the introduction, I discuss the related literature. In section 2, I develop the homogeneous products version of the model and place it in context relative to previous models in the literature. In particular, I formalize the assumptions imposed on firms' decision rules that replace the standard equilibrium assumptions. Section 3 extends the model to a differentiated products market. In section 4, I derive identification results for both homogeneous product markets and differentiated product markets. In each case, under standard instrumental variable assumptions, the parameters of the demand side of the market are point identified, while cost parameters are set-identified. The identification results in either case show that the size of the identified set for cost parameters depends on the market demand elasticity(ies). Section 5 discusses estimation. Section 6 provides a simple example of how the methodology could be applied to a homogeneous product market with a linear inverse demand and linear marginal cost specification. Section 7 presents an application to data on prices and quantities shipped by the Joint Executive Committee, a railway cartel, and details the new estimation method employed. The application illustrates how the size of the identified set for marginal cost parameters, and therefore the precision of inference that can be drawn, depends on the assumptions brought to bear as well as the elasticity of market demand. Section 8 concludes and discusses avenues for future research.

1.1 Related Literature

Recent results concerning identification of model parameters in this setting build on the recent econometrics literature on partial identification, covered in depth by Manski (2003). Of the papers cited there, the identification results of this paper are most closely related to Manski and Tamer (2002). However, while Manski and Tamer (2002) focus on single equation estimation when an outcome variable or regressor is interval measured and all regressors are exogenous, the econometric model of this paper is composed of simultaneous equations with cross-equation restrictions and endogenous regressors, which is concisely represented by a set of conditional moment inequalities. This requires some new techniques for estimation and statistical inference. Some possibilities include the inferential approaches of Andrews, Berry, and Jia (2004), Chernozhukov, Hong, and Tamer (2004), Shaikh (2005), Beresteanu and Molinari (2006), Pakes, Porter, Ho, and Ishii (2006), and Rosen (2005). For the empirical example of this paper, I employ the estimator of Rosen (2005), which is generally applicable for models based on a finite number of moment inequalities. The estimator takes the form of a minimum chi-square statistic, which has appeared previously in the statistics literature on multivariate one-sided hypothesis tests.

Research on incomplete econometric models in empirical industrial organization, particularly when the model admits multiple equilibria, is also a recent and rapidly expanding literature. Examples that deal with multiple equilibria in models of entry and exit include Andrews, Berry, and Jia (2004), Tamer (2003), and Ciliberto and Tamer (2004). In the context of an English auction, Haile and Tamer (2003) use an incomplete model that nests many possible equilibria to estimate bounds on the distribution of bidders' valuations. Additional research that seeks to estimate model parameters in the presence of multiple equilibria includes Borzekowski and Cohen (2005) and Sweeting (2005). To my knowledge, in the oligopoly setting I consider, estimation strategies based on only incomplete knowledge of firms' pricing and or quantity decisions have not been considered.[2]

Further discussion of related research in the empirical study of oligopoly models is deferred to the next section, where the model is formally introduced and placed in further context.

2 A Model For Homogeneous Product Markets

The setting I consider in this section is one in which a researcher observes product price and firms' quantities produced, as well as demand and cost covariates in a homogeneous product market for a large number T periods. In each period of observation, $t = 1, ..., T$, the market is assumed to be comprised of the same n firms.[3] The econometrician has observations of market price p_t and

[2] There is also a growing literature on estimation of dynamic models of oligopoly. To this point, however, this research abstracts away from dynamic pricing and/or quantity decisions (for example through tacit collusion), while instead allowing for dynamics in other decision variables, such as investment or entry/exit decisions. Examples from the literature include Bajari, Benkard, and Levin (forthcoming), and Pakes, Ostrovsky, and Berry (2005).

[3] This assumption can be relaxed, but allows for clearer exposition.

quantities $\mathbf{q}_t = (q_{1t}, ..., q_{nt})$. Let x_t, w_t be (vector-valued) perfectly observed demand and cost covariates, respectively, at time t. These may in general include market specific components and may include some common components across firms as well.

Demand in each market is assumed to be determined by the parametric inverse demand function

$$p_t = P(Q_t, x_t, \alpha) + u_t, \tag{2.1}$$

Where u_t is a mean zero demand shifter such that $\mathbb{E}[u_t|w_t, x_t] = 0$, Q_t is aggregate time t output, α are parameters to be estimated, and $P(\cdot, x, \alpha)$ is downward sloping for all x. Each firm $i = 1, ..., n$ in period t has marginal costs

$$mc_{it} = mc(q_{it}, w_t, \beta_i) + \epsilon_{it}, \tag{2.2}$$

where ϵ_{it} is an additive mean zero cost-shifter such that $\mathbb{E}[\epsilon_{it}|w_t, x_t] = 0$, and q_{it} is the firm's period t quantity produced. β_i are the parameters of firm i's marginal cost function to be estimated, and w_t are covariates for observation t. $\beta \equiv (\beta_1', ..., \beta_n')'$ is used to denote the composite vector of all the firms' marginal cost parameters. This does not exclude the possibility that there are some common parameters for firms' marginal cost functions, i.e. that for $i \neq j$, β_i and β_j share some common components.

To this point, the imposed assumptions are standard in the empirical IO literature. Empirical models of oligopoly markets always include a specification of market supply and demand. What ties the two together is the specification of an equilibrium assumption. Given the parametric specification above, different equilibrium firm strategies are generally derived from conditions for profit maximization. However, these profit maximizing conditions are sensitive to the particular game firms are assumed to play, the information structure, and the type of equilibrium assumed. For example, perfectly competitive, Cournot, and Stackelberg equilibria all lead to different conditions for profit maximization, and thus different structural equations from which to derive parameter estimates. Imposing one of these equilibria then yields a profit maximizing condition of the form

$$p_t = mc_{it} - \theta_{it} Q_t \frac{\partial P_t(Q_t, x_t, \alpha)}{\partial Q}, \tag{2.3}$$

for each firm. It is important to note that this equation is a tautology, as it is completely unrestrictive without specifying restrictions for θ_{it}; given values for p_t, mc_{it}, and a non-zero value for $Q_t \frac{\partial P_t(Q_t, x_t, \alpha)}{\partial Q}$, there will always exist some θ_{it} that makes this equation true.[4] However, this condition gives a concise way of embedding different static equilibrium assumptions, as different equilibria correspond to different values of θ_{it}, and thus different values for firms' markups. For example, Bertrand, Cournot, and monopoly equilibrium pricing correspond to the restrictions $\theta_{it} = 0$, $\theta_{it} = s_{it}$, and $\theta_{it} = 1$, respectively, where $s_{it} \equiv \frac{q_{it}}{Q_t}$, firm i's period t market share. Many empirical

[4] The only case where it may not seem possible to satisfy equation (2.3) is when $\frac{\partial P_t(Q_t, x_t, \alpha)}{\partial Q} = 0$. In that case, however, firms are necessarily price-takers so that price equals marginal cost, and any value of θ_{it} will suffice.

applications also use a market level version of condition (2.3) such as

$$p_t = \overline{mc}_t - \theta_t Q_t \frac{\partial P_t\left(Q_t, x_t, \alpha\right)}{\partial Q}, \tag{2.4}$$

where $\overline{mc}_t = \sum_{i=1}^{n} s_{it} \cdot mc_{it}$ and $\theta_t = \sum_{i=1}^{n} s_{it}\theta_{it}$. This market-level version of (2.3) is derived by summing over i, and weighting each equation by s_{it}, and it also embeds Bertrand, Cournot, and monopoly pricing as special cases.

Estimation strategies that invoke an equilibrium assumption have identifying power through the restrictions imposed on θ_{it}. One restriction in the empirical literature is to specify a particular static equilibrium for all observations, e.g. marginal cost pricing where $\theta_{it} = 0$ for all i, t. Another is to consider a finite number of static equilibria, e.g. competitive, Cournot, and joint profit-maximization, and then select the one that best fits the data through a formal testing procedure, e.g. Bresnahan (1987) and Gasmi, Laffont, and Vuong (1992). Yet another equilibrium restriction is the so-called conduct parameter method (CPM), which imposes constancy of θ_{it} or θ_t across time, but treats it as a parameter to be estimated. This approach is more general than the first two, because although θ_{it} is assumed constant, it's value is not fixed a priori or limited to only a small number of possibilities. The next section discusses the limitations of imposing constancy of θ_{it} (or θ_t) in order to motivate the approach developed in this paper. Section 2.2 then describes how one can replace the standard type of equilibrium assumptions with a more general restriction that allows θ_{it}, and thus firms' markups, to vary flexibly across time.

2.1 Motivation and Discussion of Restrictions on θ_{it}

The conduct parameter method embeds many different static equilibria by combining (2.3), or the market-level analog (2.4), with the restriction that θ_{it} (or θ_t, respectively) is constant across periods. Under the assumption that $\theta_{it} = \theta_i$ for all t, it follows that under particular functional form assumptions and sufficient exclusion restrictions θ_i is identified.[5] In this case, the conduct parameter θ_i can then be consistently estimated jointly with the parameters of firm i's marginal cost function. Bresnahan (1989) gives a thorough discussion of this approach and surveys its use in applied work. Variations of the CPM used in empirical research include Iwata (1974), Rubinovitz (1993), Graddy (1995), Genesove and Mullin (1998), and Wolfram (1999), to name a few.

Solving for θ_{it},(2.3) can be rewritten

$$\theta_{it} = \left(mc_{it} - p_t\right) \cdot \left(Q_t \frac{\partial P_t\left(Q_t, x_t, \alpha\right)}{\partial Q}\right)^{-1}, \tag{2.5}$$

[5] Bresnahan (1982) and Lau (1982) derive necessary and sufficient conditions for identification of a market level conduct parameter.

or equivalently,

$$\begin{aligned}\theta_{it} &= \frac{(mc_{it} - p_t)}{p_t} \cdot \eta_t \\ &= L_{it} \cdot \eta_t,\end{aligned}$$

where $\eta_t = \frac{p_t}{Q_t} \cdot \frac{\partial Q(p_t, x_t, \alpha)}{\partial p}$, the elasticity of aggregate demand, and L_{it} is the Lerner index, which is firm i's price-cost margin normalized by price. Generally speaking, there is no reason for this elasticity-adjusted markup to be constant across periods, unless the firms are playing precisely the same static equilibrium every period. Furthermore, because θ_{it} is a function of marginal cost, price, and the elasticity of demand, it is potentially correlated with any demand or cost shifters; an advantage of the incomplete model of this paper is that it is robust to this possibility.

While the CPM was in part intended to allow for the possibility of tacit collusion, Corts (1999) demonstrates that it may yield inconsistent parameter estimates when firms play a tacitly collusive dynamic equilibrium. If firms are actually playing a repeated game in which they maintain tacitly collusive behavior by playing a more favorable SPNE than repeated static Nash Equilibrium, the estimated conduct parameter may be an altogether uninformative measure of firm cooperation. The underlying issue is precisely that across periods the conduct parameter θ_{it} need not remain constant, and is in fact correlated with covariates. In Corts' framework, the fluctuations in firm conduct are due to fluctuations in market conditions, i.e. changes in demand or cost covariates that alter the highest incentive compatible profit levels. Other explanations for fluctuations in the conduct parameter, such as equilibrium selection from a set of possible SPNE, or the existence of a competitive fringe, e.g. Salvo (2005), could be posited as well. Through simulation of a duopoly market for a homogeneous product in a repeated Cournot game, Corts demonstrates that the estimated conduct parameter need not be consistent for even the average level of firm conduct, and may in fact be altogether misleading, failing to detect collusive behavior when the average level of firm conduct is actually close to the highest level achievable. This in turn leads to inconsistent estimates of firms' marginal cost parameters. Because the CPM in this case concludes that firms are behaving more competitively than they in fact are, on average, firm margins are underestimated, and marginal costs are overestimated.[6]

Corts' critique has particular bite because the CPM itself will not reveal whether firms' conduct is indeed constant across periods, unless additional a priori assumptions regarding firms' behavior are brought to bear. Corts (1999) verifies formally that the CPM is not well-suited for markets in which firm behavior changes significantly and often over time. Indeed, in the context of the theories of tacit collusion put forth by Green and Porter (1984), Abreu, Pearce, and Stachetti (1986), and Rotemberg and Saloner (1986), Bresnahan (1989, 1024) writes:

[6] There are cases where Corts' critique does not apply. For example, if it is known that firms price at marginal cost with positive probability, and such periods can be identified, then marginal costs can be inferred from those periods. This can in turn be used to infer markups in other periods as well.

> These various theories have in common the idea that in an imperfectly informed world, "successfully" collusive industries will have periods of cartel pricing and periods of competition. In general, they imply models with θ_{it} not necessarily equal to $\theta_{i\tau}$. These theories differ somewhat in the expected time-series behavior of these two regimes, as the exact equations determining passage from one regime to the other vary between theories....It is easy to imagine other theories of success or failure in tacit collusion which predict different patterns; taking all of these theories at once would lead to even more complex potential time series behavior for θ.

In this paper I address this issue by putting a bound directly on firms' markups that is compatible with many equilibrium assumptions that have been previously employed. This results in an incomplete model that is sufficiently rich to express a wide-ranging class of possibilities for time-varying levels of firm collusion. To perform estimation and inference I exploit and build on findings from the recent literature on identification and estimation in partially identified models. Unlike the conduct parameter method, I do not restrict θ_{it} to be constant over time. My goal is not estimation of any sort of conduct parameter, but rather estimation of firms' marginal cost functions. Of course, if marginal costs could be estimated with enough precision, then quite a bit could be inferred concerning firm behavior as well.

In place of an equilibrium assumption, I instead put restrictions directly on firms' price-cost margins. Exactly which restrictions are reasonable will be different for different markets, and the empirical researcher must, as always, take care to choose assumptions that are appropriate. As a baseline, I start by positing that firms make their strategic decisions so that their price-cost margins are positive, but no greater than what could be obtained by a monopolist. While this assumption is consistent with many types of equilibria, even this rather conservative restriction may be violated in some markets. Cases where this assumption would not apply are markets where firms may sometimes price below marginal cost, such as those with predatory pricing or learning by doing, or cases where some firms can garner price-cost margins greater than the monopoly margin. The methodology here can, however, be modified in order to embed those sets of price-cost margins thought most salient for the market being studied. For this paper, the restriction of firms' margins to be between those of perfect competition and monopoly levels will be the most lenient considered. Restrictions that are not nested by this one are left to future research.

