4. Competition in prices/quantities

4.1 Introduction

The decisions of how much to produce and what price to charge are fundamental determinants of firms' profits. These decisions are also main sources of strategic interactions: a firm's profit not only depends on its own decisions but also on other firms' actions. In the market for a homogeneous good, the price declines with total output such that a firm's profit also declines with the amount of output produced by its competitors. In a differentiated product industry, demand for a firm's product increases with the prices of products sold by other firms. These strategic interactions have first order importance to understand competition and outcomes in most industries. For this reason, models of competition where firms choose prices or quantities are at the core of Industrial Organization.

The answers to many economy questions in IO require not only the estimation of demand and cost functions but also the explicit specification of an equilibrium model of competition. For instance, evaluating the effects on prices, profits, and welfare of an increase in the minimum wage (or in the sales tax rate) requires to understand firms' incentives to change their prices or outputs in response to a change in costs. This incentive depends on their beliefs about what other firms will do: that is, it depends on how firms compete in the market.

The estimation of competition models can provide information on firms' marginal costs, on the form of competition, and on the demand function. In many empirical applications, the researcher has information on firms' prices and quantities sold, but information on firms' costs is not always available. The researcher may not observe the amounts of firms' inputs, such that it is not even possible to obtain costs by estimating the production function as described in chapter 3. In this context, empirical models of competition in prices or quantities may provide an approach to obtain estimates of firms' marginal costs and the structure of the marginal cost function, such as the magnitude of economies of scale or scope. Given an assumption about the form of competition (for instance, perfect competition, Cournot, Bertrand, Stackelberg, or collusion), the model predicts that a firm's marginal cost should be equal to the marginal revenue implied by that form of competition. This is the key condition that is used to estimate firms' marginal costs in this class of models. Typically, the first step in the econometric analysis of these models consists in the estimation of the demand function or demand system. Given the estimated demand, we can construct an estimate of the realized marginal revenue for every observation in the sample. This measure of marginal revenue provides, directly, an estimate of the realized marginal cost at each sample observation. Finally, we use this sample of realized marginal costs to estimate the marginal cost function, and how the marginal cost depends on the firm's output of different products (that is, economies of scale and scope), and possibly on other firm characteristics such as historical cumulative output, installed capacity, or geographic distance between the firm's production plants (that is, economies of density).

The value of a firm's marginal revenue depends on the form of competition in the industry, or *the nature of competition*. Given the same demand function, the marginal revenue is different under perfect competition, Cournot, Bertrand, or collusion. The researcher's selection of a model of competition should answer the following questions: (a) is the product homogeneous or differentiated; (b) do firms compete in prices or in quantities?; (c) is there collusion between some or all the firms in the industry?; and (d) what does a firm believe about the behavior of other firms in the market? For instance, if the researcher assumes that the product is homogenous, that firms compete in quantities, that there is no collusion in the industry, and that firms choose their levels of output under the belief that the other firms will not change their respective output levels (that is, Nash assumption), then the form of competition is the one specified in the Cournot model. In principle, some of these assumptions may be supported by the researcher's knowledge of the industry. However, in general, some of these assumptions are difficult to justify. Ideally, we would like to learn from our data about the nature of competition. Suppose that the researcher has data on firms' marginal costs (or estimates of these costs based on a production function) and has estimated the demand system. Then, given an assumption about the form of competition in this industry (for instance, perfect competition, Cournot, collusion), the researcher can use the demand to obtain firms' marginal revenues and check whether they are equal to the observed marginal costs. That is, the researcher can test if a particular form of competition is consistent with the data. In this way, it is possible to find the form of competition that is consistent with the data, for instance, identify if there is evidence of collusive behavior. We will see in this chapter that, even if the researcher does not have data on firms' costs, it is still possible to combine the demand system and the equilibrium conditions to jointly identify marginal costs and the *nature of competition* in the industry. This is the main purpose of the so called *conjectural variation approach*.

Section 4.2 presents empirical models of competition in a homogenous product industry. Section 4.3 deals with competition in a differentiated product industry. We present the conjectural variation approach both in homogenous and differentiated product industries. Section 4.4 describes models of price and quantity competition when firms have asymmetric or incomplete information.

4.2 Homogenous product industry

4.2.1 Estimating marginal costs

First, we consider the situation where the researcher does not have direct measures of marginal costs and uses the equilibrium conditions to estimate these costs.

Perfect competition

We first illustrate this approach in the context of a perfectly competitive industry for a homogeneous product. Suppose that the researcher knows, or is willing to assume, that the industry under study is perfectly competitive, and she has data on the market price and on firms' output for T periods of time (or T geographic markets) that we index by t. The dataset consists of $\{p_t, q_{it}\}$ for $i = 1,2,...,N_t$ and $t = 1,2,...,T$, where N_t is the number of firms active at period t. The variable profit of firm i is $p_t\ q_{it} - C_i(q_{it})$. Under perfect competition, the marginal revenue of any firm i is the market price, p_t. The marginal condition of profit maximization for firm i is $p_t = MC_i(q_{it})$ where $MC_i(q_{it})$ is the marginal cost, $MC_i(q_{it}) \equiv C_i'(q_{it})$. Since all the firms face the same market price, a first important implication of the first order condition of optimality under perfect competition is that all the firms should have the same realized marginal costs. This is a testable restriction of the assumption of perfect competition with a homogeneous product.

Consider a particular specification of the cost function. With a Cobb-Douglas production function, we have that (see section 3.2.1 above):

$$MC_i(q_{it}) = q_{it}^{\theta}\ w_{1it}^{\alpha_1}\ ...\ w_{Jit}^{\alpha_J}\ \exp\{\varepsilon_{it}^{MC}\} \tag{4.1}$$

w_{jit} is the price of variable input j for firm i, and α's are the technological parameters in the Cobb-Douglas production function. Variable ε_{it}^{MC} is unobservable to the researcher and it captures the cost (in)efficiency of a firm that depends on the firm's total factor productivity, and input prices that are not observable. The technological parameter θ is equal to $(1-\alpha_V)/\alpha_V$, where α_V is the sum of the Cobb-Douglas coefficients of all the variable inputs: $\alpha_V \equiv \alpha_1 + ... + \alpha_J$. Therefore, the equilibrium condition $p_t = MC_i(q_{it})$ implies the following regression model in logarithms:

$$\ln(p_t) = \theta\ \ln(q_{it}) + \alpha_1\ \ln(w_{1it}) + ... + \alpha_J\ \ln(w_{Jit}) + \varepsilon_{it}^{MC} \tag{4.2}$$

We can distinguish three cases for parameter θ. Constant Returns to Scale (CRS), with $\alpha_V = 1$ such that $\theta = 0$ and the marginal cost does not depend on the level of output; and Decreasing (Increasing) Returns to Scale, with $\alpha_V < 1$ ($\alpha_V > 1$) such that $\theta > 0$ ($\theta < 0$) and the log-marginal cost function is an increasing (decreasing) linear function of log-output.

Using data on market price, firms' quantities and firms' inputs, we can estimate the slope parameter θ in this regression equation. Even if the researcher does not have data on any of the firms' inputs, we can estimate parameter θ in this regression equation, as all the inputs become part of the error term ε_{it}^{MC}. As we explain below, we should be careful with endogeneity problems due to the correlation between this error term an a firm's output. Given an estimate of parameter θ, we can estimate ε_{it}^{MC} as a residual from this regression. Therefore, we can estimate the marginal cost function of each firm. Since the dependent variable of the regression, $\ln(p_t)$, is constant over firms, then, by construction, if $theta > 0$, then firms that produce more output should have a smaller value for the term $\alpha_1 \ln(w_{1it}) + ... + \alpha_J \ln(w_{Jit}) + \varepsilon_{it}^{MC}$: that is, they should be more cost-efficient.

Estimation of equation (4.2) by OLS suffers from an endogeneity problem. The equilibrium condition implies that firms with a large value of ε_{it}^{MC} are less cost-efficient and, all else equal, should have a lower level of output. Therefore, the regressor $\ln(q_{it})$ is negatively correlated with the error term ε_{it}^{MC}. This negative correlation between the regressor and the error term implies that the OLS estimator provides a downward biased estimate of the true θ. For instance, the OLS estimate could show increasing returns to scale, $\theta < 0$, when in fact the true technology has decreasing returns to scale, $\theta > 0$. This endogeneity problem does not disappear if we consider the model in market means.

We can deal with this endogeneity problem by using instrumental variables. Suppose that $\mathbf{x}_t^D$ is an observable variable (or vector of variables) that affects the demand of the product but not the marginal costs of the firms. The equilibrium of the model implies that these demand variables should be correlated with firms' outputs, $\ln(q_{it})$: exogenous variables that shift the demand curve should have an impact on the amount output of each firm in the market. The condition that $\mathbf{x}_t^D$ is correlated with firms' output is testable. Under the assumption that these observable demand variables $\mathbf{x}_t^D$ are not correlated with the unobserved term in the marginal cost, we can use these variables as instruments for log-output in the regression equation (4.2) to obtain a consistent estimator of θ.