2.2 Restrictions on Firms' Markups

Under the assumption that firms have perfect information (i.e. they observe ϵ_{it} and u_t), marginal cost pricing for firm i is given by $p_t = mc_{it}$, while monopoly pricing yields $p_t = mc_{it} - Q_t \frac{\partial P}{\partial Q}(Q_t, x_t, \alpha)$ for each active firm in period t. Thus, restricting price-cost margins to lie between

these two extremes yields

$$\begin{aligned} \forall t &= 1, ..., T,\ i = 1, ..., n \\ 0 &\leq p_t - mc_{it} \leq -Q_t \frac{\partial P}{\partial Q}(Q_t, x_t, \alpha). \end{aligned}$$

This restriction imposes upper and lower bounds on firms' marginal costs in each period of observation, in contrast with conventional equilibrium assumptions that pin down marginal costs to a singleton. This weaker restriction on firms' marginal costs is the reason why marginal cost parameters may only be set identified, rather than point identified.

More generally, I consider restrictions of the form

$$\begin{aligned} \forall t &= 1, ..., T,\ i = 1, ..., n \\ -\theta_{it}^0 Q_t \frac{\partial P}{\partial Q}(Q_t, x_t, \alpha) &\leq p_t - mc_{it} \leq -\theta_{it}^1 Q_t \frac{\partial P}{\partial Q}(Q_t, x_t, \alpha), \end{aligned} \tag{A**}$$

where θ_{it}^0 and θ_{it}^1, either perfectly observed or imposed a priori by the researcher, serve as known lower and upper bounds, respectively, on firm i's elasticity-adjusted Lerner index, or equivalently, on firm i's markup as measured as a portion of the monopoly markup. The perfectly competitive and monopoly bounds are given by $\theta_{it}^0 = 0$, and $\theta_{it}^1 = 1$ for all i, t. Another case would occur if firms were known to be quantity setters, and assumed never to have margins any lower than that achieved by Cournot equilibrium. Then $\theta_{it}^0 = s_{it}$ for all i, t, where s_{it} is firm i's market share.

An equivalent formulation is that

$$\begin{aligned} \forall t &= 1, ..., T, i = 1, ..., n, \exists \theta_{it} \in \left[\theta_{it}^0, \theta_{it}^1\right] \text{ such that} \\ p_t &= mc_{it} - \theta_{it} Q_t \frac{\partial P}{\partial Q}(Q_t, x_t, \alpha), \end{aligned}$$

where θ_{it} captures the value of the true price-cost margin relative to the would-be monopoly margin. As this notation suggests, the price-cost margin bounds can be explicitly derived from the conduct parameter paradigm. However, this approach is inherently different from the CPM because the period t conduct parameter is allowed to take any value between θ_{it}^0 and θ_{it}^1, and no restrictions are placed on the evolution of that parameter over time. There is no "structural" parameter θ_{it} to be estimated. Rather, θ_{it} simply offers one way to parameterize the distance of price from marginal cost, on the scale of $-Q_t \frac{\partial P}{\partial Q}(Q_t, x_t, \alpha)$. For any firm and period, the set of θ_{it} allowed imposes restrictions on firm markups, and the specific values of θ_{it} offer a parameterization of the unobserved level of firm i's period t markup.

With suitable choice of θ_{it}^0 and θ_{it}^1 the restriction (A**) can easily be made less restrictive than many standard equilibrium assumptions. Such restrictions are, however, not automatically satisfied. In particular, for a given period of observation t, $-Q_t \frac{\partial P}{\partial Q}(Q_t, x_t, \alpha)$ does not measure the

monopoly markup of the industry, but rather the implied markup if observation t *were* in fact one in which monopoly power was exercised. For example, if the firms were in fact playing a Cournot equilibrium in period t, then the value of $-Q_t \frac{\partial P}{\partial Q}(Q_t, x_t, \alpha)$ would not be the same as if Q_t were instead at the monopoly level. It is worth stressing that both of these restrictions, and particularly the former, are less stringent than those typically imposed by researchers to infer markups, either through the conduct parameter method or by imposing a static Nash Equilibrium. Nonetheless, care must be taken by the researcher to impose reasonable restrictions on firms' decision rules on a case-by-case basis. For now, I consider the implications of imposing bounds on each firms' elasticity-adjusted markups, but appropriate bounds must be carefully chosen to suit the application at hand.

2.3 Implications From Cross Section and Time Series Data

In this subsection I briefly discuss the differences between the cases of cross section and time series data. First, consider the case of cross section data, where t is instead an index for each distinct market at a fixed point in time. If the markets are truly distinct in the sense that there is no cross-market ownership, then the equilibrium behavior that dictates the outcomes in each market can be treated as independent, conditional on market covariates. That is, it may be reasonable to invoke an assumption such as $\mathbb{P}\{\theta_{it}|w, x\} = \mathbb{P}\{\theta_{i\tau}|w, x\}$ for all $t \neq \tau$. If, instead, some of the same firms are active in different markets, such an assumption may not be plausible. However, it may in this case be reasonable to impose a restriction on firm conduct when the same firms appear in different markets, e.g. $\theta_{it} = \theta_{i\tau}$ for all markets t, τ with the same set of firms present.

This is different from the case of time-series data in a single market. In this case, reasonable patterns of firm behavior across time can be considered. The folk theorem dictates that repeated interactions among firms results in a wealth of SPNE. However, if the researcher is able to credibly impose restrictions on the time-varying behavior of firms, such as those considered by Green and Porter (1984)[7] or Rotemberg and Saloner (1986), then these restrictions could potentially be imposed for greater identifying power.

The key point is that the time series and cross section data have very different implications for plausible restrictions on firm behavior. When the researcher has panel data, both types of assumptions on firm behavior could potentially be imposed, with those of the first type being applied across markets at each point in time, and the latter being applied in each market (or for each set of firms in a single market) across time. For now I simply restrict attention to firm behavior resulting in a wide range of static outcomes.

[7] The Green and Porter framework is one in which firms have incomplete information with respect to demand, while I focus attention on situations in which firms have perfect information. The current methodology could, at least in theory and with some care, also be extended to situations where firms have incomplete information as well.

3 A Model for Differentiated Product Markets

Now I turn to the more complex case of differentiated product markets. For this case, I assume throughout that each firm's marginal costs are constant with respect to quantity, but may differ with period t covariates. As a starting point, I consider the differentiated product oligopoly framework of Nevo (1998) with the added simplification that each firm in the market produces one and only one differentiated product. There are J differentiated products in the industry, and thus J firms, and the demand for each product at a given point in time in a single market is

$$Q_{jt} = Q_j\left(\mathbf{p}_t, x_t, \alpha\right) + u_{jt}, \tag{3.1}$$

where $\mathbf{p}_t = (p_{1t}, ..., p_{jt})$ is the price vector of differentiated products, x_t are covariates, α are parameters to be estimated, and u_{jt} is a demand shifter for product j that is observed by the firms prior to making their decisions, but is not observed by the econometrician. The notation $Q\left(\mathbf{p}, x, \alpha\right)$ denotes the vector-valued function $\left(Q_1\left(\mathbf{p}_t, x_t, \alpha\right), ..., Q_J\left(\mathbf{p}_t, x_t, \alpha\right)\right)'$. The variable profit (i.e. not counting fixed costs) of firm j in period t is thus

$$\pi_{jt} = \left(p_{jt} - mc_{jt}\right) \cdot Q_{jt}, \tag{3.2}$$

where $mc_{jt} = mc\left(w_t, \beta_j\right) + \epsilon_{jt}$ is a function of the exogenous cost shifters w_t, the parameters β_j, and the cost shifter ϵ_{jt}. Like u_{jt}, each ϵ_{jt} is assumed to be observed by the firms prior to making their decisions, but unobserved by the econometrician. A Bertrand-Nash equilibrium in prices is then characterized by the equations

$$Q_{jt} + \mathcal{D}_{jjt}\left(p_{jt} - mc_{jt}\right) = 0, \tag{3.3}$$

for all $j \in \{1, ..., J\}$, where w, x, α, β_j have been suppressed for convenience, and where $\mathcal{D}_{kjt} \equiv \frac{\partial Q_k(\mathbf{p}_t, x_t, \alpha)}{\partial p_j}$, the time t cross-price elasticity of product k with respect to p_j.[8] This set of equations derived from the assumption of Bertrand-Nash equilibrium can then be used to solve for each product's markup. If, however, the firms are not playing a Bertrand-Nash equilibrium every period, estimated markups and marginal cost parameters based on these equations will be inconsistent. There are many reasons for departure from Bertrand-Nash equilibrium, including the possibility of other static equilibria (e.g. Cournot), collusion on the part of some or all of the firms, and market dynamics.

Another way of performing inference on model parameters is the so-called menu approach. The idea is to consider a finite number of possibilities for equilibrium behavior and corresponding first order conditions. Which condition is appropriate in a given setting can be tested, as done by

[8] $\mathcal{D}_{kjt}$ is not a function of t, but may vary with $\mathbf{p}_t$ and x_t.

Gasmi, Laffont, and Vuong (1992), for example, or can be chosen according to the researcher's judgement. Nevo (1998) summarizes this approach by writing (3.3) in vector notation

$$\mathbf{Q}_t - \Omega\left(\mathbf{p}_t - \mathbf{mc}_t\right) = 0,$$

where Ω is the identity matrix, and $\mathbf{Q}_t$, $\mathbf{p}_t$, $\mathbf{mc}_t$ denote the vector of all firms' period t quantities produced, prices, and marginal costs. The menu approach entertains the possibility that subsets of firms in the industry collude. If a subset of firms in the industry are colluding, the system of equations is modified by setting $\Omega_{ij} = 1$ as well for any combination of firms i and j that collude. The conduct parameter approach instead allows each Ω_{ij}, $i \neq j$, to take any value on $[0, 1]$, but maintains the restriction that the Ω_{ij} are constant across observations. Given this restriction, the researcher could, in principle, hope to estimate the matrix of conduct parameters Ω. Nevo (1998) makes the point that even in this case, it will often be extremely difficult to have enough instrumental variables to identify these parameters. Even if we did have enough exclusion restrictions to identify the conduct parameters, the argument of Corts (1999) still applies. Both the menu approach and the conduct parameter method have the inherent shortcoming that Ω_{ij} is fixed across periods of observation; as in the homogeneous products case, this is exactly the restriction I relax, and I thus add the subscript t to Ω_{ij} to reflect this.

Equation by equation, we have $\forall j \in \{1, ..., J\}$, $\forall t$,

$$Q_{jt} + \left(p_{jt} - mc_{jt}\right)\mathcal{D}_{jjt} + \sum_{k \neq j} \Omega_{jkt}\left(p_{kt} - mc_{kt}\right)\mathcal{D}_{kjt} = 0.$$

for some set of values of Ω_{ijt}, where again $\mathcal{D}_{kjt} \equiv \frac{\partial Q_k(\mathbf{p}_t, x_t, \alpha)}{\partial p_j}$. As in the case of homogeneous products, the requirements for identification of $Q\left(\mathbf{p}_t, x_t, \alpha\right)$ are exactly the same as they are under the assumption of Bertrand-Nash equilibrium. Estimation of the demand-side parameters is well-studied[9], so I will again assume that these are observed. Isolating mc_{jt} in the above equation, it follows that for all products $j = 1, ..., J$,

$$mc_{jt}\mathcal{D}_{jjt} = p_{jt}D_{jjt} + Q_{jt} + \sum_{k \neq j} \Omega_{jkt}\left(p_{kt} - mc_{kt}\right)\mathcal{D}_{kjt}. \tag{3.4}$$

In place of assuming a particular value for each Ω_{jkt}, I only restrict them to lie in the interval $\left[\Omega^0_{jkt}, \Omega^1_{jkt}\right]$, encapsulating a range of possible firm decision rules. A natural special case is that where $\left[\Omega^0_{jkt}, \Omega^1_{jkt}\right] = [0, 1]$ for all j, k, t combinations, encapsulating a set of decision rules ranging from Bertrand-Nash Equilibrium to jointly collusive behavior. Assuming that $p_{kt} - mc_{kt} \geq 0$ for

[9] See Nevo (2000) for references.

all observations,[10] $\mathcal{D}_{kjt} \geq 0$ for $j \neq k$ and $\mathcal{D}_{jjt} \leq 0$, it follows that for all products $j = 1, ..., J$,

$$mc_{jt}\mathcal{D}_{jjt} \leq p_{jt}D_{jjt} + Q_{jt} + \sum_{k \neq j} \Omega^1_{jkt} \left(p_{kt} - mc_{kt}\right) \mathcal{D}_{kjt}, \tag{3.5}$$

$$mc_{jt}\mathcal{D}_{jjt} \geq p_{jt}D_{jjt} + Q_{jt} + \sum_{k \neq j} \Omega^0_{jkt} \left(p_{kt} - mc_{kt}\right) \mathcal{D}_{kjt}. \tag{3.6}$$

In a perfect information game in which all marginal cost and demand errors are perfectly observed, it simply follows that these inequalities must hold over all observations. Since demand can be identified under the standard assumptions, consistent estimates for demand parameters can be obtained in a first stage regression. For ease of notation I impose throughout the remainder of the discussion of the differentiated products case that $\left[\Omega^0_{jkt}, \Omega^1_{jkt}\right] = [0, 1]$ for all j, k, t, although working with (3.5) and (3.6) for other sets of $\left\{\Omega^0_{jkt}, \Omega^1_{jkt}\right\}$ follows the same logical steps.

Thus, in a perfect information setup only marginal costs are unobserved to the econometrician. In order to compare the implications of these restrictions to models that impose equilibrium behavior that pins down marginal costs, it is useful to rewrite the above inequalities isolating marginal costs:

$$mc_{jt} \geq p_{jt} + \mathcal{D}^{-1}_{jjt} \left[Q_{jt} + \sum_{k \neq j} \left(p_{kt} - mc_{kt}\right) \mathcal{D}_{kjt} \right], \tag{3.7}$$

$$mc_{jt} \leq p_{jt} + \mathcal{D}^{-1}_{jjt} Q_{jt}, \tag{3.8}$$

for all $j = 1, ..., J$. This system of inequalities essentially says that the period t vector of firms' marginal costs $\mathbf{mc}_t = (mc_{1t}, ..., mc_{Jt})'$ is known to belong to a subset of $\mathbb{R}^J$. Because of the existence of cross-price elasticities in a differentiated product market, there are a plethora of restrictions across the bounds on each firm's marginal costs. That is, in contrast to the homogeneous product case, the bounds for each firm's marginal costs include all of the other firms' marginal costs. If marginal costs are equal across all firms, so that $\mathbf{mc}_t$ is a scalar, then this complication no longer exists. Formal identification results are provided in section 4.2.

4 Identification

4.1 Homogeneous Product Industries

Let $\mathbf{q}_t \equiv (q_{1t}, ..., q_{nt})'$ and $\boldsymbol{\epsilon}_t \equiv (\epsilon_{1t}, ..., \epsilon_{nt})'$ denote the vector of period t firm quantities produced and marginal cost errors. Throughout, I use $(\mathbf{q}, p, x, w, u, \boldsymbol{\epsilon})$ to denote a representative realization of $(\mathbf{q}_t, p_t, x_t, w_t, u_t, \boldsymbol{\epsilon}_t)$. Let the observations $\{(u_t, \boldsymbol{\epsilon}_t) : t = 1, ..., T\}$ be a random sample distributed

[10] I am assuming here that all pairs of differentiated products are gross substitutes. This formulation excludes the case where two of the products are complements.

$\mathbb{P}$ such that $\mathbb{E}[u, \boldsymbol{\epsilon}|x, w] = 0$. (x, w) are assumed to have bounded support, and $(\mathbf{q}, p)$ are assumed to have nonnegative, bounded support. I assume throughout that firms observe $(u, \boldsymbol{\epsilon})$ prior to making their decisions. Firms' decision rules and market demand jointly determine the realization of $(\mathbf{q}, p)$ as a function of $(x, w, u, \boldsymbol{\epsilon})$. For clarity I use X, W to explicitly denote X, W as random variables and x, w to denote representative realizations. Thus, $\mathbb{E}[mc(q_i, W, \beta_i)|X, W]$ is a random variable, while $\mathbb{E}[mc(q_i, w, \beta_i)|x, w]$ is the particular value of $\mathbb{E}[mc(q_i, W, \beta_i)|X, W]$ for $X = x$ and $W = w$. Formally, these assumptions are summarized below. B denotes the parameter space within which β is assumed to belong, and B^* is used to denote the identified set for β. That is, B^* is the subset of marginal cost parameters that are logically consistent with the restrictions of the model.

Assumption: (A1) Each component of $(\mathbf{q}, p)$ has nonnegative, bounded support. (X, W) have bounded support $\mathcal{X}$,$\mathcal{W}$, respectively.

Assumption: (A2) $\mathbb{E}[u, \boldsymbol{\epsilon}|X, W] = 0$.

Both of these assumptions are standard. A further assumption is

Assumption: (A3) The inverse demand function $P(\cdot, \cdot, \alpha)$ and thus its derivative $\frac{\partial P(\cdot,\cdot,\alpha)}{\partial Q}$ are identified.

Assumption (A3) is motivated by the observation that the parameters of the inverse demand equation are identified under the same exclusion restrictions as in traditional, point identified models. In the inverse demand equation (2.1) Q_t is endogenous, since the firms observe u_t before making their decisions. As in standard models, endogeneity can be handled with appropriate instruments for Q_t, for which the cost covariates w_t are immediate candidates. Because identification of the parameters of a single equation using instrumental variables is well studied and not the focus here, I assume that the researcher does indeed have enough exclusion restrictions to identify and estimate inverse demand independently of the marginal cost function. In the linear example of section 6, and the empirical application of section 7, this condition is easily seen to be met, so that the parameters of the demand equation are identified.[11] Thus, for the purpose of identification, I restrict attention to the case where the parameters of the demand function are identified and can be consistently estimated, perhaps, for example, by a "first-step" regression. Thus for the remainder of this section I treat the inverse demand function and hence its derivative with respect to Q as observed.

I also assume throughout that the inverse demand function is continuously differentiable and downward sloping.

Assumption: (A4) $P(\cdot, x, \alpha)$ is continuously differentiable in Q and $\frac{\partial P(\cdot,x,\alpha)}{\partial Q} < 0$ for all $x \in \mathcal{X}$.

As a preliminary step for identification the following proposition derives a bound for expected

[11] If $P(Q_t, x_t, \alpha)$ is linear in Q_t, then all that is required is that for some component of w_t, w_{1t}, $cov(Q_t, w_{1t}) \neq 0$ and $E[u_t w_{1t}] = 0$. The first condition is satisfied because firms take the demand unobservable into account when choosing their strategy, while the second will be satisfied by assumption; indeed I invoke the stronger assumption $E[u_t|w_t] = 0$.

marginal costs conditional on exogenous covariates under (A**) and (A1)-(A4).

Proposition 1 *Let (A1)-(A4) and (A**) hold. Then* $\forall (x, w) \in (\mathcal{X}, \mathcal{W})$*, for each* $i = 1, ..., n$,

$$\underline{mc}_i (x, w) \leq \mathbb{E} [mc (q_i, w, \beta_i) | x, w] \leq \overline{mc}_i (x, w), \tag{4.1}$$

where

$$\begin{aligned} \underline{mc}_i (x, w) &\equiv \mathbb{E} [p_t | x, w] + \mathbb{E} \left[\theta_{it}^1 Q \frac{\partial P}{\partial Q} | x, w \right], \\ \overline{mc}_i (x, w) &\equiv \mathbb{E} [p_t | x, w] + \mathbb{E} \left[\theta_{it}^0 Q \frac{\partial P}{\partial Q} | x, w \right]. \end{aligned}$$

Proof. All proofs are included in the Appendix. ■

When instead a particular equilibrium is imposed, then a moment equation is obtained from each firm's first order condition, which can then be used to estimate the parameters of their marginal costs. Now however, due to the incomplete knowledge of firms' decision rules, a set of possible moment equations must be considered, as $mc (q_i, w, \beta_i)$ is only known to lie between $\underline{mc}_i (x, w)$ and $\overline{mc}_i (x, w)$. The previous proposition shows that the conditional expectation of (A**) with respect to the distribution of $(u, \epsilon_1, ..., \epsilon_N)$ bounds the value of firms' marginal costs conditional on exogenous covariates.

Condition (4.1) is a system of conditional moment inequalities, for each firm i, that are implied by the restrictions on firms' decision rules. This system may in general have cross-inequality restrictions, complicating the analysis, much like simultaneity in a system of equations. In the special case where firms share no common cost parameters, this complication is not present. Furthermore, the presence of the endogenous variables $\mathbf{q}$ is a complication in this model just like it is in traditional, point identified models. In particular, the endogeneity of $\mathbf{q}$ means that marginal costs may vary with the demand covariates x indirectly through $\mathbf{q}$.

While β will in general not be point identified, the above inequalities do place meaningful restrictions on expected marginal costs conditional on covariates, and thus on β. Exactly what can be learned as a consequence of the bounds given by Proposition 1 will depend on the particular functional form assumptions imposed, and what quantities are identified by the sampling process. For example, if one is only interested in bounds for $\mathbb{E} [mc (q_i, w, \beta_i) | x, w]$ for each combination of (x, w), then identification of the conditional expectation functions $\underline{mc}_i (x, w)$ and $\overline{mc}_i (x, w)$ will suffice. If instead the goal is inference on the marginal cost parameters β, then more will be required if $mc (q_i, w, \beta_i)$ depends on q_i. The extent to which (4.1) has identifying power for β depends on which components of the inequality are identified by the sampling process. Thus, to further characterize the implications of (A**) for the identified set for β, I restrict $mc (q_i, w, \beta_i)$ to fall in the class of parametric forms of assumption (A5).

Assumption: (A5: additive separability) $mc(q_i, w, \beta_i) = f(w, \beta_i) + g(q_i, \beta_i)$.

This assumption embeds a number of functional forms from the literature. In particular, constant marginal costs and linear marginal costs are included. When assumption (A5) is violated, Proposition 1 still restricts the feasible values for β, and these could be examined on a case-by-case basis. When (A5) holds, the following additional high-level assumption, along with condition (4.1), can be used to obtain the identified set for β of Proposition 2.

Assumption: (A6: identification of conditional expectations) For all $(x, w) \in (\mathcal{X}, \mathcal{W})$, $i = 1, ..., n$, $\underline{mc}_i(x, w)$, $\overline{mc}_i(x, w)$, and $\mathbb{E}[g(q_i, \beta_i)|x, w]$ are identified.

Exactly which conditions guarantee that assumption (A6) holds depend on the functional form of both the inverse demand and marginal cost function. If, for example, inverse demand and marginal costs are linear, then identification of the conditional means $\mathbb{E}[p_t|x, w]$, $\mathbb{E}\left[\theta_{it}^0 Q|x, w\right]$, $\mathbb{E}\left[\theta_{it}^1 Q|x, w\right]$ and $\mathbb{E}[q_{it}|x, w]$ will suffice. Given sufficient regularity conditions, each of these conditional expectations can then be consistently estimated.

Proposition 2 proceeds to characterize the identified set for β, the composite vector of all firms' marginal cost parameters β_i, under these additional assumptions.

Proposition 2 *Assume that (A**) and (A1)-(A6) hold. Then the identified set for β is*

$$B^* = \cap_{i=1,...,n} \left\{ \begin{array}{c} \beta \in B : \underline{mc}_i(X, W) \leq f(W, \beta_i) + \mathbb{E}[g(q_i, \beta_i)|X, W] \leq \overline{mc}_i(X, W) \\ \textit{with probability 1} \end{array} \right\}. \tag{4.2}$$

Corollary 1 *Assume that (A**) and (A1)-(A6) hold. For any $\tilde{x}$ contained in the support of X, define $\mathcal{Q}(\tilde{x}) \equiv \left\{Q : \frac{\partial P}{\partial Q}(Q, \tilde{x}, \alpha) = 0\right\}$, the set of values of Q such that $\frac{\partial P}{\partial Q}(Q, \tilde{x}) = 0$. Then $\mathbb{E}[mc(q_i, W, \beta_i)|W = w]$ is identified if $\Pr\{Q \in \mathcal{Q}(\tilde{x})|\tilde{x}, w\} = 1$. Furthermore, if there is a value of the demand covarites x^* such that $\frac{\partial P}{\partial Q}(Q, x^*) = 0$ for any Q, then $\mathbb{E}[mc(q_i, W, \beta_i)|W = w]$ is identified for all w such that (x^*, w) lies on the support of X, W.*

Proposition 2 characterizes the identified set for the marginal cost parameters, and forms the basis of the estimation strategy developed in section 5. The first corollary verifies that firms' expected marginal costs conditional on covariates w are point-identified if, conditional on w, there is probability one of a realization of demand covariates $\tilde{x}$ such that demand is infinitely elastic. While it may not in general be reasonable to expect demand to be infinitely elastic for any hypothesized value of demand covariates, this corollary is still worth mention for two reasons. First, if demand is indeed infinitely elastic conditional on some set of observables, then firms must be price-takers, and marginal costs can be recovered since it must be that price equals marginal cost. Thus, this presents one case in which we would expect the researcher to be able to identify mean marginal costs, and indeed this methodology satisfies that criteria. Second, identification need not be viewed as an "all-or-nothing" proposition. The corollary presents the limiting case of the observation that

the more elastic the aggregate demand function, the smaller will be the identified set for marginal costs and, in turn, the identified set for marginal cost parameters B^*.

In particular, this characterization of the identification region reveals that the size of the identified set is sensitive to the distribution of $Q\frac{\partial P}{\partial Q}(Q, x, \alpha) = p \cdot \eta^{-1}$, where η is the elasticity of quantity demanded with respect to price. This result is rather intuitive: if demand is highly elastic, then there is less potential for price manipulation on the part of the firms, since small changes in price have a large effect on quantity demanded. If, on the other hand, demand is inelastic, then there is more potential for collusive firms to profit, as consumers will not modify their purchasing behavior as much in the face of higher prices. Furthermore, as the polar case of Corollary 1 also illustrates, a relatively small identification region does not require the elasticity to be uniformly high. Rather, if the elasticity is sufficiently high on average over a broad enough range of covariate values, a small identification region may be obtained.

A second corollary shows that if marginal costs are linear, then B^* is convex. Knowledge that the identified set is convex may also be useful to ease the computational burden of estimating B^*.

Corollary 2 *If for all i, $mc(q_i, w, \beta_i) = \beta_{iq} \cdot q_i + w'\beta_{iw}$ where $\beta_i = \left(\beta_{iq}, \beta'_{iw}\right)'$, then B^* is convex.*

4.2 Differentiated Product Industries

In this section I derive formal identification results for marginal cost parameters for the case of differentiated products. The starting point for characterizing the identified set is equation (3.4). The environment and the notation follow that of section 3 as closely as possible. Let $\boldsymbol{\epsilon}_t \equiv (\epsilon_{1t}, ..., \epsilon_{nt})'$ and $\mathbf{u}_t = (u_{1t}, ..., u_{Jt})'$. Using notation similar to that of the homogeneous products framework, I use $(\mathbf{q}, \mathbf{p}, w, x, \mathbf{u}, \boldsymbol{\epsilon})$ to denote representative observations of $(\mathbf{q}_t, \mathbf{p}_t, w_t, x_t, u_t, \boldsymbol{\epsilon}_t)$. Formally, the sampling process consists of a random sample of draws of $(x, w, \mathbf{u}, \boldsymbol{\epsilon})$ from population $(\Omega, \Sigma, \mathbb{P})$. The measure $\mathbb{P}$ induces a conditional probability measure over $(\mathbf{u}, \boldsymbol{\epsilon})$ for each possible realization of (x, w) such that the conditional expectation of $(\mathbf{u}, \boldsymbol{\epsilon})$ given (x, w) is zero. Firms' decision rules and market demand determine the realization of $(\mathbf{q}, \mathbf{p})$ as a function of $(x, w, \mathbf{u}, \boldsymbol{\epsilon})$. As before, I use X, W to explicitly denote the covariates as random variables and x, w to denote representative realizations. B denotes the parameter space within which β is assumed to belong, and B^* is used to denote the identified set for β.

Assumption: (B1) Each component of $(\mathbf{q}, \mathbf{p})$ has nonnegative, bounded support. (X, W) have bounded support $\mathcal{X} \times \mathcal{W}$, respectively.

Assumption: (B2) $\mathbb{E}[u, \boldsymbol{\epsilon}|X, W] = 0$.

As in the case of homogeneous products, I assume that the dimension of w, the variables excluded from the demand specification, are sufficient to identify the parameters of the demand specification, α. The requirements for identification of α are no different that those of models where firms are assumed to play a particular equilibrium.

Assumption: (B3) α is identified.

Replacing mc_{jt} with $mc\left(w_t, \beta_j\right) + \epsilon_{jt}$ for each j and taking conditional expectations of the inequalities (3.5) and (3.6), it follows that the true parameter β must satisfy the restrictions

$$\mathbb{E}\left[p_j|x,w\right] + \mathbb{E}\left[\mathcal{D}_{jj}^{-1}Q_j\left(\mathbf{p}\right)|x,w\right] \geq mc_j\left(w,\beta_j\right) \tag{4.3}$$

$$\geq \mathbb{E}\left[p_j|x,w\right] + \mathbb{E}\left[\mathcal{D}_{jj}^{-1}Q_j\left(\mathbf{p}\right)|x,w\right] + \mathbb{E}\left[\mathcal{D}_{jj}^{-1}\sum_{k\neq j}\mathcal{D}_{kj}\left(p_k - mc_k\left(w,\beta_k\right) - \epsilon_k\right)|x,w\right],$$

for each product j. However, the quantity $\mathbb{E}\left[\mathcal{D}_{jj}^{-1}\mathcal{D}_{kj}\epsilon_k|x,w\right]$ is not identified without adding more structure, since the marginal cost shifters ϵ_k are not observed by the econometrician. In particular, if the demand derivatives $\mathcal{D}_{kj}$ depend on the price of goods aside from good j, then we might expect that $\mathcal{D}_{kj}$ and ϵ_k are correlated, even after conditioning on x and w, since ϵ_k is correlated with p_k. Thus, I impose the restriction that the demand function for each product j is linear in prices, which implies that the cross-price derivatives $\mathcal{D}_{kj}$ are constant.