Cournot competition

Now, suppose that the researcher assumes that the market is not perfectly competitive and that firms compete à la Nash-Cournot. Demand can be represented using the inverse demand function $p_t = P\left(Q_t, \mathbf{x}_t^D\right)$, where $Q_t \equiv \sum_{i=1}^N q_{it}$ is the market total output, and $\mathbf{x}_t^D$ is a vector of exogenous market characteristics affecting demand. Each firm chooses its own output q_{it} to maximize profit. Profit maximization implies the condition that marginal revenue equals marginal cost, where the marginal revenue function is:

$$MR_{it} \quad = \quad p_t + P_Q'\left(Q_t, \mathbf{x}_t^D\right) \left[1 + \frac{dQ_{(-i)t}}{dq_{it}}\right] q_{it} \tag{4.3}$$

where $P_Q'\left(Q_t, \mathbf{x}_t^D\right)$ is the derivative of the inverse demand function with respect to total output. Variable $Q_{(-i)t}$ is the aggregate output of firms other than i. The derivative $dQ_{(-i)t}/dq_{it}$ represents the *belief* or *conjecture* that firm i has about how other firms will respond by changing their output when firm i changes marginally its own output. Under the assumption of Nash-Cournot competition, this *belief* or *conjecture* is zero:

$$Nash - Cournot \;\Leftrightarrow\; \frac{dQ_{(-i)t}}{dq_{it}} = 0 \tag{4.4}$$

Firm i takes as fixed the quantity produced by the rest of the firms, $Q_{(-i)t}$, and chooses her own output q_{it} to maximize her profit. Therefore, the first order condition of optimality under Nash-Cournot competition is:

$$MR_{it} \;=\; p_t + P_Q'\left(Q_t, \mathbf{x}_t^D\right)\; q_{it} \;=\; MC_i(q_{it}) \tag{4.5}$$

We assume that the profit function is globally concave in q_{it} for any positive value of $Q_{(-i)t}$, such that there is a unique value of q_{it} that maximizes the firm's profit, and it is fully characterized by the marginal condition of optimality that establishes that marginal revenue equals marginal cost.

Suppose that the demand function has been estimated in a first step such that there is a consistent estimate of demand. Therefore, the researcher can construct consistent estimates of marginal revenues $MR_{it} \equiv p_t + P'_Q\left(Q_t, \mathbf{x}_t^D\right) q_{it}$ for every firm i. Consider the same Cobb-Douglas specification of the cost function as in equation (4.1). Then, the econometric model can be described in terms of the following linear regression model in logarithms:[1]

$$\ln\left(MR_{it}\right) = \theta \ \ln(q_{it}) + \alpha_1 \ \ln(w_{1it}) + ... + \alpha_J \ \ln(w_{Jit}) + \varepsilon_{it}^{MC} \tag{4.6}$$

We are interested in the estimation of the parameters θ and α's, and in the firms' cost inefficiency, ε_{it}^{MC}.

OLS estimation of this regression function suffers from the same endogeneity problem as in the perfect competition case described above. The model implies a negative correlation between a firm's output and its unobserved inefficiency. To deal with this endogeneity problem, we can use instrumental variables. As in the case of perfect competition, we can use observable variables that affect demand but not costs as instruments. With Cournot competition, we may have additional types of instruments, as we explain next.

Suppose that the researcher observes some exogenous input prices $\mathbf{w}_{it} = (w_{1it}, ..., w_{Jit})$ and that at least one of these prices has cross-sectional variation over firms. For instance, suppose that there is information at the firm level on the firm's wage rate, or its capital stock, or its installed capacity. Note that, in equilibrium, the input prices of competitors have an effect on the level of output of a firm. That is, given its own input prices $\mathbf{w}_{it}$, log-output $\ln(q_{it})$ still depends on the input prices of other firms competing in the market, $\mathbf{w}_{jt}$ for $j \neq i$. A firm's output increases if, all else equal, the wage rates of a competitor increase. Note that the partial correlation between $\mathbf{w}_{jt}$ and $\ln(q_{it})$ is a testable condition. Under the assumption that the vector $\mathbf{w}_{jt}$ is exogenous, that is, $\mathbb{E}(\mathbf{w}_{jt}\ \varepsilon_{it}^{MC}) = 0$, a standard approach to estimate this model is using IV or GMM based on moment conditions that use the characteristics of other firms as an instrument for output. For instance, the moment conditions can be:

$$\mathbb{E}\left(\left[\begin{array}{c} \ln(\mathbf{w}_{it}) \\ \sum\limits_{j \neq i} \ln(\mathbf{w}_{jt}) \end{array}\right] \left[\ln\left(MR_{it}\right) - \theta \ \ln(q_{it}) - \ln \mathbf{w}'_{it} \ \alpha\right]\right) = \mathbf{0} \tag{4.7}$$

4.2.2 The nature of competition

Model

Consider an industry where the inverse demand curve is $p_t = P\left(Q_t, \mathbf{x}_t^D\right)$, and firms, indexed by i, have cost functions $C_i(q_{it})$. Every firm i chooses its amount of output, q_{it}, to maximize its profit, $p_t\ q_{it} - C_i(q_{it})$. The marginal condition for the profit maximization implies marginal revenue equals marginal cost. The marginal revenue of firm i has the

[1] For notational simplicity, here I omit the estimation error from the estimation of the demand function in the first step. Note that, in this case, this estimation error only implies measurement error in the dependent variable and it does not affect the consistency of the instrumental variables estimator described below or the estimation of robust standard errors.

expression in equation (4.3). As mentioned above, the term $dQ_{(-i)t}/dq_{it}$ represents the *belief* that firm i has about how the other firms in the market will respond if it changes its own output marginally. We denote this belief as the *conjectural variation* of firm i at period t, and denote it as CV_{it}.

As researchers, we can choose between different assumptions about firms' beliefs or conjectural variations. An assumption about CVs implies a model of competition with its corresponding equilibrium outcomes. Nash (1951) proposed the following conjecture: when a player constructs her best response, she believes that the other players will not respond to a change in her decision. In the Cournot model, the Nash conjecture implies that $CV_{it} = 0$. For every firm i, the "perceived" marginal revenue is $MR_{it} = p_t + P'_Q\left(Q_t\ \mathbf{x}_t^D\right)\ q_{it}$, and the condition $p_t + P'_Q\left(Q_t\ \mathbf{x}_t^D\right)\ q_{it} = MC_i(q_{it})$ implies the Cournot equilibrium.

Similarly, there are assumptions about CVs that generate the perfect competition equilibrium and the collusive or cartel equilibrium.

Perfect competition. For every firm i, $CV_{it} = -1$. A firm believes that if it increases (reduces) its own output in, say, q units, the other firms will respond by reducing (increasing) their output by the same amount such that total market output does not change. That is, a firm believes that it cannot have any influence on total market output. This conjecture implies that: $MR_{it} = p_t$, and the equilibrium conditions $p_t = MC_i(q_{it})$ under perfect competition.

Perfect collusion. For every firm i, $CV_{it} = N_t - 1$. A firm believes that if it increases (reduces) its own output in, say, q units, each of the other firms in the market will imitate this decision, increasing (reducing) its output by the same amount, such that total market output increases (declines) in $N_t q$ units. This conjecture implies that $MR_{it} = p_t + P'_Q\left(Q_t, \mathbf{x}_t^D\right)\ N_t\ q_{it}$, which generates the equilibrium conditions $p_t + P'_Q\left(Q_t, \mathbf{x}_t^D\right)$ $N_t\ q_{it} = MC_i(q_{it})$. When firms have constant and homogeneous MCs, this condition implies $p_t + P'_Q\left(Q_t, \mathbf{x}_t^D\right)\ Q_t = MC_t$, as $Q_t = N_t q_t$, which is the equilibrium condition under monopoly.

The value of the beliefs / CV parameters are related to the nature of competition:

$$\begin{cases} \textit{Perfect competition:} & CV_{it} = -1; \quad MR_{it} = p_t \\ \textit{Nash-Cournot:} & CV_{it} = 0; \quad MR_{it} = p_t + P'_Q\left(Q_t\right)\ q_{it} \\ \textit{Collusion of n firms:} & CV_{it} = n-1; \quad MR_{it} = p_t + P'_Q\left(Q_t\right)\ n\ q_{it} \\ \textit{Perfect collusion:} & CV_{it} = N_t - 1; \quad MR_{it} = p_t + P'_Q\left(Q_t\right)\ Q_t \end{cases} \tag{4.8}$$

The expressions in (**??**) show that firms' beliefs about competitors' behavior, as represented by CVs, are closely related to the nature of competition. Importantly, these results are not making any assumption about how firms' conjectures CV are determined. These beliefs can endogenously determined, together with the other outcomes of the model, price and outputs. However, the model is silent about how CVs are achieved.[2] For the

[2] A possible approach for endogenizing CVs is to consider a dynamic game with multiple periods $t = 1, 2, \ldots$ where every period firms choose their CVs and their amounts of output. Firms can learn

moment, we do not specify the determinants of CVs, but it is important to keep in mind that they are endogenous objects. Interpreting CV_{it} as an exogenous parameter is not correct. Conjectural variations represent firms' beliefs, and as such they are endogenous outcomes from the model.

Some applications or views of the *conjectural variations model* go beyond the results above for perfect competition, Cournot, and collusion, and consider that CVs can take any continuous value between -1 and $N_t - 1$. Under this interpretation, if CV is negative, the degree of competition is stronger than Cournot, and the closer to -1, the more competitive. If CV is positive, the degree of competition is weaker than Cournot, and the closer to $N_t - 1$, the less competitive. It seems reasonable to expect that CV should not be smaller than -1 or greater than $N_t - 1$. Values smaller than -1 imply a competitors' respond that generates negative profits. Values greater than $N_t - 1$ imply that the cartel is not maximizing the joint profit.[3] However, this view of the conjectural variations approach has been criticized as these "intermediate values" of CVs cannot be obtained as equilibrium values of a dynamic game. See Corts (1999) for an analysis of this issue that has been influential in empirical IO.

Estimation with information on marginal costs

Consider a homogeneous product industry and a researcher with data on firms' quantities and marginal costs, and on market prices over T periods of time: $\{p_t,\, MC_{it},\, q_{it}\}$ for $i = 1,2,...,N_t$ and $t = 1,2,...,T$. Under the assumption that every firm chooses the amount of output that maximizes its profit given its belief CV_{it}, we have that the following condition holds:

$$p_t + P'_Q \left(Q_t\ \mathbf{x}_t^D\right)\ [1 + CV_{it}]\ q_{it} = MC_{it} \tag{4.9}$$

And solving for the conjectural variation, we have:

$$CV_{it} = \frac{p_t - MC_{it}}{-P'_Q \left(Q_t.\mathbf{x}_t^D\right)\ q_{it}} - 1 = \left[\frac{p_t - MC_{it}}{p_t}\right]\ \left[\frac{1}{q_{it}/Q_t}\right]\ \eta_t - 1 \tag{4.10}$$

where η_t is the price elasticity of demand (in absolute value): that is, $\eta_t = -(p_t/Q_t)$ $(1/P'_Q(Q_t, \mathbf{x}_t^D))$. Note that $(p_t - MC_{it})/p_t$ is the Lerner index, and q_{it}/Q_t is the market share of firm i. This equation shows that, given data on output, price, demand, and marginal cost, we can identify a firm's belief that is consistent with these data and profit maximization.

Let us denote $\left[\frac{p_t - MC_{it}}{p_t}\right]\ \left[\frac{1}{q_{it}/Q_t}\right]$ as the Lerner-index-to-market-share ratio of a firm. If this ratio is close to zero, then the estimated value of CV is close to -1 unless the absolute demand elasticity is large. In contrast, if the Lerner-index-to-market-share ratio is large (that is, larger than the inverse demand elasticity), then the estimate of CV is greater than zero, and the researcher can reject the hypothesis of Cournot competition in favor of some collusion.

---

over time and update their beliefs CV. In the equilibrium of this dynamic game, CVs are determined endogenously. We present this type of dynamic game in chapter 8.

[3] Nevertheless, in the context of dynamic games, we could have values of CV that can be smaller than -1 due to competitive wars that try to induce other firms' exit from the market. For instance, see the dynamic game in Beviá, Corchón, and Yasuda (2020).

Under the restriction that all the firms have both the same marginal costs and conjectural variations, equation (4.10) becomes:

$$\frac{p_t - MC_t}{p_t} = \left[\frac{1+CV_t}{N_t}\right] \frac{1}{\eta_t} \tag{4.11}$$

where N_t is the number of firms in the market. This is the equation that we use in the empirical application that we describe at the end of this section. According to this expression, market power, as measured by the Lerner Index, is related to the elasticity of demand (negatively), the number of firms in the market (negatively), and the conjectural variation (positively). Importantly, one *should not* interpret equation (4.11) as a causal relationship where the Lerner index (in the left hand side) depends on exogenous variables in the right hand side. In this equation, all the variables – Lerner index, conjectural variation, and number of firms – are endogenous, and are jointly determined in the equilibrium of this industry as functions of exogenous variables affecting demand and costs. Nevertheless, equation 4.11 is still a very useful equation for empirical analysis and estimation, as it determines the value of one of the endogenous variables once we have measures for the others.

Estimation without information on marginal costs

So far, we have considered the estimation of CV parameters when the researcher knows both demand and firms' marginal costs. We now consider the case where the researcher knows the demand, but it does not know firms' marginal costs. Identification of CVs requires also the identification of marginal costs. We show here that, under some conditions, we can jointly identify CVs and MCs using the marginal condition of optimality and demand.

The researcher observes data $\left\{p_t,\, q_{it},\, \mathbf{x}_t^D,\, \mathbf{w}_t :\; i = 1,...N_t;\, t = 1,...,T\right\}$, where $\mathbf{x}_t^D$ are exogenous variables affecting consumer demand, for instance, average income or population, and $\mathbf{w}_t$ are variables affecting marginal costs, for instance, some input prices. Consider the linear (inverse) demand equation:

$$p_t \;=\; \alpha_0 + \mathbf{x}_t^D \alpha_1 \; - \alpha_2\, Q_t + \varepsilon_t^D \tag{4.12}$$

with $\alpha_2 \geq 0$, and ε_t^D is unobservable to the researcher. Consider the marginal cost function:

$$MC_{it} \;=\; \beta_0 + \mathbf{w}_t \beta_1 + \beta_2\, q_{it} + \varepsilon_{it}^{MC} \tag{4.13}$$

with $\beta_2 \geq 0$, and ε_{it}^{MC} is unobservable to the researcher. Profit maximization implies the marginal condition $p_t + dP_t/dQ_t\,[1+CV_{it}]\; q_{it} = MC_{it}$. Since the demand function is linear and $dP_t/dQ_t = -\alpha_2$, we can write the marginal condition as follows:

$$p_t \;=\; \beta_0 + \mathbf{w}_t \beta_1 + [\beta_2 + \alpha_2(1+CV_{it})]\; q_{it} + \varepsilon_{it}^{MC} \tag{4.14}$$

This model is typically completed with the assumption that conjectural variations are constant over time: $CV_{it} = CV_i$. The assumption of CV constant over time is plausible when the industry is mature and has not experienced structural changes during the sample period. Nevertheless, some empirical studies analyze specific events, such as important regulatory changes or mergers, and allow CV to be different before and after this change. Some empirical applications impose also the restriction that CV is the same for all the firms in the market. This assumption of homogeneous CVs across firms is not always plausible. For instance, there may be leaders and followers in the industry, or cartels that include only some firms. Furthermore, this restriction is not necessary if the researcher has data on output at the firm level, q_{it}, and not only on total market output, Q_t. Accordingly, in this section we assume that firms' beliefs/conjectures are constant over time but can vary across firms.

The structural equations of the model are the demand equation in (4.12) and the equilibrium condition in (4.14). Using this model and data, can we identify (that is, estimate consistently) the CV parameter? Without further restrictions, the answer to this question is negative. However, we show below that a simple and plausible condition in this model implies the identification of both CV and MC parameters. We first describe the identification problem.

Identification of demand parameters. The estimation of the regression equation for the demand function needs to deal with the well-known simultaneity problem. In equilibrium, output Q_t is correlated with the error term ε_t^D. The model implies a valid instrument to estimate demand. In equilibrium, Q_t depends on the exogenous cost variables $\mathbf{w}_t$. This variable does not enter in the demand equation. If $\mathbf{w}_t$ is not correlated with ε_t^D, then this variable(s) satisfies all the conditions of a valid instrument. Parameters α_0, α_1, and α_2 are therefore identified using this IV estimator.

Identification of CV and MCs. In the regression equation (4.14), we also need to deal with an endogeneity problem. In equilibrium, output q_{it} is correlated with the error term ε_{it}^{MC}. The model implies a valid instrument to estimate this equation. In equilibrium, q_{it} depends on the exogenous demand shifters $\mathbf{x}_t^D$. Note that $\mathbf{x}_t^D$ does not enter in the marginal cost and in the right hand side of the regression equation (4.14). If $\mathbf{x}_t^D$ is not correlated with ε_{it}^{MC}, then this variable satisfies all the conditions for being a valid instrument such that the parameters β_0, β_1, and $\gamma_i \equiv \beta_2 + \alpha_2(1 + CV_i)$ are identified using this IV estimator.

Now, the identification of parameter $\gamma_i \equiv \beta_2 + \alpha_2(1 + CV_i)$ and of the slope of the inverse demand function, α_2, is not sufficient to identify separately CV_i and the slope of the marginal cost function, β_2. That is, given known values for γ_i and α_2, equation

$$\gamma_i \;=\; \beta_2 \;+\; \alpha_2\;(1 + CV_i) \tag{4.15}$$

implies a linear relationship between CV_i and β_2 and there are infinite values of these parameters that satisfy this restriction. Even if we restrict CV_i to belonging to the values consistent with an equilibrium concept, such that $CV_i \in \{-1, 0, N-1\}$ and β_2 to being greater or equal than zero, we do not have point identification of these parameters. For instance, suppose that $N = 2$, $\gamma_i = 2$, and $\alpha_2 = 1$ such that equation (4.15) becomes $2 = \beta_2 + (1 + CV_i)$, or equivalently, $\beta_2 + CV_i = 1$. This equation is satisfied by any of the following forms of competition and values of $\beta_2 \geq 0$: perfect competition, with $CV_i = -1$ and $\beta_2 = 2$; Cournot competition, with $CV_i = 0$ and $\beta_2 = 1$; and perfect collusion, with $CV = 1$ and $\beta_2 = 0$.