Assumption: (B4) $Q_j\left(\mathbf{p},x,\alpha\right) = \alpha_{pj}\cdot\mathbf{p} + h\left(\alpha_{xj},x\right)$, where $\alpha_{pj} = \left(\mathcal{D}_{j1},...,\mathcal{D}_{jJ}\right)'$, and α_{xj} are those components of α through which firm j's marginal cost function depends on covariates x.

Assumption: (B5) $\mathbb{E}\left[\mathbf{p}|x,w\right]$ is identified on $\mathcal{X}\times\mathcal{W}$.

Under these additional assumptions, proposition 3 gives the identified set for the marginal cost parameters β.

Proposition 3 *Let $\mathcal{D}_{kj} > 0$ for all $k \neq j$, and $\mathcal{D}_{jj} < 0$. Assume that (B1)-(B5) hold, and that inequalities (3.7) and (3.8) both hold. Then the identified set for the demand parameters β is*

$$B^* = \cap_{j=1}^{J}\left\{\begin{array}{c}\beta\in B : \mathbb{E}\left[p_j|X,W\right] + \mathcal{D}_{jj}^{-1}\alpha_{pj}\mathbb{E}\left[\left(\mathbf{p}\right)|X,W\right] + h\left(\alpha_{xj},X\right) \geq mc_j\left(W,\beta_j\right) \geq \\ \left(\begin{array}{c}\mathbb{E}\left[p_j|X,W\right] + \mathcal{D}_{jj}^{-1}\alpha_{pj}\mathbb{E}\left[\mathbf{p}|X,W\right] + h\left(\alpha_{xj},X\right) \\ +\mathcal{D}_{jj}^{-1}\mathbb{E}\left[\sum_{k\neq j}\mathcal{D}_{kj}\left(p_k - mc_k\left(w,\beta_k\right) - \epsilon_k\right)|X,W\right]\end{array}\right) \\ \textit{with probability } 1\end{array}\right\}.$$

Corollary 3 *If $mc_j\left(w,\beta\right) = w_j\cdot\beta_j$ for all $j\in J$, where w_j,β_j are subvectors of w and β, respectively, then B^* is convex.*

As was the case for homogeneous good industries, the identification result given by proposition 3 can be used as the basis to estimate B^* as outlined in the following section. However, doing so appears to be generally rather difficult computationally. In comparison with the homogeneous products model, there are many more terms in the inequalities that define the identified set. Furthermore, the bounds on each firms' marginal cost functions are functions of the other firms'

(set-identified) marginal costs. The corollary provides sufficient conditions for the identified set to be convex, which as in the homogeneous product case may be useful for estimation.

5 Estimation

The approach I take here in either homogeneous product or differentiated product industries is to estimate the model in two steps. The idea is to first estimate the parameters of the demand specification, and then use these "first-step" estimates to estimate the identified set for the marginal cost parameters. To perform the second stage, I formulate a population level objective function that attains its minimum only on those values of the cost parameters that belong to the identified set. Estimation then proceeds by computing the set of approximate minimizers of the sample analog of this objective function.

For estimation of the model one can use a standard GMM procedure as a first step to estimate demand parameters, and a modified minimum distance (MMD) estimator as defined by Manski and Tamer (2002) to estimate the identified set B^*. MMD estimation is a straightforward method for constructing a sample analog of the identified set. First, the researcher defines a population objective function of model parameters Q that is minimized only on the identified set. This is achieved by constructing a nonnegative valued function that takes the value zero if and only if $\beta \in B^*$. By construction, an equivalent formulation for the identified set for the parameters β is

$$B^* = \arg\min_{b \in B} Q(b).$$

Estimation of B^* then proceeds by constructing a sample analog for the objective function, $\hat{Q}_T$ such that $\hat{Q}_T$ converges uniformly almost surely to Q over the parameter space B. Then one computes the set of approximate minimizers of $\hat{Q}_T$,

$$\hat{B}^* = \left\{ \beta \in B : \hat{Q}_T(\beta) \leq \min_{b \in B} Q_T(b) + \epsilon_T \right\},$$

where $\epsilon_T \downarrow 0$ at an appropriate rate as $T \to \infty$. Under sufficient regularity conditions, $\hat{B}^*$ converges to B^* in the Hausdorff metric as $T \to \infty$. The need for the sequence of ϵ_T is the same as that of Manski and Tamer (2002). That is, in order to guarantee that $\hat{B}^*$ contains the boundary of B^* asymptotically as $T \to \infty$, the inequalities that define the identified set must be relaxed slightly. The consistency proofs of their paper are generally applicable for MMD estimation based on any nonnegative objective function that is only equal to zero on the identified set, and for which there is a uniformly consistent estimator. Under sufficient regularity conditions, $\hat{B}^*$ is consistent for B^* in the Hausdorff metric, as long as ϵ_T converges to zero no faster than $\hat{Q}$ converges to Q.

While Manski and Tamer (2002) showed that MMD estimators for set identified parameters are

generally consistent, they did not devise a method for statistical inference for MMD estimators. Indeed, constructing confidence regions for partially identified parameters is a field of current research. To perform inference for the empirical example of section 7, I use the modified minimum distance estimator developed in Rosen (2005), which is applicable in models that can be written as a finite number of conditional moment inequalities. The MMD objective function of that paper has the advantage that the asymptotic distribution of $T \cdot \hat{Q}_T(\beta)$ is chi-bar-square. As described in that paper, this result can then be used to construct asymptotically valid confidence sets for the parameter β, even if it is only partially identified.

While the model used for the application of section 7 can be written as a finite set of moment inequalities, the general methodology of this paper can result in models that do not fit within that framework. In this case, one would need to use a different inferential approach to construct confidence sets for β. Other recent papers that devise a way to perform inference in partially identified models and could potentially be applied include Andrews, Berry, and Jia (2004), Chernozhukov, Hong, and Tamer (2004), Shaikh (2005), Beresteanu and Molinari (2006), and Pakes, Porter, Ho, and Ishii (2006).

The inferential approach I employ from Rosen (2005) uses the set of moment inequalities implied by the model as a basis for estimation, and I briefly outline that approach here. Because the covariates in the application have discrete support, the identified set for β implied by proposition 2 can be written as a finite set of moment inequalities of the form

$$\mathbb{E}\left[m\left(\mathbf{q}, p, x, w, \alpha, \beta\right)\right] = \mathbb{E}\begin{bmatrix} m_1\left(\mathbf{q}, p, x, w, \alpha, \beta\right) \\ \vdots \\ m_K\left(\mathbf{q}, p, x, w, \alpha, \beta\right) \end{bmatrix} \geq \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{5.1}$$

where the inequality applies to each element of the moment vector $\mathbb{E}\left[m\left(\mathbf{q}, p, x, w, \alpha, \beta\right)\right]$. α is consistently estimated in a first stage by GMM, using the observed cost shifter w as an instrumental variable for market price. The modified minimum distance objective function I use then takes the form

$$Q\left(\alpha, \beta\right) = \min_{t \geq 0} \left(\mathbb{E}\left[m\left(\mathbf{q}, p, x, w, \alpha, \beta\right)\right] - t\right) V_{\alpha,\beta}^{-1} \left(\mathbb{E}\left[m\left(\mathbf{q}, p, x, w, \alpha, \beta\right)\right] - t\right),$$

where V is the variance matrix of the vector $m\left(\mathbf{q}, p, x, w, \alpha, \beta\right)$. Notice that, as a function of β, evaluated at the true value of α, $Q\left(\alpha, \beta\right) = 0$ if and only if (5.1) is satisfied, so that $Q\left(\alpha, \beta\right)$ is indeed a modified minimum distance objective function. To estimate the identified set for β, I use the sample objective function

$$\hat{Q}_T\left(\beta\right) = \min_{t \geq 0} \left(\hat{E}_n\left[m\left(\mathbf{q}, p, x, w, \hat{\alpha}, \beta\right)\right] - t\right) \hat{V}_\beta^{-1} \left(\hat{E}_n\left[m\left(\mathbf{q}, p, x, w, \hat{\alpha}, \beta\right)\right] - t\right),$$

where $\hat{E}_n$ denotes the sample mean and $\hat{V}_\beta$ is the sample variance of $m\left(\mathbf{q}, p, x, w, \hat{\alpha}, \beta\right)$, corrected

by means of the delta method to account for the fact that $\hat{\alpha}$ is separately estimated. Under mild regularity conditions, if $\beta \in B^*$, violations of $\hat{Q}_T(\beta) = 0$ are attributable to no more than sampling variation. To build confidence sets for β, I then make use of the asymptotic distribution of statistics of the form $T\hat{Q}_T(\beta)$ as derived in Rosen (2005).

6 A Simple Example: A Linear Model with Instruments

In order to illustrate the proposed methodology before proceeding to the empirical example of section 7, in this section I consider a simple linear model. The model has both demand and supply shifters that serve as instruments. The market is for a homogeneous good with single period inverse demand function

$$p_t = a_0 + a_1 x_t + a_2 Q_t + u_t. \tag{6.1}$$

Further suppose that two firms compete for consumers in each period of observation, and that each has constant marginal costs given by the following functional form

$$mc_i(q_{it}; w_t) = c_{0i} + c_{1i} w_t + \epsilon_{it}. \tag{6.2}$$

As before, t indexes the observation, while $i = 1, 2$ indexes the firms. u_t is an unobserved period t demand shifter, while $\epsilon_{1t}, \epsilon_{2t}$ are unobserved marginal cost shifters. $\epsilon_{1t}, \epsilon_{2t}$ and u_t are assumed to be iid mean zero, independent of each other, and uncorrelated with all of the regressors, so that $E[\epsilon_{1t}, \epsilon_{2t}, u_t | x_t, w_t] = 0$. These additive demand and marginal cost shifters (u_t, ϵ_{it}) are perfectly observed by each firm i, but unobserved by the econometrician. The variables $p_t, q_{1t}, q_{2t}, w_t, x_t$ are all observed by the firms and by the econometrician.

Firms play a repeated quantity-setting game in a market characterized by these linear inverse demand and linear marginal cost functions. Following the behavioral assumptions considered in section 2, they are assumed to play strategies in each period that satisfy the inequalities (A**). In this model, the conditional expectations of these inequalities are

$$-a_2 \mathbb{E}\left[\theta_{it}^0 Q_t | x_t, w_t\right] \leq a_0 + a_1 x_t + a_2 Q_t - c_{0i} - c_{1i} w_t \leq -a_2 \mathbb{E}\left[\theta_{it}^1 Q_t | x_t, w_t\right], \tag{6.3}$$

for all i, t pairs. The structural parameters to be estimated in this model are $a_0, a_1, a_2, c_{01}, c_{02}, c_{11}$ and c_{12}. In this case, w_t is correlated with Q_t through the firms' supply relation, but is not correlated with the demand errors by assumption. Thus, w_t can be used as an instrument for quantity supplied in estimation of the linear demand equation, and the demand parameters are identified and can be consistently estimated. An instrument for quantity demanded, x_t is also available in this model, but the traditional IV techniques are not applicable for estimation of the supply-side parameters c_0 and c_1, which are not point identified.

The conditional moment inequalities implied by the model are given by

$$\begin{aligned}
\mathbb{E}\left\{p_t - (a_0 + a_1 x_t + a_2 Q_t)\,|x_t, w_t\right\} &= 0, \\
\mathbb{E}\left\{a_0 + a_1 x_t + a_2 Q_t - c_{01} - c_{11} w_t + \theta_{1t}^0 a_2 Q_t | x_t, w_t\right\} &\geq 0, \\
\mathbb{E}\left\{a_0 + a_1 x_t + a_2 Q_t - c_{02} - c_{12} w_t + \theta_{2t}^0 a_2 Q_t | x_t, w_t\right\} &\geq 0, \\
\mathbb{E}\left\{a_0 + a_1 x_t + a_2 Q_t - c_{01} - c_{11} w_t + \theta_{1t}^1 a_2 Q_t | x_t, w_t\right\} &\leq 0, \\
\mathbb{E}\left\{a_0 + a_1 x_t + a_2 Q_t - c_{02} - c_{12} w_t + \theta_{2t}^1 a_2 Q_t | x_t, w_t\right\} &\leq 0.
\end{aligned}$$

The first restriction implies that

$$\begin{aligned}
a_0 + a_1 \mathbb{E}\left\{x_t\right\} &= \mathbb{E}\left\{p_t\right\} - a_2 \mathbb{E}\left\{Q_t\right\}, \\
a_0 \mathbb{E}\left\{x_t\right\} + a_1 \mathbb{E}\left\{x_t^2\right\} &= \mathbb{E}\left\{p_t x_t\right\} - a_2 \mathbb{E}\left\{Q_t x_t\right\}, \\
a_0 \mathbb{E}\left\{w_t\right\} + a_1 \mathbb{E}\left\{x_t w_t\right\} &= \mathbb{E}\left\{p_t w_t\right\} - a_2 \mathbb{E}\left\{Q_t w_t\right\},
\end{aligned}$$

from which it is clear that the demand parameters are identified, barring linear dependence of these three equations. Furthermore, a_0, a_1, and a_2 can be estimated consistently by their sample analogs.