This identification problem has an intuitive interpretation. The identified parameter γ_i captures the true causal effect of firm i's output on market price. There are two different channels for this causal effect: through the change in marginal cost; and through the change in marginal revenue, that depends on the firm's conjectural variation.

Identification of the causal effect parameter γ_i is not sufficient to disentangle the relative contribution of the two channels.

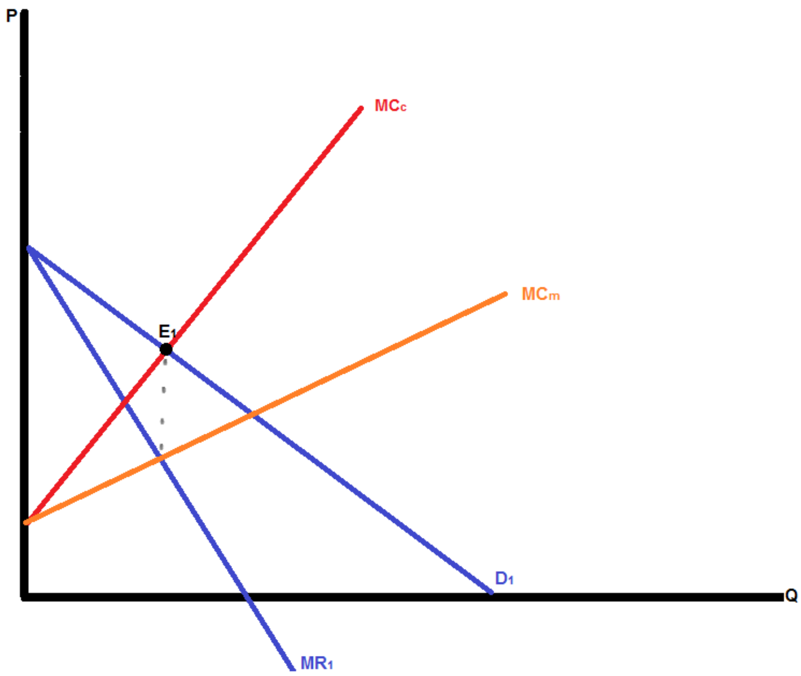

Figure 4.1: One data point: No identification of competition (c) vs. collusion (m)

Following Bresnahan (1981), we can provide a graphical representation of this identification problem. Suppose that we have followed the approach described above to estimate consistently the demand parameters – that imply the demand curve D_1 and the monopoly marginal revenue curve MR_1 in figure 4.1 – the marginal cost parameters β_0 and β_1, and the parameter γ_i. We can define two hypothetical marginal cost functions: the marginal cost under the hypothesis of perfect competition ($CV_i = -1$ such that $\beta_2 = \gamma_i$), $MC_c = \beta_0 + \mathbf{w}\beta_1 + \gamma_i q$; and the marginal cost under the hypothesis of monopoly or perfect collusion ($CV = N - 1$ such that $\beta_2 = \gamma_i - \alpha_2 N$), $MC_m = \beta_0 + \mathbf{w}\beta_1(\gamma_i - \alpha_2 N)$ q. That is, MC_c and MC_m are the marginal cost functions that rationalize the observed price and output under the hypotheses of perfect competition and monopoly, respectively. Figure 4.1 shows that the observed price and quantity – represented by the point $E_1 = (q_1, p_1)$ – can be rationalized either as the point where the demand function D_1 crosses the competitive marginal cost MC_c, or as the monopoly outcome defined by the marginal revenue MR_1 and the monopoly marginal cost MC_m.

Data on prices and quantities at multiple time periods do not help to solve this identification problem. This is illustrated in Figure 4.2. Consider the demand curves D_1 and D_2 at periods $t = 1$ and $t = 2$, respectively. Importantly, under the demand function in equation (4.12), every change in the demand curve (that is, a change in $\mathbf{x}_t^D$ or in ε_t^D) implies a parallel vertical shift, keeping the slope constant. Therefore, demand curves D_1 and D_2 are parallel, and so they are the corresponding marginal revenue curves MR_1 and MR_2, as shown in Figure 4.2. Again, the observed points $E_1 = (q_1, p_1)$ and $E_2 = (q_2, p_2)$ can be rationalized either as perfectly competitive equilibria that come from the intersection of demand curve D_t and marginal cost MC_c, or as monopoly

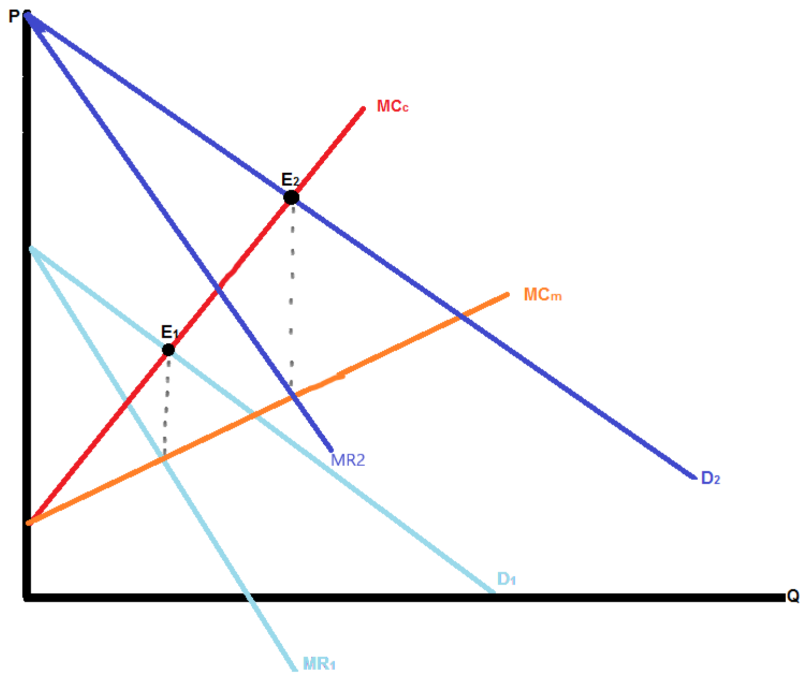

Figure 4.2: Multiple data points: No identification of competition (c) vs. collusion (m)

outcomes that are determined by the intersection of the marginal revenue curve MR_t and the marginal cost MC_m.

This graphical analysis provides also an intuitive interpretation of a solution to this identification problem. This solution involves generalizing the demand function so that changes in exogenous variables do more than just a parallel shift in the demand curve and the marginal revenue. We introduce additional exogenous variables that are capable of **rotating** the demand curve. Consider Figure 4.3. Now, the demand curve D_2 represents a rotation of demand curve D_1 around point E_1. Under perfect competition, this rotation in the demand curve should not have any effect in equilibrium prices and quantities. Therefore, under perfect competition, E_1 is the equilibrium point under the two demand curves. This is not the case under monopoly (collusion). When firms have market power, a change in the slope of the demand has an effect on prices and quantities. Therefore, given demand curves D_1 and D_2, if the data shows different values of the (quantity, price) points E_1 and E_2, as in Figure 4.3, we can reject the hypothesis of perfect competition in favor of firms having market power.

We now present more formally the identification of the model illustrated in Figure 4.3. Consider now the following demand equation:

$$p_t \;=\; \alpha_0 + \mathbf{x}_t^D \alpha_1 - \alpha_2\, Q_t - \alpha_3\, [z_t\; Q_t] + \varepsilon_t^D \tag{4.16}$$

Variable z_t is observable to the researcher and affects the slope of the demand. Some possible candidates for these variables are the price of a substitute or complement product, seasonal dummies capturing changes in the composition of the population of consumers during the year, or consumer demographics. The key condition is that the parameter α_3 is different from zero. That is, when z_t varies, there is a rotation in the demand curve. Note that this condition is testable. Given this demand model, we have

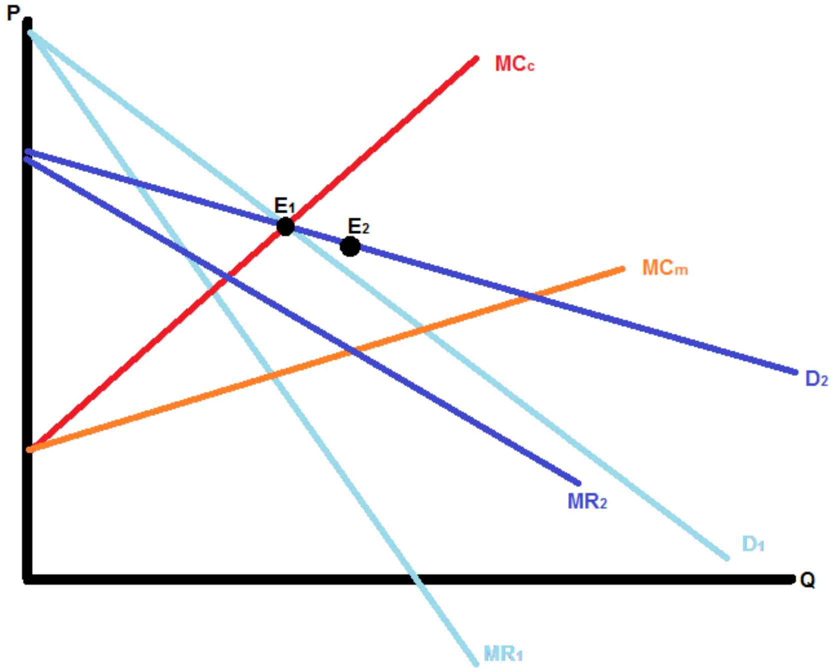

Figure 4.3: Rotating demand curve: Rejecting perfect competition

that the slope of the demand curve is $dP_t/dQ_t = -\alpha_2 - \alpha_3\ z_t$, and the marginal condition for profit maximization implies the following regression model:

$$p_t \ = \ \beta_0 + \mathbf{w}_t \beta_1 + \gamma_{1,i}\ q_{it} + \gamma_{2,i}\ (z_t\ q_{it}) + \varepsilon_{it}^{MC} \tag{4.17}$$

with $\gamma_{1,i} \equiv \beta_2 + \alpha_2\,[1 + CV_i]$ and $\gamma_{2,i} \equiv \alpha_3\,[1 + CV_i]$.

Equations (4.16) and (4.17) describe the structural model. Using this model and data, we now show that we can separately identify CV_i and the slope of the marginal cost, $\beta 2$. Demand parameters can be identified similarly as before, using $\mathbf{w}_t$ as instruments for output. Parameters α_0, α_1, α_2, and α_3 are identified using this IV estimator. The model also implies valid instruments to estimate the parameters in the equilibrium equation in (4.17). We can instrument q_{it} using $\mathbf{x}_t^D$ and $z_t\ q_{it}$ using $z_t\ \mathbf{x}_t^D$. Parameters β_0, β_1, $\gamma_{1,i}$, and $\gamma_{2,i}$ are identified. Note that:

$$\left\{ \begin{array}{rcl} \gamma_{1,i} & = & \beta_2 + \alpha_2\,[1 + CV_i] \\ \\ \gamma_{2,i} & = & \alpha_3\,[1 + CV_i] \end{array} \right. \tag{4.18}$$

Given estimates of α_2, α_3, $\gamma_{1,i}$, and $\gamma_{2,i}$, we have that (4.18) is a system of two equations with two unknowns (β_2 and CV_i) that has a unique solution if and only if α_3 is different to zero. The solution of this system implies that $1 + CV_i = \gamma_{2,i}/\alpha_3$. The conjectural variation is identified by the ratio between the sensitivity of price with respect to $(z_t\ q_{it})$ in the equilibrium equation and the sensitivity of price with respect to $(z_t\ Q_t)$ in the demand equation.

The sample variation in the slope of the inverse demand plays a key role in the identification of the CV parameter. An increase in the slope means that the demand becomes less price sensitive, more inelastic. For a monopolist, when the demand becomes more inelastic, the optimal price should increase. In general, for a firm with a high level of market power (high CV), we should observe an important increase in prices associated with an increase in the slope. On the contrary, if the industry is characterized by very low market power (low CV), the increase in price should be practically zero. Therefore, the response of price to an exogenous change in the slope of the demand contains key information for the estimation of CV.

This identification result still holds if the exogenous variable z_t that generates the change in the slope of the demand curve also also an effect in the marginal cost. That is, z_t can be included in the vector of cost shifters $\mathbf{w}_t$. However, a key identifying restriction that cannot relaxed in this approach is that z_t cannot affect the slope of the marginal cost function. That is, variation in z_t does not affect the degree of diseconomies of scale in the production of the good.

An application: The sugar industry

Genesove and Mullin (1998) (GM) study competition in the US sugar industry during the period 1890-1914. One of the purposes of this study is to test the validity of the conjectural variation approach by focusing on an industry where firms' marginal costs can be very accurately measured. This motivation plays an important role in the authors' selection of this industry and historical period. During this period, the production technology of refined sugar was very simple, and the marginal cost function was characterized in terms of a simple linear function of the cost of raw sugar, the main intermediate input in the production of refined sugar. Furthermore, during this period there was an investigation of the industry by the US antitrust authority. As a result of that investigation, there were reports from multiple expert witnesses who provided a very coherent description of the structure and magnitude of production costs in this industry. GM use this information on marginal costs to test the validity of the standard conjectural variation approach for the estimation of price cost margins and marginal costs.

Let $p_t = P(Q_t, S_t)$ be the inverse demand function at year t in the industry, where S_t represents exogenous variables affecting demand, and that we specify below. Under the assumption that all the firms are identical in their marginal costs and in their conjectural variations, the marginal revenue at period t is:

$$MR_t \;=\; p_t - P'(Q_t, S_t)\; [1 + CV_t]\; \frac{Q_t}{N_t} \tag{4.19}$$

where $P'(Q_t, S_t)$ is the slope of the demand curve. The condition for profit maximization (marginal revenue equals marginal cost) implies the following relationship between the Lerner Index and the conjectural variation:

$$\frac{p_t - MC_t}{p_t} = \left[\frac{1 + CV_t}{N_t}\right] \frac{1}{\eta_t} \tag{4.20}$$

where η_t is the price demand elasticity, in absolute value.

Given equation 4.20, if we observe price and marginal cost and we can estimate the demand elasticity, then there is a simple and direct estimate of the conjectural variation. Without knowledge of the marginal cost, the estimation of the CV should depend on two conditions: (a) the existence of an observable exogenous variable that rotates the demand curve; and (b) the exclusion restrictions that observable demand shifters do not affect marginal costs. If assumptions (a) or (b) are not correct, our estimation of the CV (and of the Lerner Index) will be biased. GM evaluate these assumptions by comparing the estimate of CV obtained under conditions (a) and (b) (and without using data on marginal costs) with the direct estimate of this parameter using data on marginal costs.

The rest of this section describes the following aspects of this empirical application: (i) the industry; (ii) the data; (iii) estimation of demand parameters; (iv) predicted markups under different conduct parameters; and (v) estimation of CV.

(i) The industry

During 1890-1914, refined sugar was a homogeneous, and the industry in the US was highly concentrated. The industry leader, the *American Sugar Refining Company (ASR)*, had more than 65% of the market share during most of these years.[4]

Production technology. Refined sugar companies bought raw sugar from suppliers in national and international markets, transform it into refined sugar, and sell it to grocers. They sent sugar to grocers in barrels, without any product differentiation. Raw sugar is 96% sucrose and 4% water. Refined sugar is 100% sucrose. The process of transforming raw sugar into refined sugar was called "melting", and it consisted of eliminating the 4% of water in raw sugar. Industry experts reported that firms in the industry used a fixed coefficient (or Leontieff) production technology that can be described by the following production function:

$$Q_t \ = \ \min\{\lambda \ Q_t^{raw} \ ; \ f(L_t, K_t)\} \tag{4.21}$$

where Q_t is refined sugar output, Q_t^{raw} is the input of raw sugar, $\lambda \in (0,1)$ is a technological parameter, and $f(L_t, K_t)$ is a function of labor and capital inputs. Production efficiency and cost minimization imply that $Q_t = \lambda \ Q_t^{raw} = f(L_t, K_t)$. That is, 1 ton of raw sugar generates λ tons units of refined sugar. Since raw sugar is only 96% sucrose, the largest possible value of λ is 0.96. Industry experts at that time unanimously reported that there was some loss of sugar in the refining process such that the value of the parameter λ was close to 0.93.

Marginal cost function. For this production technology, the marginal cost function is:

$$MC_t \ = \ c_0 + c_1 \ p_t^{raw} + c_2 \ q_t \tag{4.22}$$

where c_0, c_1, and c_2 are parameters, p_t^{raw} is the price of raw sugar in dollars per pound, and q_t is output per firm. The Leontieff production function in equation (4.21) implies that $c_2 = 0$, $c_1 = 1/\lambda$, and c_0 is a component of the marginal cost that depends on labor. According to industry experts, during the sample period the values of the parameters in the marginal cost were $c_0 = \$0.26$ per pound, $c_1 = 1/\lambda = 1/0.93 = 1.075$, and $c_2 = 0$. Therefore, the marginal cost at period (quarter) t, in dollars per pound of sugar, was:

$$MC_t \ = \ 0.26 + 1.075 \ p_t^{raw} \tag{4.23}$$

[4] ASR operated one of the world's largest sugar refineries at that time, the Domino Sugar Refinery in Brooklyn, New York. The ASR company became known as Domino Sugar in 1900.