To use the full identifying power of the restrictions imposed, I apply Proposition 2 to characterize the identification region. To characterize sufficient conditions for identification of α, let

$$M = \begin{pmatrix} 1 & \mathbb{E}(x) & \mathbb{E}\{Q\} \\ \mathbb{E}(x) & \mathbb{E}\left(x^2\right) & \mathbb{E}\{Qx\} \\ \mathbb{E}\{w\} & \mathbb{E}\{xw\} & \mathbb{E}\{Qw\} \end{pmatrix}.$$

The conditions that M is non-singular and that $\mathbb{E}\left[\left(1+\theta_t^1\right) Q|X,W\right]$ and $\mathbb{E}\left[\left(1+\theta_t^0\right) Q|X,W\right]$ are identified yield the following result.[12]

Corollary 4 *(to Proposition 2) In the linear duopoly model of this section, assume that* $\mathbb{E}\left[\left(1+\theta_t^1\right) Q|X,W\right]$ *and* $\mathbb{E}\left[\left(1+\theta_t^0\right) Q|X,W\right]$*, are identified, and that* M *has full rank. Then* α *is identified and the identified set for* $\beta = (c_{01}, c_{11}, c_{02}, c_{12})$ *is*

$$B^* = \left\{ \begin{array}{c} \beta \in B : a_0 + a_1 x_t + a_2 \mathbb{E}\left[\left(1+\theta_{1t}^1\right) Q_t|X,W\right] \leq c_{01} + c_{11} W \leq a_0 + a_1 x_t + a_2 \mathbb{E}\left[\left(1+\theta_{1t}^0\right) Q_t|X,W\right] \\ \textit{and } a_0 + a_1 x_t + a_2 \mathbb{E}\left[\left(1+\theta_{2t}^1\right) Q_t|X,W\right] \leq c_{02} + c_{12} W \leq a_0 + a_1 x_t + a_2 \mathbb{E}\left[\left(1+\theta_{2t}^0\right) Q_t|X,W\right] \\ \textit{with probability } 1 \end{array} \right\}$$

[12] Sufficient conditions for identification of the condtional expectations $\mathbb{E}\left[\left(1+\theta_t^1\right) Q_t|X,W\right]$ and $\mathbb{E}\left[\left(1+\theta_t^0\right) Q_t|X,W\right]$ are straightforward. Suppose, for example, that the bounds on θ_{it}, θ_t^0 and θ_t^1 are constant across t. Then all that is required is identification of $\mathbb{E}\left[Q_t|X,W\right]$. If $Q_t = \eta\left(x_t, w_t\right) + \delta_t$, for some continuous but unknown function η, where δ_t are iid and independent of x_t and w_t. The conditional mean functions $\mathbb{E}\left[\left(1+\theta_t^1\right) Q_t|X,W\right]$ and $\mathbb{E}\left[\left(1+\theta_t^0\right) Q_t|X,W\right]$ could then be consistently estimated by kernel estimation, for example.

7 Application to the Joint Executive Committee

In this section I demonstrate the methodology with an application to data on the Joint Executive Committee, a late 19th century railway cartel.

7.1 Background and Data

Formed on April 18, 1879, the Joint Executive Committee (henceforth JEC) was a cartel that sought to sustain cooperation among rail carriers over the so called trunk-line territory between the Atlantic seaboard and midwest commercial centers.[13] Both eastbound and westbound traffic was monitored and regulated by the cartel, each under different organizations within the JEC. Because the JEC predated the formation of the Interstate Commerce Commission (1887) and the passage of the Sherman Act (1890), the operation of the cartel was publicly acknowledged, and detailed accounts of its operation are well-documented by, for example, MacAvoy (1965) and Ulen (1979).

The dataset is a time series of weekly observations from the beginning of 1880 through the 16th week of 1886. For each week, tons of grain, flour, and provisions shipped on each of 8 different rail lines from Chicago to the east coast is recorded. These rail lines were owned by the 5 different firms that participated in the cartel for at least some subset of the years considered here. These weekly quantities as well as percentage allotments and cartel prices for each of the goods shipped were recorded by the JEC itself in order to assist in the task of maintaining collusion. Table 1 provides a description of the variables employed from this dataset, and Table 2 provides various summary statistics.

Also available is an indicator L_t that equals 1 if the Great Lakes were open to navigation in week t, and equals zero otherwise. Shipment via lake was the main competitor for rail shipments during this period, so L_t was an important determinant of demand for rail shipments. As noted by Porter (1983) and Lee and Porter (1984), one would ideally have data on prices charged for lake shipments, but this data is not available to me. It is, however, known that the JEC at no point in time managed to explicitly include the lake shippers in their collusive endeavors.[14].

There has been a good deal of prior research on the JEC that has employed the dataset used here. This includes Porter (1983), Lee and Porter (1984), Cosslett and Lee (1985), Porter (1985), Berry and Briggs (1988), Hajivassiliou (1989), and Ellison (1994). These studies focus primarily on examining explicit models of collusion, and the extent to which the observed data can be explained by particular theoretical models. Their focus is not on the estimation of firms' marginal costs per se, and in these studies marginal costs are not estimated. In fact, in this line of research firms'

[13] Ulen (1979)

[14] Ulen (1979, p. 229) notes that the cartel did once discuss the idea of colluding with lake shippers in September 1886, but that this idea was never pursued. One potential reason for this, however, is that some of the lake steamers were owned by some of the same companies as the railroads.

Table 1: List of Variables

p	price of shipping grain in dollars per 100 lbs.**
Q	total quantity of grain shipped, in tons.**
q_i	quantity of grain shipped by firm i, in tons.
A_i	market share allotment for firm i as dictated by the JEC.
L	1 if Great Lakes were open for shipping, 0 otherwise.**
S_1	seasonal dummy; 1 if observation is from $Q2$, 0 otherwise.
S_2	seasonal dummy; 1 if observation is from $Q3$, 0 otherwise.
S_3	seasonal dummy; 1 if observation is from $Q4$, 0 otherwise.
DM_1	1 from week 28 in 1880 to week 10 in 1883, 0 otherwise; reflecting entry by the Grand Trunk Railway.**
DM_2	1 from week 11 in 1883 to week 25 in 1883, 0 otherwise; reflecting an addition to the New York Central.**
DM_3	1 from week 26 in 1883 to week 11 in 1886, 0 otherwise; reflecting entry by the Chicago and Atlantic.**
DM_4	1 from week 12 in 1886 to week 16 in 1886, 0 otherwise; reflecting departure of the Chicago and Atlantic from the JEC.**

* The Sample is from week 1 in 1880 to week 16 in 1886.

** Source: Porter (1983), Table 1.

equilibrium conditions have been aggregated in such a way so that marginal cost parameters may not be identified by the model used, unless further assumptions are imposed. Here, rather than posit a particular theory of collusive behavior, I entertain a wide range of possibilities, and my focus is primarily on the estimation of marginal cost parameters. Of course, in a model that yields precise inference of marginal cost parameters, much could potentially be said about the magnitude of firms' markups and the extent to which they compete or collude.

7.2 Demand Specification and Estimation

I take the functional form of Porter (1983) for the market demand equation as a starting point, but relax the assumption of normal errors. Instead, I assume that the demand unobservable has zero expectation conditional on L and the season dummies $S = (S_1, S_2, S_3)'$. Market demand in period t is assumed to have the log-linear form

$$\ln Q_t = \alpha_0 + \alpha_1 \ln p_t + \alpha_2 L_t + \sum_{j=3}^{5} \alpha_j S_{j-2,t} + u_t. \tag{D1}$$

The functional form differs from that of Porter (1983) only in that I employ three rather than twelve seasonal dummies, dividing the calendar into quarters rather than four week periods. This structure is imposed to make the large sample approximations for the supply side of the market reasonable with the available data. Each *Lake*, *Season* combination implies two moment inequalities for each

Table 2: Summary Statistics

Variables	Mean	Median	Standard Deviation	Minimum	Maximum
p	.2465	.25	.06653	.125	.40
Q	25384	23101	11632	4810	76407
L	.6189	1	.4864	0	1
q_1	13111	12722	6303	2403	35973
q_2	5846	4961	3413	1291	24258
q_3	2201	1716	1611	189	11277
q_4^*	3098	2690	2053	0	9592
q_5^*	3197	2810	2204	0	13732

* Firm 4 was active from week 28 through the end of the sample.
** Firm 5 was active from week 182 through week 323.
The mean, median, and standard deviation for firms 4 and 5 are reported conditional on that firm being active.

firm; dividing the year into thirteen seasons rather than four results in some *Lake*, *Season* pairs with too few observations to expect reliable inference.

Furthermore, I allow u_t to be governed by an AR(1) process, as in Ellison (1994) and Cosslett and Lee (1985), who found significant serial correlation in the demand errors. That is, I impose that

$$u_t = \rho \cdot u_{t-1} + \xi_t, \tag{7.1}$$

where ξ_t is iid and ρ is estimated as a parameter of the model. I maintain the assumption that u_t, and thus u_{t-1} and ξ_t are uncorrelated with both L_t and all the season dummies, and I use a two step efficient GMM procedure to estimate the parameters of the demand equation and the autoregressive parameter separately from the marginal cost parameters. The variables DM_1, DM_2, DM_3, DM_4 and the allotments of four of the firms are used as instruments for $\ln p_t$; that is, I make use of the moment equation $\mathbb{E}\left[u|DM_j, A_k\right] = 0$ for $j = 1, ..., 4$ and $k = 1, ..., 4$.[15].

I also consider the specification

$$\ln Q_t = \alpha_0 + \alpha_1 \ln p_t + \alpha_2 L_t + \sum_{j=3}^{5} \alpha_j S_{j-2,t} + \alpha_6 L_t \ln p_t + u_t, \tag{D2}$$

along with the AR(1) process for the error term of (7.1). This allows for the elasticity of market demand to vary with whether or not the Great Lakes were open for shipment. Again the parameters are estimated by an efficient GMM procedure, and firm allotments interacted with L are used as additional instruments. The estimation results of both specifications are reported in Table 3.

The estimates from the more flexible demand specification suggest that demand was in fact

[15] Only the allotments of the first four firms are used because the sum of all firms' allotments is 1 by definition.

Table 3: Demand Estimates

	Specification 1: No interaction	*Specification 2: $L_t, \ln p_t$ interaction*
α_0 (const)	8.9655***	9.6226***
	(0.1671)	(0.1317)
α_1 ($\ln p_t$)	-0.9143***	-0.4157**
	(0.2449)	(0.1848)
α_2 (Lake dummy)	-0.4374***	-2.1148***
	(0.1232)	(0.0879)
α_3 (Q2 dummy)	0.1229*	0.1510**
	(0.0738)	(0.0605)
α_4 (Q3 dummy)	-0.1278*	-0.0906
	(0.074)	(0.0569)
α_5 (Q4 dummy)	0.076	0.0723
	(0.0732)	(0.0553)
α_6 ($L_t * \ln p_t$)	—	-1.1315***
	—	(0.1306)
ρ	0.6225***	0.2188***
	(0.0033)	(0.0114)

*** denotes significance at the 0.01 level.
** denotes significance at the 0.05 level.
* denotes significance at the 0.10 level.

considerably more elastic when the lakes were open for navigation. This finding seems reasonable; when the lakes were open, the railway prices for grain shipping may have been, at least to some extent, constrained by the prices of lake shippers, making quantity more sensitive to price changes in those periods. A further difference between the two sets of demand estimates is the estimated autoregressive parameter ρ. The estimate is roughly 3 times as high when demand elasticity is assumed constant across all periods, although it is very highly statistically significant under both specifications. The autoregressive parameter has no bearing on marginal cost estimates, however.

I proceed to estimate marginal costs under each of the demand specifications, separately for each of the three firms that are active over the entire dataset. It is important to note that given these functional form specifications, the conduct parameter method, which assumes $\theta_{it} = \theta_i$ for each firm i and all t, identifies none of the θ_i nor the marginal cost parameters, even when marginal costs are assumed constant. Due to the lack of marginal cost shifters, marginal costs cannot be separately identified from θ_i. The methodology of this paper is, however, still applicable.

7.3 Marginal Cost Specification and Estimation

I consider two different specifications for firms' marginal cost functions, the first being constant marginal costs, and the second being linear marginal costs. There are no observed marginal cost shifters aside from quantity, which is endogenous. The specification is

$$mc\left(q_{it}\right) = c_{0i} + \epsilon_{it}, \tag{7.2}$$

when marginal costs are assumed constant, and

$$mc\left(q_{it}\right) = c_{0i} + c_{1i}q_{it} + \epsilon_{it}, \tag{7.3}$$

when marginal costs are assumed to be linear. ϵ_{it} is an unobserved marginal cost shifter such that $\mathbb{E}\left[\epsilon_{it}|L,S\right] = 0$. The composite vector of firms' marginal cost parameters to be estimated is $\beta = \left(c_{01}, c_{02}, c_{03}\right)'$ when marginal costs are given by (7.2) and $\beta = \left(c_{01}, c_{11}, c_{02}, c_{12}, c_{03}, c_{13}\right)$ when marginal costs are given by (7.3). The units of the variables are such that $mc\left(q_{it}\right)$ is measured in dollars per 100 pounds of grain. It is assumed throughout that all components of β are nonnegative.

I assume that in each period, firms' strategies are such that market outcomes satisfy the restriction (A**),

$$-\theta_{it}^0 Q_t \frac{dP\left(Q_t, L_t, S_t\right)}{dQ} \leq p_t - mc_{it} \leq -\theta_{it}^1 Q_t \frac{dP\left(Q_t, L_t, S_t\right)}{dQ}, \tag{7.4}$$

for all active firms i at time t, and I consider the identifying power of two different sequences of values for θ_{it}^0 and θ_{it}^1, $\left(\theta_{it}^0, \theta_{it}^1\right) = (0,1)$ and $\left(\theta_{it}^0, \theta_{it}^1\right) = (0, s_{it})$. Given the demand specification, it

is convenient to rewrite this restriction as

$$\left(1+\frac{\theta_{it}^{1}}{\eta_t}\right)\cdot p_t \leq mc_{it} \leq \left(1+\frac{\theta_{it}^{0}}{\eta_t}\right)\cdot p_t, \tag{7.5}$$

where η_t is the elasticity of aggregate demand in period t. Under (D1), η_t is given by

$$\eta_t = \alpha_1,$$

while under (D2),

$$\eta_t = \alpha_1 + \alpha_6 L_t.$$

Looking at the estimated values of the demand equation, the estimated elasticity under the first specification is constant at $\hat{\alpha}_1 = -0.9143$, while under the second specification, the demand elasticity is -0.4157 when the Great Lakes are closed and -1.5472 when the Great Lakes are open. From (7.5), and more generally from Proposition 1, it is clear that the bounds on mc_{it}, are smaller when demand is more elastic. Thus, the difference in demand elasticities under the two specifications should be expected to have an effect on marginal cost parameter estimates. However, the direction of this difference is not clear a priori, since demand is more elastic under (D1) when the Lakes are closed than under (D2), but demand is more elastic under (D2) than (D1) when the Lakes are open. The effect of the different estimates for demand elasticity is reflected in the estimates of the marginal cost parameters in the remainder of this section.

In this model, Proposition 2 implies that the identified set for the marginal cost parameters β is given by those value for β that satisfy

$$\mathbb{E}\left\{\left(1+\frac{\theta_i^1}{\eta}\right)\cdot p|L,S\right\} \leq \mathbb{E}\left\{mc\left(\beta_i\right)|L,S\right\} \leq \mathbb{E}\left\{\left(1+\frac{\theta_i^0}{\eta}\right)\cdot p|L,S\right\}. \tag{7.6}$$

Because the Great Lakes were always open in Q3, and always closed in Q1, there are six *Lake*, *Season* combinations observed in the sample. As a result, (7.6) yields twelve moment inequalities for each firm, two for each *Lake*, *Season* combination. The set of parameter values that comprise the identified set, B^* are those values of β such that each firm's marginal cost parameters satisfy each of these twelve moment inequalities. In the case of constant marginal costs, with specification (7.2), these moment inequalities are

$$\mathbb{E}\left\{\left(1+\frac{\theta_i^1}{\eta}\right)\cdot p|L,S\right\} \leq c_{0i} \leq \mathbb{E}\left\{\left(1+\frac{\theta_i^0}{\eta}\right)\cdot p|L,S\right\},$$

while for specification (7.3) they are

$$\mathbb{E}\left\{\left(1+\frac{\theta_i^1}{\eta}\right)\cdot p|L,S\right\} \leq c_{0i} + c_{1i}\cdot\mathbb{E}\left\{q_i|L,S\right\} \leq \mathbb{E}\left\{\left(1+\frac{\theta_i^0}{\eta}\right)\cdot p|L,S\right\}.$$

7.3.1 Marginal Cost Estimates Given Demand Specification 1

First, I estimate the identified set for marginal cost parameters when demand is assumed to be given by (D1)[16]. Table 4 shows the estimated bounds on marginal costs for firms 1-3 when marginal costs are assumed constant. Under the assumption that $\theta_{it} \in [0, 1]$ for all t, the bounds on marginal costs for each firm are identical, ranging from marginal costs of 0 to 22¢. In this case, there is no firm specific component to (7.5), as neither marginal costs nor the bounds on marginal costs implied from the demand estimates depend on firms' quantities produced. When, however, it is instead assumed that $\theta_{it} \in [0, s_{it}]$, where s_{it} is firm i's period t market share, the restriction (7.5) does have different implications for each of the firms, and the estimated bounds are substantially smaller. Furthermore, the restriction that $\theta_{it} \in [0, s_{it}]$ embeds perfect competition and Cournot Equilibrium as polar cases. Under this restriction, the marginal costs of firm 2 have a very narrow estimated identification region, falling between 21.8¢ and 22¢, with a 95% confidence interval of [20.6¢, 23.1¢]. Turning to firm 3, the estimated identified set for c_{03} is empty, as is the associated 95% confidence interval. For firm 3, the data are found to be inconsistent with all of the restrictions of the model, at the 95% level. The value of c_{03} closest to satisfying the imposed restrictions is 22.4¢.

Now I turn to estimation of marginal cost parameters under the linear specification (7.3). The estimated identified set for the marginal cost parameters (c_{0i}, c_{1i}), are depicted graphically for each firm i in figures $1 - 3$. The figures illustrate very clearly the identifying power of the assumption $\theta_{it} \in [0, s_{it}]$ relative to $\theta_{it} \in [0, 1]$, as the estimated identified sets are much smaller under the former assumption for each of the firms. Comparing the estimates across each of the three firms, judging by the shape of the identified set, it seems that firm 1's marginal costs are potentially the least sensitive to quantity, while firm 3's are potentially the most sensitive. This reflects the variation in each firms' quantity produced relative to the bounds on marginal costs implied by (7.5). Under the restriction that $\theta_{it} \in [0, s_{it}]$, the estimated identified set for firm 2's marginal cost parameters is very informative, with the slope of firm 2's marginal cost function estimated to be no greater than approximately 0.000007.[17] Interestingly, in the case of firm 3, the specification of constant marginal costs was rejected at the 95% level, but when marginal costs are allowed to be linear in quantity, the model is no longer rejected at the 95% level, although no value of the marginal cost parameters yields a value of 0 for the sample objective function. The parameter values that yielded the smallest value were $(c_{03}, c_{13}) = \left(0.178, 2.8 \times 10^{-5}\right)$.

[16] The reported estimates for the identified set are those values of the marginal cost parameter(s) β such that $\hat{Q}_T(\beta) \leq 0.0001$. This cutoff was chosen to match the degree of computation error allowed by Matlab's minimization algorithm.

[17] The slope parameter c_{1i} measures the effect of increasing quantity by one ton on firm i's marginal costs measured in dollars per 100 pounds shipped. Because quantities shipped were as high as tens of thousands of tons for each firm, values of c_{1i} on the order of 10^{-5} are economically significant. For example, a value of c_{12} of 0.000007 would contribute about 4¢ to marginal costs at firm 2's mean quantity of 5846 tons.

Table 4: Estimated bounds and confidence intervals for firms' marginal costs under the assumption that marginal costs are constant, for the constant elasticity demand specification (D1). The units of the estimated bounds are dollars per 100 pounds of grain shipped.

		$[\underline{\mathbf{mc}}, \overline{\mathbf{mc}}]$	75% CI	95% CI
Firm 1:	$\theta_{it} \in [0, 1]$	$[0, 0.220]$	$[0, 0.226]$	$[0, 0.231]$
	$\theta_{it} \in [0, s_{it}]$	$[0.118, 0.220]$	$[0.111, 0.226]$	$[0.106, 0.231]$
Firm 2:	$\theta_{it} \in [0, 1]$	$[0, 0.220]$	$[0, 0.226]$	$[0, 0.231]$
	$\theta_{it} \in [0, s_{it}]$	$[0.218, 0.220]$	$[0.211, 0.226]$	$[0.205, 0.231]$
Firm 3:	$\theta_{it} \in [0, 1]$	$[0, 0.220]$	$[0, 0.226]$	$[0, 0.231]$
	$\theta_{it} \in [0, s_{it}]$	$\emptyset$ $(0.224)^*$	$\emptyset$	$\emptyset$

$*$ denotes the unique value at which the sample objective function is minimized.

7.3.2 Marginal Cost Estimates Given Demand Specification 2

Now I consider the implications of using the estimates from demand specification 2 to estimate the identified set for marginal cost parameters. Again, I begin by considering the constant marginal costs specification, and these results are reported in table 5. In contrast to the estimates under (D1), the presence of the interaction term in the demand specification leads to rejection of marginal costs being 0 for each of the three firms. This conclusion is driven by the periods in which the lakes were open, as the demand elasticity was sufficiently high in those periods to indicate that, were the firms actually engaging in perfectly collusive behavior, their implied marginal costs would still be well above zero. Furthermore, the change in demand elasticities has no effect on the upper bounds of the estimated marginal costs. This is because the upper bound on marginal costs is derived from setting θ_{it} at its lower bound for each observation. This lower bound is 0 for both cases considered, which cancels the effect of the demand elasticity. If the lower bound on θ_{it} were non-zero, the different elasticity would have had an effect on the upper bound on marginal costs.

For the case where $\theta_{it} \in [0, s_{it}]$, using the estimated demand elasticities from (D2) does change the implied lower bound on marginal costs, but the direction of the effect, relative to the estimates under (D1), is not uniform across firms. For firm 1, the estimated lower bound changes from 11.8¢ to 16.6¢, while for firm 2 the lower bound decreases from 21.8¢ to 19.6¢. For firm 3, there is still

Table 5: Estimated bounds and confidence intervals for firms' marginal costs under the assumption that marginal costs are constant, for demand specification (D2), where elasticity is allowed to depend on whether or not the Great Lakes are open for shipment.

		$[\underline{\mathbf{mc}}, \overline{\mathbf{mc}}]$	75% CI	95% CI
Firm 1:	$\theta_{it} \in [0, 1]$	$[0.086, 0.220]$	$[0.081, 0.226]$	$[0.078, 0.231]$
	$\theta_{it} \in [0, s_{it}]$	$[0.166, 0.220]$	$[0.161, 0.226]$	$[0.155, 0.231]$
Firm 2:	$\theta_{it} \in [0, 1]$	$[0.086, 0.220]$	$[0.081, 0.226]$	$[0.078, 0.231]$
	$\theta_{it} \in [0, s_{it}]$	$[0.196, 0.220]$	$[0.190, 0.226]$	$[0.186, 0.231]$
Firm 3:	$\theta_{it} \in [0, 1]$	$[0.086, 0.220]$	$[0.081, 0.226]$	$[0.078, 0.231]$
	$\theta_{it} \in [0, s_{it}]$	$\emptyset$ $(0.223)^*$	$[0.221, 0.226]$	$[0.216, 0.231]$

$*$ denotes the unique value at which the sample objective function is minimized.

no value of c_{01} that satisfies all of the restrictions of the model, but now that the interaction term has been included in the demand specification, constant marginal costs are no longer rejected at reasonable levels.

Turning now to the set estimates for the parameters of the linear marginal cost specification, we immediately see from figures 4-6 that zero marginal costs are rejected for each of the three firms, as the origin is always excluded. Again, there is a large difference between the estimated identified sets for the two restrictions on firms' markups. Interestingly, the shapes of implied identified sets do not change by much.

As was the case when demand specification D1 was imposed, there exist no parameter values for firm 3's marginal cost function at which the sample objective function takes the value zero under the assumption that $0 \leq \theta_{3t} \leq s_{it}$. However, the specification is not rejected at the 95% level, and in contrast to the results of section 7.3.1, constant marginal costs cannot be rejected. The parameter values that minimize the sample objective function for firm 3 in this case are $(c_{03}, c_{13}) = \left(0.199,\ 1.3 \times 10^{-5}\right)$.

8 Conclusion

This paper provides a methodology for estimating firms' marginal cost functions under relatively lax restrictions on firm behavior in an oligopoly setting. The empirical literature has long recognized that empirical results may be sensitive to the particular equilibrium assumptions imposed. To the extent that it is unclear how firms actually make their strategic decisions in any given market, imposing assumptions on firm behavior may be problematic, and could lead to inconsistent parameter estimates. The methodology of this paper allows the empirical researcher to relax potentially unverifiable assumptions on firm behavior, and determine what conclusions can be drawn even if the remaining assumptions are insufficient for point identification. If more restrictions can be credibly imposed regarding firm behavior across observations, then this framework provides a starting point for evaluating the impact of those additional restrictions.

References

Abreu, D., D. Pearce, and E. Stachetti (1986): "Optimal Cartel Equilibria with Imperfect Monitoring," *Journal of Economic Theory*, 39, 251–269.

Andrews, D. K., S. T. Berry, and P. Jia (2004): "Confidence Regions for Parameters in Discrete Games with Multiple Equilibria, with an Application to Discount Chain Store Location," working paper, Yale University.

Bajari, P., C. L. Benkard, and J. Levin (forthcoming): "Estimating Dynamic Models of Imperfect Competition," .

Beresteanu, A., and F. Molinari (2006): "Asymptotic Properties for a Class of Partially Identified Models," working paper, Cornell University.

Berry, S., and H. Briggs (1988): "A Non-parametric Test of a First-Order Markov Process for Regimes in a Non-cooperatively Collusive Industry," *Economics Letters*, 27, 73–77.

Borzekowski, R., and A. M. Cohen (2005): "Estimating Strategic Complementarities in Credit Unions' Outsourcing Decisions," working paper, Federal Reserve Board of Governors.

Bresnahan, T. F. (1982): "The Oligopoly Solution Concept is Identified," *Economics Letters*, 10, 87–92.

——— (1987): "Competition and Collusion in the American Automobile Industry: the 1955 Price War," *Journal of Industrial Economics*, 35(4), 457–482.

——— (1989): "Empirical Studies of Industries with Market Power," in *The Handbook of Industrial Organization*, ed. by R. Schmalansee, and R. Willing, vol. 2, chap. 17. Elsevier Science Publishers B.V., Amsterdam.

Chernozhukov, V., H. Hong, and E. Tamer (2004): "Parameter Set Inference in a Class of Econometric Models," working paper, Princeton University.

Ciliberto, F., and E. Tamer (2004): "Market Structure and Multiple Equilibria in Airline Markets," working paper, Northwestern University.

Corts, K. S. (1999): "Conduct Parameters and the Measurement of Market Power," *Journal of Econometrics*, 88, 227–250.

Cosslett, S. R., and L.-F. Lee (1985): "Serial Correlation in Latent Discrete Variable Models," *The Journal of Econometrics*, 27, 79–97.

Ellison, G. (1994): "Theories of Cartel Stability and the Joint Executive Committee," *Rand Journal of Economics*, 25(1), 37–57.

Gasmi, F., J. Laffont, and Q. Vuong (1992): "Econometric Analysis of Collusive Behavior in a Soft-Drink Market," *Journal of Economics and Management Strategy*, 1, 277–311.

Genesove, D., and W. P. Mullin (1998): "Testing Static Oligopoly Models: Conduct and Cost in the Sugar Industry," *Rand Journal of Economics*, 29(2), 355–377.

Graddy, K. (1995): "Testing for Imperfect Competition at the Fulton Fish Market," *Rand Journal of Economics*, 26(1), 75–92.

Green, E. J., and R. H. Porter (1984): "Noncooperative Collusion under Imperfect Price Information," *Econometrica*, 52, 87–100.

Haile, P. A., and E. Tamer (2003): "Inference with an Incomplete Model of English Auctions," *Journal of Political Economy*, 111(1), 1–51.

Hajivassiliou, V. A. (1989): "Testing Game-Theoretic Models of Price Fixing Behaviour," working paper, Cowles Foundation, Yale University.

Iwata, G. (1974): "Measurement of Conjectural Variations in Oligopoly," *Econometrica*, 42, 947–966.

Lau, L. (1982): "On Identifying the Degree of Competitiveness From Industry Price and Output Data," *Economics Letters*, 10, 93–99.

Lee, L.-F., and R. H. Porter (1984): "Switching Regression Models with Imperfect Sample Separation Information – with an Application on Cartel Stability," *Econometrica*, 52(2), 391–418.

MacAvoy, P. W. (1965): *The Economic Effects of Regulation*. The MIT Press, Cambridge, MA.

Manski, C. F. (2003): *Partial Identification of Probability Distributions*. Springer-Verlag, New York.

Manski, C. F., and E. Tamer (2002): "Inference on Regressions with Interval Data on a Regressor or Outcome," *Econometrica*, 70(2), 519–546.

Mas-Colell, A., M. D. Whinston, and J. R. Green (1995): *Microeconomic Theory*. Oxford University Press, New York.

Nevo, A. (1998): "Identification of the Oligopoly Solution Concept in a Differentiated-Products Industry," *Economics Letters*, 59, 391–395.

——— (2000): "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand," *Journal of Economics and Management Strategy*, 9(4), 513–548.

Pakes, A., M. Ostrovsky, and S. Berry (2005): "Simple Estimators for Parameters of Discrete Dynamic Games (with Entry/Exit Examples)," working paper, Harvard University.

Pakes, A., J. Porter, K. Ho, and J. Ishii (2006): "The Method of Moments with Inequality Constraints," working paper, Harvard University.

Porter, R. H. (1983): "A Study of Cartel Stability: The Joint Executive Committee, 1880-1886," *Bell Journal of Economics*, 14, 301–314.

——— (1985): "On the Incidence and Duration of Price Wars," *Journal of Industrial Economics*, 33(4), 415–426.