(ii) The data
The dataset contains 97 quarterly observations on industry output, price, price of raw sugar, imports of raw sugar, and a seasonal dummy.

$$\text{Data} = \{\ Q_t,\ p_t,\ p_t^{raw}, IMP_t, S_t : t = 1,2,...,97\} \tag{4.24}$$

IMP_t represents imports of raw sugar from Cuba, and S_t is a dummy variable for the Summer season: $S_t = 1$ if observation t is a Summer quarter, and $S_t = 0$ otherwise. The summer was a high demand season for sugar because most the production of canned fruits was concentrated during that season, and the canned fruit industry accounted for an important fraction of the demand of sugar.

(iii) Estimation of demand parameters
GM estimate four different models of demand: linear, quadratic, log-linear, and exponential. The main results are consistent for the four models. Here we concentrate on results using the linear (inverse) demand function:

$$p_t \ = \ \alpha_0 + \alpha_1\ S_t - \alpha_2\ Q_t - \alpha_3\ S_t Q_t + \varepsilon_t^D \tag{4.25}$$

Parameters α_0 and α_2 represent the intercept and slope of the demand curve during the "Low season" (when $S_t = 0$). Similarly, parameters $\alpha_0 + \alpha_1$ and $\alpha_2 + \alpha_3$ are the intercept and slope of the demand curve in the "High season" (when $S_t = 1$).

As we have discussed before, Q_t is an endogenous regressor in this regression equation. We need to use IV to deal with this endogeneity problem. In principle, it seems that we could use p_t^{raw} as an instrument. However, GM have a reasonable concern about the validity of this instrument. The demand of raw sugar from the US accounts for a significant fraction of the world demand of raw sugar. Therefore, shocks in the US domestic demand of refined sugar, as represented by ε_t^D, can generate an increase in the world demand of raw sugar and in p_t^{raw} such that p_t^{raw} and ε_t^D can be positively correlated. Instead, GM use imports of raw sugar from Cuba as an instrument. Almost 100% of the production of raw sugar in Cuba was exported to the US, and the authors claim that variations in Cuban production of raw sugar was driven by supply/weather conditions and not by the demand from the US.

Table 4.1: Genesove and Mullin: Demand estimates
Based on Table 3 (column 2) in Genesove and Mullin (1998)

Parameter	Estimate	Standard Error
Intercept Low, α_0	5.813	(0.826)
Intercept High, $\alpha_0 + \alpha_1$	7.897	(1.154)
Slope Low, α_2	0.434	(0.194)
Slope High, $\alpha_2 + \alpha_3$	0.735	(0.321)
Average elasticity Low, η_L	2.24	
Average elasticity High, η_H	1.04	

Table 4.1 presents parameter estimates of demand parameters. In the high season, the demand shifts upwards by \$2.09 per ton ($\alpha_1 = \$2.09 > 0$) and becomes steeper ($\alpha_3 = 0.301 > 0$). The estimated price elasticities of demand in the low and the high season are $\eta_L = 2.24$ and $\eta_H = 1.04$, respectively. According to this, any model of oligopoly competition where firms have some market power predicts that the price cost margin should increase during the high season due to the lower price sensitivity of demand.

(iv) Predicted markups under different conduct parameters

Before we discuss the estimates of the conjectural variation parameter, it is interesting to illustrate the errors that researchers can make when – in the absence of information about marginal costs – they estimate price cost margins by making an incorrect assumption about the value of CV in the industry.

As mentioned above, the industry was highly concentrated during this period. Though there were approximately 6 firms active during most of the sample period, one of the firms accounted for more than two-thirds of total output. Consider three different researchers investigating this industry, that we label as researchers M, C, and S. These researchers do not know the true marginal cost and they have different views about the nature of competition in this industry.Researcher M considers that the industry was basically a Monopoly/Cartel during this period.[5] Therefore, she assumes that $[1+CV]/N = 1$. Researcher C considers that the industry can be characterized by Cournot competition between the 6 firms, such that $[1+CV]/N = 1/6$. Finally, researcher S thinks that this industry can be better described by a Stackelberg model with 1 leader and 5 Cournot followers, and therefore $[1+CV]/N = 1/(2*6-1) = 1/11$.

Table 4.2: Genesove and Mullin: Markups under different conduct parameters

Assumption	Predicted Lerner Low season $\frac{1+CV}{N\,\eta_L}$	Actual Lerner Low season $\frac{p_L-MC}{p_L}$	Predicted Lerner High season $\frac{1+CV}{N\,\eta_H}$	Actual Lerner High season $\frac{p_H-MC}{p_H}$
Monopoly: $\frac{1+CV}{N} = 1$	44.6%	3.8%	96.1%	6.5%
Cournot: $\frac{1+CV}{N} = \frac{1}{6}$	7.4%	3.8%	16.0%	6.5%
Stackelberg: $\frac{1+CV}{N} = \frac{1}{11}$	4.0%	3.8%	8.7%	6.5%

Table 4.2 presents the predictions of the Lerner index – in the low and high season – from these three researchers and also the actual value of the Lerner index based on our information on marginal costs. Researcher M makes a very seriously biased prediction of market power. Since the elasticity of demand is quite low in this industry, especially during the high season, the assumption of Cartel implies a very high Lerner index, much

[5] In fact, there was an anti-trust investigation, such that there were some suspicions of collusive behavior.

higher than the actual one. Researcher C also over-estimates the actual Lerner index. The estimates of researcher S are only slightly upward biased.

Consider the judge of an anti-trust case in which there is not reliable information on the actual value of MCs. The picture of industry competition that this judge gets from the three researchers is very different. This judge would be interested in measures of market power in this industry that are based on a scientific estimate of the conjectural parameter.

(iv) Estimation of conjectural variation

Now, GM consider the hypothetical scenario where the researcher does not observe the marginal cost and applies the method described above to jointly estimate CV and marginal cost parameters, c_0, c_1, and c_2. The marginal condition for profit maximization implies the following equation:

$$p_t \ = \ c_0 + c_1 \ p_t^{raw} + \gamma_1 \ Q_t + \gamma_2 \ S_t Q_t + \varepsilon_t^{MC} \tag{4.26}$$

with $\gamma_1 \equiv [c_2 + \alpha_2(1+CV)]/N$, and $\gamma_2 \equiv \alpha_3(1+CV)]/N$. We treat c_0 and c_1 as parameters to estimate because we consider the estimation of CV under the hypothetical situation where the researcher does not know that $c_0 = 0.26$, $c_1 = 1.075$, and $c_2 = 0$.

Since Q_t is endogeneously determined, it should be correlated with ε_t^{MC}. To deal with this endogeneity problem, GM use instrumental variables. Again, they use imports from Cuba as an instrument for Q_t. Table 4.3 presents the IV estimates of c_0, c_1 and $(1+CV)/N$ and their standard errors (in parentheses). For comparison, we also include the "true" values of these parameters based on the information on marginal costs.

Table 4.3: Genesove and Mullin:
Estimates of conduct and marginal cost parameters

Parameter	Estimate (s.e.)	"True" value
$(1+CV)/N$	0.038 (0.024)	0.10
c_0	0.466 (0.285)	0.26
c_1	1.052 (0.085)	1.075

The estimates of $(1+CV)/N$, c_0, and c_1, are not too far from their "true" values. This seems to validate the CV approach for this particular industry and historical period. Based on this estimate of $(1+CV)/N$, the predicted values for the Lerner index is $0.038/2.24 = 1.7\%$ in the low season, and $0.038/1.04 = 3.6\%$ in the high season. Remember that the true values of the Lerner index using information on marginal costs were 3.8% in the low season and 6.5% in the high season. Therefore, the estimates using the CV method only slightly under-estimate the actual market power in the industry. Furthermore, using either information on marginal costs or the CV method, we can clearly reject the null hypothesis of a perfect cartel, that is, $(1+CV)/N = 1$

4.3 Differentiated product industry

4.3.1 Model

Consider an industry with J differentiated products, for instance, automobiles, indexed by $j \in \mathscr{J} = \{1, 2, ..., J\}$. Consumer demand for each of these products can be represented using the demand system:

$$q_j = D_j(\mathbf{p}, \mathbf{x}) \quad \text{for} \, j \in \mathscr{J} \tag{4.27}$$

where $\mathbf{p} = (p_1, p_2, ..., p_J)$ is the vector of prices, and $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_J)$ is a vector of other product attributes. There are F firms in the industry, indexed by $f \in \{1, 2, ..., F\}$. Each firm f owns a subset $\mathscr{J}_f \subset \mathscr{J}$ of the brands. The profit of firm f is the sum of the profits from each product it owns. That is:

$$\Pi_f = \sum_{j \in \mathscr{J}_f} p_j \, q_j - C_j(q_j) \tag{4.28}$$

where $C_j(q_j)$ is the cost of producing a quantity q_j of product j. Firms compete in prices.