Rosen, A. (2005): "Confidence Sets for Partially Identified Parameters that Satisfy a Finite Number of Moment Inequalities," working paper, Northwestern University.

Rosse, J. N. (1970): "Estimating Cost Function Parameters Without Using Cost Data: Illustrated Methodology," *Econometrica*, 38(2), 256–275.

Rotemberg, J. J., and G. Saloner (1986): "A Supergame Theoretic Model of Price Wars During Booms," *American Economic Review*, 76, 390–407.

Rubinovitz, R. N. (1993): "Market Power and Price Increases for Basic Cable Service Since Deregulation," *Rand Journal of Economics*, 24(1), 1–18.

Salvo, A. (2005): "Inferring Conduct Under the Threat of Entry, The Case of the Brazilian Cement Industry," working paper, Northwestern University.

Shaikh, A. M. (2005): "Inference for Partially Identified Econometric Models," working paper, Stanford University.

Sweeting, A. (2005): "Coordination Games, Multiple Equilibria and the Timing of Radio Commercials," working paper, Northwestern University.

Tamer, E. (2003): "Incomplete Simultaneous Discrete Response Models with Multiple Equilibria," *Review of Economic Studies*, 70(1), 147–167.

Ulen, T. S. (1979): "Cartels and Regulation," Ph.D. thesis, Stanford University.

Wolfram, C. D. (1999): "Measuring Duopoly Power in the British Electricity Spot Market," *American Economic Review*, 89(4), 805–826.

Appendix A: Proofs

8.1 Proposition 1

Proof. By (A**), $\exists \theta_{it} \in \left[\theta_{it}^0, \theta_{it}^1\right]$ such that $p_t - mc_{it} + \theta_{it} Q_t \frac{\partial P}{\partial Q}(Q_t, x_t, \alpha) = 0$, or equivalently,

$$mc\left(q_{it}, w, \beta_i\right) = P\left(Q_t, x_t, \alpha\right) + \theta_{it} Q_t \frac{\partial P}{\partial Q}\left(Q_t, x_t, \alpha\right) + u_t - \epsilon_{it}.$$

Taking expectations over $(\mathbf{q}_t, p_t, \theta_{it}, u_t, \boldsymbol{\epsilon}_t)$ conditional on (x, w) implies

$$\mathbb{E}\left[mc\left(q_i, W, \beta_i\right)|x, w\right] = \mathbb{E}\left[p_t|x, w\right] + \mathbb{E}\left[\theta_{it} Q_t \frac{\partial P}{\partial Q}\left(Q_t, x_t, \alpha\right)|x, w\right].$$

Because $\theta_{it} \in \left[\theta_{it}^0, \theta_{it}^1\right]$ and $Q_t \frac{\partial P}{\partial Q} \leq 0$, it follows that

$$\mathbb{E}\left[\theta_{it}^1 Q_t \frac{\partial P}{\partial Q}\left(Q_t, x_t, \alpha\right)|x, w\right] \leq \mathbb{E}\left[\theta_{it} Q_t \frac{\partial P}{\partial Q}\left(Q_t, x_t, \alpha\right)|x, w\right] \leq \mathbb{E}\left[\theta_{it}^0 Q_t \frac{\partial P}{\partial Q}\left(Q_t, x_t, \alpha\right)|x, w\right],$$

and therefore

$$\underline{mc}_i\left(x, w\right) \leq \mathbb{E}\left[mc\left(q_i, w, \beta_i\right)|x, w\right] \leq \overline{mc}_i\left(x, w\right).$$

■

8.2 Proposition 2

Proof. From Proposition 1,

$$\underline{mc}_i(x,w) \leq \mathbb{E}\left[mc\left(q_i,w,\beta_i\right)|x,w\right] \leq \overline{mc}_i(x,w)$$

for all $(x,w) \in (\mathcal{X},\mathcal{W})$, for each $i = 1,...,N$. Under the additive separability assumption A5, this becomes

$$\underline{mc}_i(x,w) \leq f(w,\beta_i) + \mathbb{E}\left[g\left(q_i,\beta_i\right)|X,W\right] \leq \overline{mc}_i(x,w).$$

for all $(x,w) \in (\mathcal{X},\mathcal{W})$, for each $i = 1,...,N$. The identification region for β is then given by the set of values of β consistent with this statement with probability 1. ■

8.2.1 Corollary 1

Proof. This is a direct implication of Proposition 2 as $\frac{\partial P}{\partial Q}(Q,x,\alpha) = 0$ implies that $\underline{mc}_i(x,w) = \overline{mc}_i(x,w) = \mathbb{E}\left[p_t|x,w\right]$, so that $\mathbb{E}\left[mc\left(q_i,w,\beta_i\right)|x,w\right] = \mathbb{E}\left[p_t|x,w\right]$ for any (x,w) pair such that $\frac{\partial P}{\partial Q}(Q,x,\alpha) = 0$ with probability 1 conditional on (x,w). ■

8.2.2 Corollary 2

Proof. Let $\beta',\beta'' \in B^*$ and define $\beta^\gamma = \gamma\beta' + (1-\gamma)\beta''$ for some $\gamma \in [0,1]$. Fix i. For any $\beta \in B^*$ it follows that

$$\overline{mc}_i(x,w) \geq \beta_{iq}\mathbb{E}\left[q_i|x,w\right] + w'\beta_{iw} \geq \underline{mc}_i(x,w)$$

so that

$$\gamma \cdot \overline{mc}_i(x,w) \geq \gamma \cdot \beta'_{iq}\mathbb{E}\left[q_i|x,w\right] + w'\left(\gamma \cdot \beta'_{iw}\right) \geq \gamma \cdot \underline{mc}_i(x,w)$$

and

$$(1-\gamma) \cdot \overline{mc}_i(x,w) \geq (1-\gamma) \cdot \beta''_{iq}\mathbb{E}\left[q_i|x,w\right] + w'\left((1-\gamma) \cdot \beta''_{iw}\right) \geq (1-\gamma) \cdot \underline{mc}_i(x,w)$$

hold. Combining these two inequalities yields

$$\overline{mc}_i(x,w) \geq \beta_{iq}\mathbb{E}\left[q_i|x,w\right] + w'\beta_{iw} \geq \underline{mc}_i(x,w).$$

This shows that the set

$$\{\beta_i : \underline{mc}_i(x,w) \leq f(w,\beta_i) + \mathbb{E}\left[g\left(q_i,\beta_i\right)|x,w\right] \leq \overline{mc}_i(x,w)\}$$

is convex. Since B^* is an intersection of sets of this form over the support of (x, w) and across i, B^* is convex. ■

8.3 Proposition 3

Proof. The result follows directly from taking expectations of the inequalities (3.5) and (3.6) conditional on covariates (x, w) as in (4.3) and imposing $Q_j(\mathbf{p}, x, \alpha) = \alpha_{pj} \cdot p + h(\alpha_{xj}, x)$. Doing so gives

$$\mathbb{E}[p_j|x, w] + \mathcal{D}_{jj}^{-1}\alpha_{pj} \cdot \mathbb{E}[(\mathbf{p})|x, w] + h(\alpha_{xj}, x) \geq mc_j(w, \beta_j)$$
$$\geq \mathbb{E}[p_j|x, w] + \mathcal{D}_{jj}^{-1}\alpha_{pj} \cdot \mathbb{E}[\mathbf{p}|x, w] + h(\alpha_{xj}, x) + \mathcal{D}_{jj}^{-1} \cdot \mathbb{E}\left[\sum_{k \neq j} \mathcal{D}_{kj}(p_k - mc_k(w, \beta_k))|x, w\right].$$

The set B^* is exactly the set of β that satisfy all of these inequalities for all J differentiated products over the support of X, W with probability 1. Any value of β that satisfies both (3.5) and (3.6) belongs to this set by construction. ■

8.3.1 Corollary 3

Proof. Let $\beta', \beta'' \in B^*$ and $\beta_\gamma = \gamma\beta' + (1-\gamma)\beta''$ for some $\gamma \in [0, 1]$. Then under the restriction that $mc_j(w, \beta) = w_j \cdot \beta_j$, it follows that w.p. 1 (X, W),

$$\mathbb{E}[p_j|X, W] + \mathcal{D}_{jj}^{-1}\alpha_{pj} \cdot \mathbb{E}[(\mathbf{p})|X, W] + h(\alpha_{xj}, X) \geq W_j \cdot \beta'_j \geq$$
$$\mathbb{E}[p_j|X, W] + \mathcal{D}_{jj}^{-1}\alpha_{pj} \cdot \mathbb{E}[\mathbf{p}|X, W] + h(\alpha_{xj}, X) + \mathcal{D}_{jj}^{-1} \cdot \mathbb{E}\left[\sum_{k \neq j} \mathcal{D}_{kj}\left(p_k - W_k \cdot \beta'_k\right)|X, W\right]$$

and

$$\mathbb{E}[p_j|X, W] + \mathcal{D}_{jj}^{-1}\alpha_{pj} \cdot \mathbb{E}[(\mathbf{p})|X, W] + h(\alpha_{xj}, X) \geq W_j \cdot \beta''_j \geq$$
$$\mathbb{E}[p_j|X, W] + \mathcal{D}_{jj}^{-1}\alpha_{pj} \cdot \mathbb{E}[\mathbf{p}|X, W] + h(\alpha_{xj}, X) + \mathcal{D}_{jj}^{-1} \cdot \mathbb{E}\left[\sum_{k \neq j} \mathcal{D}_{kj}\left(p_k - W_k \cdot \beta''_k\right)|X, W\right].$$

Multiplying the first inequality by γ the second by $(1-\gamma)$ and summing them yields

$$\mathbb{E}[p_j|X, W] + \mathcal{D}_{jj}^{-1}\alpha_{pj} \cdot \mathbb{E}[(\mathbf{p})|X, W] + h(\alpha_{xj}, X) \geq W_j \cdot \beta_\gamma \geq$$
$$\mathbb{E}[p_j|X, W] + \mathcal{D}_{jj}^{-1}\alpha_{pj} \cdot \mathbb{E}[\mathbf{p}|X, W] + h(\alpha_{xj}, X) + \mathcal{D}_{jj}^{-1} \cdot \mathbb{E}\left[\sum_{k \neq j} \mathcal{D}_{kj}\left(p_k - W_k \cdot \beta_\gamma\right)|X, W\right],$$

so that $\beta_\gamma \in B^*$. ■

Appendix B: Figures

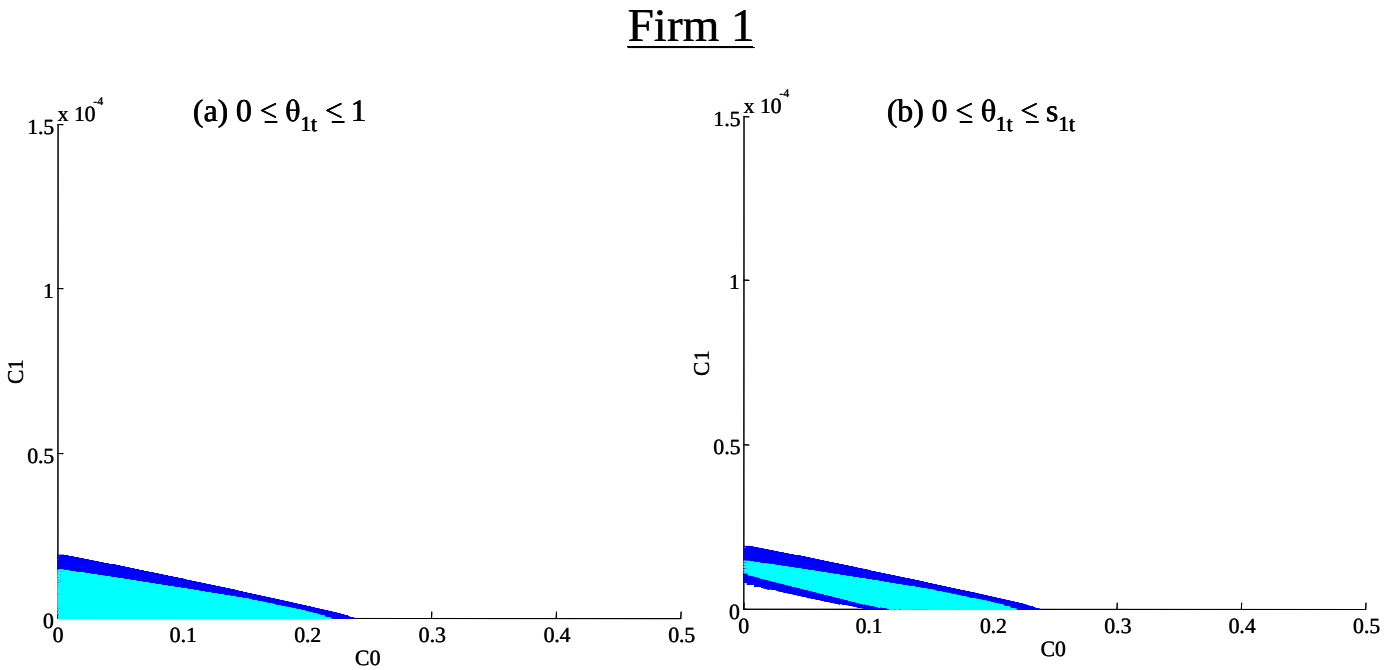

Figure 1: The estimated identified set and 95% confidence set for the marginal cost parameters of firm 1, using demand specification (D1). The estimated identified set is shown in light blue, and additional parameter values that belong to the 95% confidence region are shown in dark blue. (a) depicts the identified set under the restriction that $0 \leq \theta_{1t} \leq 1$, and (b) the identified set under the restriction that $0 \leq \theta_{1t} \leq s_{1t}$.