(i) Nash-Bertrand competition

We start with the case where firms compete in prices a la Nash-Bertrand. Each firm chooses its own prices to maximize profits and takes the prices of other firms as given. The first order conditions of optimality for profit maximization of firm f are: for any $j \in \mathscr{J}_f$

$$q_j + \sum_{k \in \mathscr{J}_f} [p_k - MC_k] \, \frac{\partial D_k}{\partial p_j} = 0 \tag{4.29}$$

where MC_j is the marginal cost $C_j'(q_j)$. We can write this system in vector form. Let $\mathbf{q}^f$, $\mathbf{p}^f$, and $\mathbf{MC}^f$ be the column vectors with the quantities, prices, and marginal costs, respectively, for every product $j \in \mathscr{J}_f$. And let $\Delta\mathbf{D}^f$ be the square Jacobian matrix with the demand-price derivatives $\partial D_k / \partial p_j$ for every $j, k \in \mathscr{J}_f$. Then, the system of optimality conditions for firm f has the following vector form:

$$\mathbf{q}^f + \Delta\mathbf{D}^f \left[\mathbf{p}^f - \mathbf{MC}^f \right] = 0 \tag{4.30}$$

Under the condition that the Jacobian matrix is non-singular, we can solve for price-cost margins in this system:

$$\mathbf{p}^f - \mathbf{MC}^f = - \left[\Delta\mathbf{D}^f \right]^{-1} \mathbf{q}^f \tag{4.31}$$

The right-hand-side of this equation depends only on demand parameters, and not on costs. Given an estimated demand system and an ownership structure of brands, the vector of Price-Cost Margins under Nash-Bertrand competition is known to the researcher.

(ii) Example: Single product firms with Logit demand

For single product firms, the marginal condition of optimality is:

$$p_j - MC_j = -\left[\frac{\partial D_j}{\partial p_j}\right]^{-1} q_j \tag{4.32}$$

In the logit demand system, we have that:

$$D_j(\mathbf{p}, \mathbf{x}) = H \frac{\exp\left\{\mathbf{x}_j'\beta - \alpha\ p_j\right\}}{1 + \sum_{k=1}^{J} \exp\left\{\mathbf{x}_k'\beta - \alpha\ p_k\right\}} \tag{4.33}$$

where H represents market size, and β and α are parameters. This logit demand system implies that $\partial D_j / \partial p_j = -\alpha\ H\ s_j(1 - s_j)$ where s_j is the market share $s_j \equiv q_j / H$. Therefore, in this model:

$$PCM_j \equiv p_j - MC_j = \frac{1}{\alpha(1 - s_j)} \tag{4.34}$$

We see that in this model the price-cost margin of a firm declines with the price sensitivity of demand, α, and increases with the own market share, s_j.

(iii) Example: Multi-product firms with Logit demand

In the logit demand system, we have that $\partial D_j / \partial p_j = -\alpha\ H\ s_j(1 - s_j)$, and for $k \neq j$, $\partial D_j / \partial p_k = \alpha\ H\ s_j\ s_k$. Plugging these expressions into the first order conditions of optimality in equation (4.29), we get:

$$PCM_j \quad = \quad \frac{1}{\alpha} + \sum_{k \in \mathscr{J}_f} PCM_k\ s_k \tag{4.35}$$

The right-hand-side is firm-specific but it does not vary across products within the same firm. This condition implies that all the products owned by a firm have the same price-cost margin. According to this condition, the price-cost margin is:

$$PCM_j = \overline{PCM}_f = \frac{1}{\alpha\left(1 - \sum_{k \in \mathscr{J}_f} s_k\right)} \tag{4.36}$$

For the Logit demand model, a multi-product firm charges the same price-cost margin for all of its products. This prediction does not extend to more general demand systems.

(iv) Owning multiple products implies higher price-cost margins

In the logit model, the difference between the price-cost margins of a multi-product and a single-product firm is:

$$\frac{1}{\alpha\left(1 - \sum_{k \in \mathscr{J}_f} s_k\right)} - \frac{1}{\alpha\left(1 - s_j\right)} > 0 \tag{4.37}$$

which is always positive. This prediction extends to a general demand system as long as products are substitutes. For a general demand system, the marginal condition for multi-product firm f and product j can be written as:

$$\begin{aligned} PCM_j &= \left[\frac{-\partial D_j}{\partial p_j}\right]^{-1} q_j \\ &+ \left[\frac{-\partial D_j}{\partial p_j}\right]^{-1} \left[\sum_{k \in J_f;\, k \neq j} PCM_k \frac{\partial D_k}{\partial p_j}\right] \end{aligned} \tag{4.38}$$

In the right-hand-side, the first term is the price-cost margin of a single-product firm. When products are substitutes, the second term is positive because $\partial D_k / \partial p_j > 0$ and $PCM_k > 0$ for every $k \neq j$. Selling multiple products contributes to increasing the price-cost margin of each of the products. This has an intuitive interpretation in terms of a multi-product firm's concern for the cannibalization of its own products. A reduction in the price of product j implies stealing market share from other competing firms, but also cannibalizing market share of the firm's own products other than j.

(v) Collusion and nature of competition

In the homogeneous product case, we have represented the *nature of competition* using firms' conjectural variations or beliefs CV. In section 4.3.5 below, we present the conjectural variation approach in the context of this model with differentiated products. For the moment, we consider here a different representation of the *nature of competition*. We can represent a collusive setting – or the nature of competition – as a $F \times F$ matrix Θ of zeroes and ones. Element (f, g) in this matrix, that we represent as $\theta_{f,g}$, is a dummy variable that equals one if firm f believes that it is colluding with firm g, and it zero otherwise. Of course, all the elements in the diagonal of Θ are ones. But other than this, there are not other restrictions in this matrix. For instance, the matrix can be asymmetric if some firms have not been able to coordinate their collusion beliefs. If there is no collusion at all in the industry, Θ is the identity matrix. The other extreme case is when all the firms in the industry form a cartel: in this case, Θ is a matrix of ones. This representation of the nature of competition can be extended to allow the elements $\theta_{f,g}$ to be real numbers in the interval $[0,1]$ such that they can be interpreted as probabilistic beliefs or as the degree collusion.

A firm f chooses the prices of its own products to maximize the profit of its collusion rink, that has the following expression:

$$\Pi_f^{\Theta} = \sum_{g=1}^{F} \theta_{f,g} \sum_{j \in \mathscr{J}_g} \left[p_j\, q_j - C_j(q_j)\right] \tag{4.39}$$

The marginal condition of optimality for firm f and product $j \in \mathscr{J}_f$ is:

$$q_j + \sum_{g=1}^{F} \theta_{f,g} \sum_{j \in \mathscr{J}_g} \left[p_k - MC_k\right] \frac{\partial D_k}{\partial p_j} = 0 \tag{4.40}$$

In vector form, using all the J products, we have:

$$\mathbf{q} + \Theta^* \, \Delta\mathbf{D}\, \mathbf{PCM} = 0. \tag{4.41}$$

$\mathbf{q}$ and $\mathbf{PCM}$ are $J \times 1$ vectors of quantities and price-cost margins for all the products; Δ is the $J \times J$ Jacobian matrix of demand-price derivatives $\partial D_k / \partial p_j$; and Θ^* is a $J \times J$ matrix with elements $\theta_{f(j),f(k)}$, where $f(j)$ represents the index of the firm that owns product j. If this matrix $\Theta^* \; \Delta\mathbf{D}$ is non-singular, we can obtain price cost margins as:

$$\mathbf{PCM} \;=\; -\left[\Theta^* \; \Delta\mathbf{D}\right]^{-1} \mathbf{q} \tag{4.42}$$

4.3.2 Estimating marginal costs

We first consider the estimation of marginal costs given that the researcher knows the nature of competition, as represented by matrix Θ. For instance, a standard set of assumptions in this context is that there is no collusion (that is, Θ is the identity matrix) and firms compete a la Nash-Bertrand to maximize their own profit.

The researcher has data on J products over T markets, and knows the ownership structure: $\left\{p_{jt}, q_{jt}, \mathbf{x}_{jt} : \; j = 1,...,J; t = 1,2,...,T\right\}$. Suppose that the researcher has estimated in a first step the parameters in the demand system, such that there is a consistent estimator of the Jacobian matrix $\Delta\mathbf{D}$. Therefore, using the marginal conditions of optimality in equation (4.42), we can solve for the vector of marginal costs to obtain:

$$\mathbf{MC}_t \;=\; \mathbf{p}_t + \left[\Theta^* \; \Delta\mathbf{D}_t\right]^{-1} \mathbf{q}_t \tag{4.43}$$

Given the same demand system, different hypotheses about collusion or ownership structures of products (for instance, mergers), imply different estimates of price-cost margins and of marginal costs.

Under the assumption of constant marginal costs – that is, these costs do not depend on the level of output – the realized marginal costs that we recover from equation (4.43) provide the whole marginal cost function. However, in some industries, the assumption of constant marginal costs may not be plausible, and the researcher needs to estimate how these costs depend on the level of output. The estimation of this function is necessary for predictions and counterfactual experiments involving substantial changes in output relative to those observed in the data. In this case, identification of realized marginal costs is not enough and we need to estimate the marginal cost function.