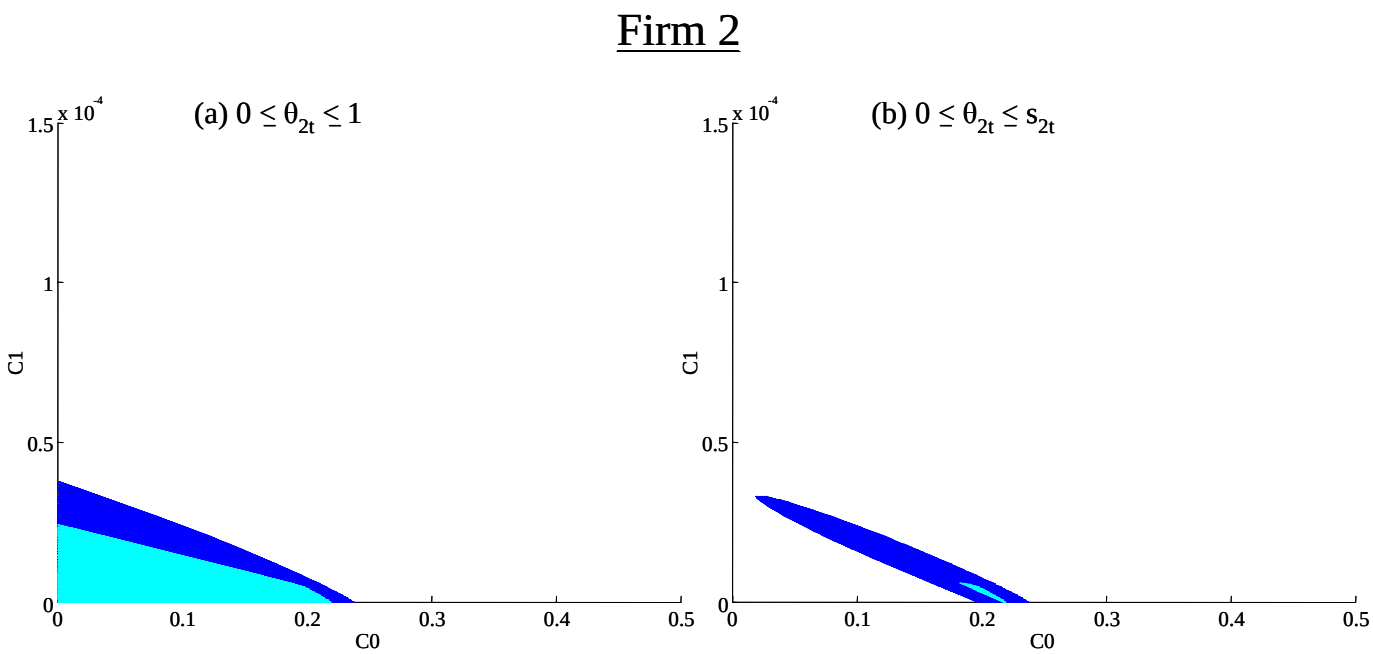

Figure 2: The estimated identified set and 95% confidence set for the marginal cost parameters of firm 2, using demand specification (D1). The estimated identified set is shown in light blue, and additional parameter values that belong to the 95% confidence region are shown in dark blue. (a) depicts the identified set under the restriction that $0 \leq \theta_{2t} \leq 1$, and (b) the identified set under the restriction that $0 \leq \theta_{2t} \leq s_{2t}$.

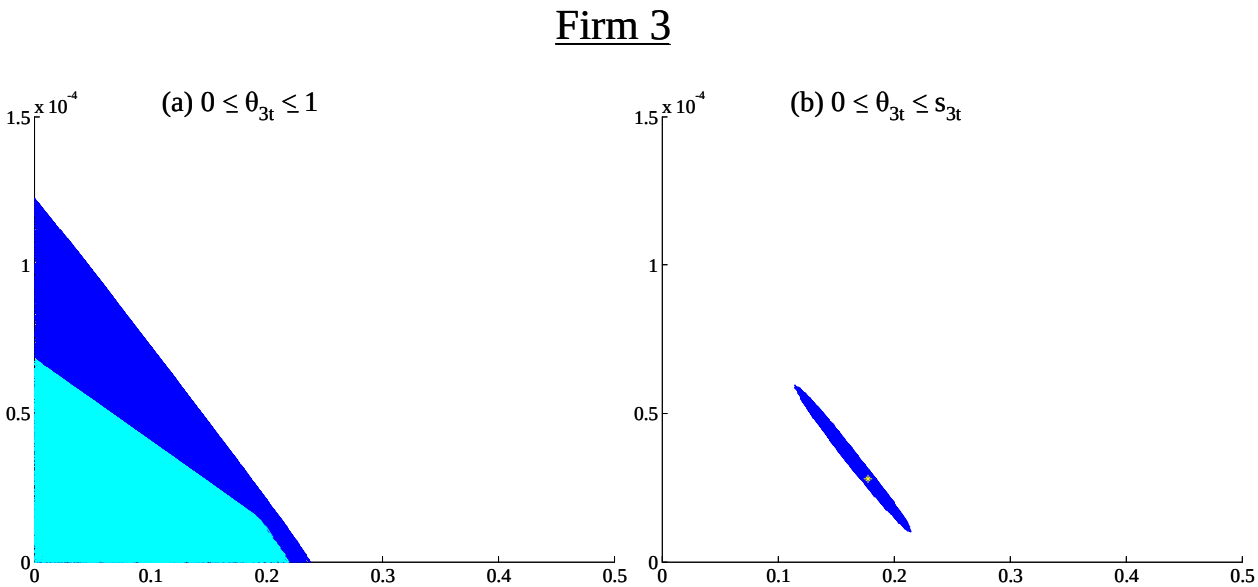

Figure 3: The estimated identified set and 95% confidence set for the marginal cost parameters of firm 3, using demand specification (D1). The estimated identified set is shown in light blue, and additional parameter values that belong to the 95% confidence region are shown in dark blue. (a) depicts the identified set under the restriction that $0 \leq \theta_{3t} \leq 1$, and (b) the identified set under the restriction that $0 \leq \theta_{3t} \leq s_{3t}$. For case (b), there is no parameter value at which the sample objective function is zero, but the 95% confidence set is non-empty. The point at which the sample objective function is minimized is $\left(0.178,\ 2.8 \times 10^{-5}\right)$, shown in yellow.

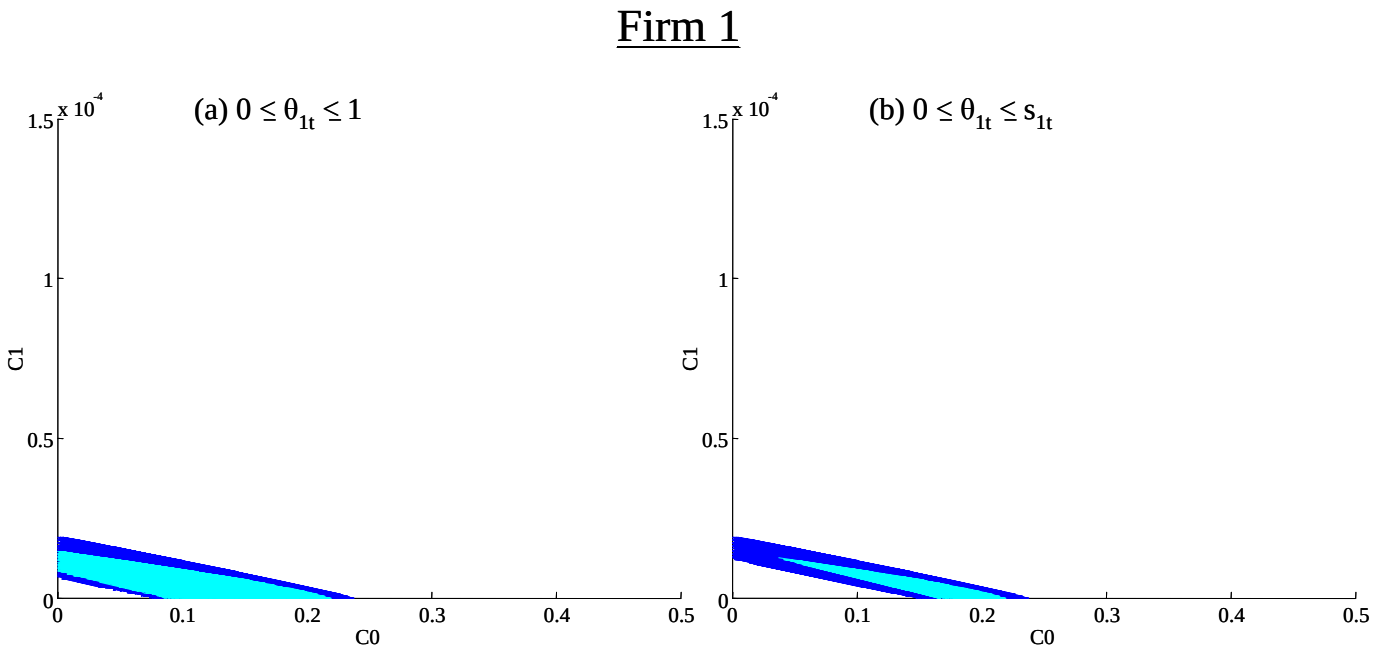

Figure 4: The estimated identified set and 95% confidence set for the marginal cost parameters of firm 1, using demand specification (D2). The estimated identified set is shown in light blue, and additional parameter values that belong to the 95% confidence region are shown in dark blue. (a) depicts the identified set under the restriction that $0 \leq \theta_{1t} \leq 1$, and (b) the identified set under the restriction that $0 \leq \theta_{1t} \leq s_{1t}$.

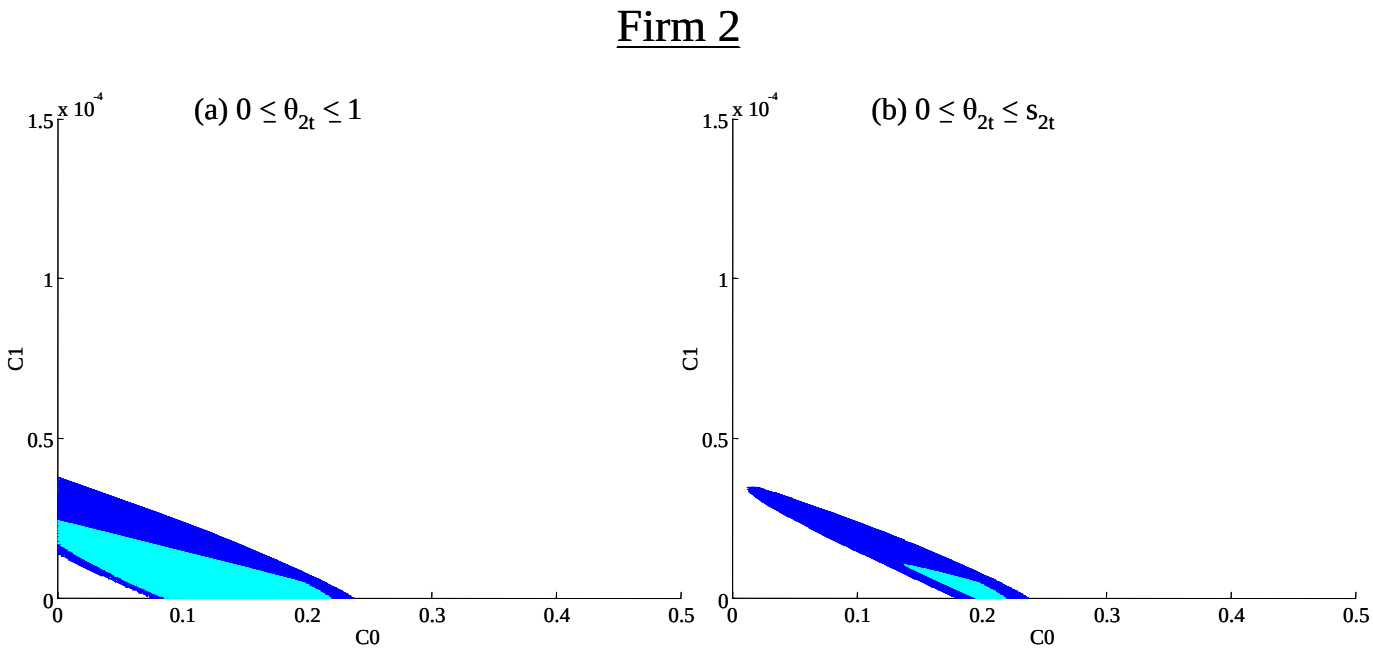

Figure 5: The estimated identified set and 95% confidence set for the marginal cost parameters of firm 2, using demand specification (D2). The estimated identified set is shown in light blue, and additional parameter values that belong to the 95% confidence region are shown in dark blue. (a) depicts the identified set under the restriction that $0 \leq \theta_{2t} \leq 1$, and (b) the identified set under the restriction that $0 \leq \theta_{2t} \leq s_{2t}$.

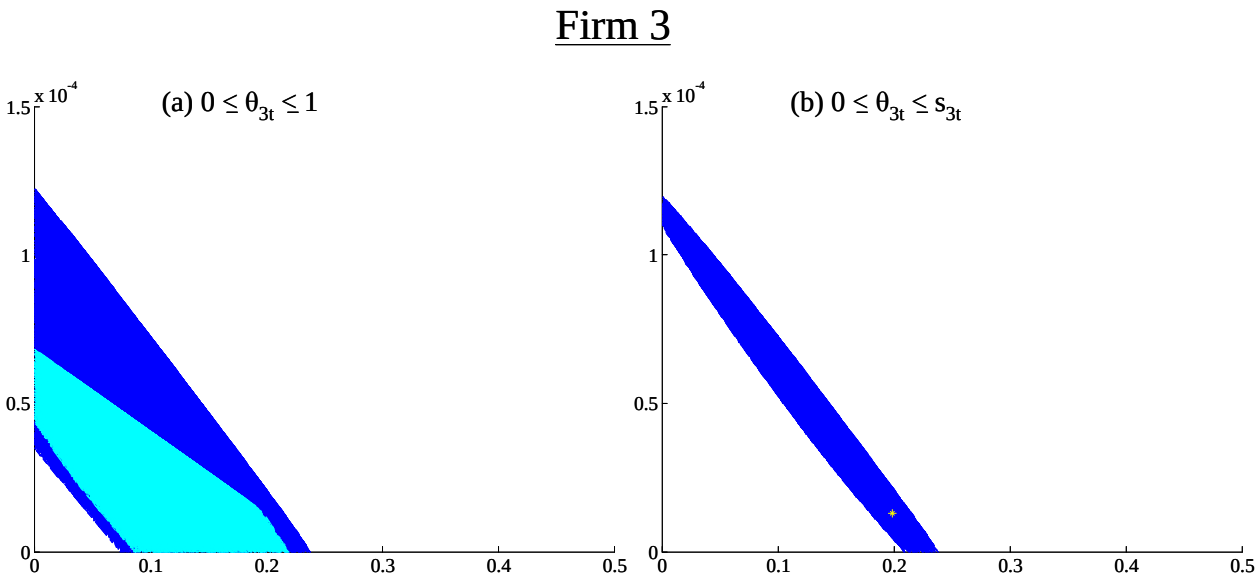

Figure 6: The estimated identified set and 95% confidence set for the marginal cost parameters of firm 3, using demand specification (D2). The estimated identified set is shown in light blue, and additional parameter values that belong to the 95% confidence region are shown in dark blue. (a) depicts the identified set under the restriction that $0 \leq \theta_{3t} \leq 1$, and (b) the identified set under the restriction that $0 \leq \theta_{3t} \leq s_{3t}$. For case (b), there is no parameter value at which the sample objective function is zero, but the 95% confidence set is non-empty. The point at which the sample objective function is minimized is $\left(0.199,\ 1.3 \times 10^{-5}\right)$, shown in yellow.