Consider the following cost function,

$$C(q_{jt}) \;=\; \frac{1}{\gamma_q + 1} q_{jt}^{\gamma_q + 1} \; \exp\{\mathbf{x}_{jt}' \gamma_x + \varepsilon_{jt}^{MC}\}, \tag{4.44}$$

with the corresponding marginal cost function,

$$MC_{jt} \;=\; q_{jt}^{\gamma_q} \; \exp\{\mathbf{x}_{jt}' \gamma_x + \varepsilon_{jt}^{MC}\}, \tag{4.45}$$

where γ_q and γ_x are parameters, and ε_{jt}^{MC} is unobservable to the researcher. Taking logarithms, we have the following linear-in-parameters regression model:

$$\ln\left(MC_{jt}\right) \;=\; \gamma_q \; \ln\left(q_{jt}\right) + \mathbf{x}_{jt}' \gamma_x + \varepsilon_{jt}^{MC} \tag{4.46}$$

Note that the realized log- marginal cost, $\left(MC_{jt}\right)$, is known to the researcher as it has been identified using equation (4.43). We are interested in the estimation of the parameters γ_q and γ_x.

The equilibrium model implies that the amount of output q_{jt} is negatively correlated with the unobservable cost inefficiency ε_{jt}^{MC}. Firms/products with larger ε_{jt}^{MC} are less cost-efficient, and this, all else equal, implies a smaller amount of output. Therefore, regressor $\ln(q_{jt})$ is endogenous in the regression equation that represents the log- marginal cost function. Fortunately, the model implies a exclusion restriction that can be used to obtain valid instruments. Note that, given the own characteristics of product j, $\mathbf{x}_{jt}$, and the own amount of output q_{jt}, the marginal cost of this product does not depend on the characteristics of other products in the market, $\{\mathbf{x}_{kt} : k \neq j\}$. However, the model of competition implies that the equilibrium amount of output for a product depends not only on the own characteristics but also on the attributes of competing products. Suppose that $\mathbf{x}_{jt}$ is not correlated with the unobservable cost inefficiency ε_{jt}^{MC}. Then, the model implies that we can use $\{\mathbf{x}_{kt} : k \neq j\}$ as instruments for the endogenous regressor $\ln(q_{jt})$ in the regression equation (4.46).

4.3.3 Testing hypotheses on nature of competition

Researchers can be interested in using price data to learn about the nature of competition, instead of imposing an assumption about matrix Θ. Can we identify collusive behavior? Can we identify matrix Θ? As in the homogeneous product case, we can distinguish two cases for this identification problem: with and without data on marginal costs.

Suppose that the researcher observes the true MC_{jt}. Perhaps more realistically, suppose that the researcher observes some measures of marginal costs that we represent using $q \times 1$ vector $\mathbf{c}^{MC}$. For instance, $\mathbf{c}^{MC}$ may include the mean value of marginal costs for all the products and firms in the industry and for one year during the sample period; or the mean value of realized marginal costs for a particular firm. In the best case scenario, $\mathbf{c}^{MC}$ includes the marginal cost of every product at every sample period.

Given an estimated demand system and a hypothesis about the nature of competition, as represented by a matrix Θ, we can use equation (4.43) to obtain the corresponding vector of marginal costs, and then we can use these values to construct the predicted value of $\mathbf{c}^{MC}$ implied by this value of Θ. We denote this predicted value as $\mathbf{c}^{MC}(\Theta)$. Then, we can compare the actual value $\mathbf{c}^{MC}$ and the predicted value $\mathbf{c}^{MC}(\Theta)$. More formally, we can use the difference between actual and predicted value to construct a test for the null hypothesis that our conjecture about Θ is correct. For instance, if $\mathbf{c}^{MC}$ is a vector of sample means (or more generally, a vector of moments) we can construct a Chi-square goodness-of-fit test. Under the null hypothesis:

$$\left[\mathbf{c}^{MC} - \mathbf{c}^{MC}(\Theta)\right]' \left[Var(\mathbf{c}^{MC}(\Theta))\right]^{-1} \left[\mathbf{c}^{MC} - \mathbf{c}^{MC}(\Theta)\right] \sim \chi_q^2 \tag{4.47}$$

This approach has been used in a good number of papers to test collusion and other hypotheses about the nature of competition. Bresnahan (1987) on the US automobile industry was the pioneering study using this approach. He finds evidence of collusive behavior. Other very influential paper using this method is Nevo (2001) on the US Ready-to-Eat cereal industry. Nevo rejects collusive behavior, and finds that the multi-product feature of firms accounts for a very substantial fraction of market power. Some authors, such as Gasmi, Laffont, and Vuong (1992), have used non-nested testing procedures (e.g., Vuong-Test) to select between alternative hypothesis about the nature of competition. Gasmi, Laffont, and Vuong (1992) study competition in prices and advertising between

Coca-Cola and Pepsi-Cola during 1968-1986. They cannot reject the null hypothesis of collusion between these firms.

4.3.4 Estimating the nature of competition

Suppose that the elements of matrix Θ are real numbers within the interval $[0,1]$. That is, $\theta_{f,g}$ represents the degree to which firm f internalizes the profits of firm g when setting prices of its own products. Can we identify these parameters without data on marginal costs? We show here conditions under which these parameters are identified, and describe an estimation method.

It is helpful to illustrate identification and estimation using a simple version of the model with only two single-product firms, firms 1 and 2. This model has two conjectural parameters, θ_{12} and θ_{21}. The marginal condition of optimality in equation (4.40) has the following form, for firm 1:

$$q_{1t} + (p_{1t} - MC_{1t}) \frac{\partial D_{1t}}{\partial p_{1t}} + \theta_{12} \ (p_{2t} - MC_{2t}) \frac{\partial D_{2t}}{\partial p_{1t}} = 0 \tag{4.48}$$

We can re-write the equation as follows

$$p_{1t} - \left(\frac{\partial D_{1t}}{\partial p_{1t}}\right)^{-1} q_{1t} = MC_{1t} + \theta_{12} \ (p_{2t} - MC_{2t}) \left(\frac{\partial D_{1t}}{\partial p_{1t}}\right)^{-1} \frac{\partial D_{2t}}{\partial p_{1t}} \tag{4.49}$$

The econometric model is completed with a specification of the marginal cost function. For instance:

$$MC_{jt} = \mathbf{x}'_{jt}\gamma + \varepsilon^{MC}_{jt} \tag{4.50}$$

Plugging this marginal cost function into the marginal condition of optimality, we get the following regression equation for product 1:

$$y_{1t} = \theta_{12} \ \widetilde{p}_{2t} + \mathbf{x}'_{1t}\gamma + \widetilde{\mathbf{x}}'_{2t}\pi_1 + u_{1t} \tag{4.51}$$

with: $y_{1t} \equiv p_{1t} - (\partial D_{1t}/\partial p_{1t})^{-1} q_{1t}$; $\widetilde{p}_{2t} \equiv p_{2t}(\partial D_{1t}/\partial p_{1t})^{-1}(\partial D_{2t}/\partial p_{1t})$; $\widetilde{\mathbf{x}}_{2t} \equiv \mathbf{x}_{2t}$ $(\partial D_{1t}/\partial p_{1t})^{-1}(\partial D_{2t}/\partial p_{1t})$; $u_{1t} \equiv \varepsilon^{MC}_{1t} - \theta_{12}\varepsilon^{MC}_{2t}(\partial D_{1t}/\partial p_{1t})^{-1}(\partial D_{2t}/\partial p_{1t})$; and $\pi_1 \equiv -\theta_{12}\gamma$. We have a similar regression equation for product 2.

The estimation of the parameters in this regression equation needs to deal with the endogeneity of prices. Regressors $\widetilde{p}_{2t}$ and $\widetilde{\mathbf{x}}_{2t}$ are endogenous because they depend on prices, and prices are correlated with the unobserved cost inefficiencies ε^{MC}_{jt} which enter into the error term u_{1t}. In this context, using the so-called *BLP instruments* (that is, the observable characteristics $\mathbf{x}$ of other products is tricky because these variables already enter in the regressor $\widetilde{\mathbf{x}}_{2t}$. We need additional instruments for the endogenous regressor $\widetilde{p}_{2t}$. Possible identification strategies are: *Hausman-Nevo instruments*, when the dataset includes multiple geographic markets and demand unobservables are not correlated across markets after controlling for product fixed effects; or *Arellano-Bond instruments*, when the dataset includes multiple time periods and demand unobservables are not serially correlated after controlling for product fixed effects. We discuss below an empirical application that uses a different identification strategy.[6]

[6] In principle, we could use BLP instruments if we impose the restrictions between the parameters θ_{12}, γ, and π_1: that is, for any product attribute, say k, in the vector $\mathbf{x}$, we have that $\pi_{1k}/\gamma_k = \theta 12$.

Example: Collusion in the Ready-to-Eat (RTE) cereal industry.

Michel and Weiergraeber (2018) study competition in the US RTE cereal industry during the period 1991-1996. There were two important events in this industry during this period: the merger of two leading firms, *Post* and *Nabisco* in 1993; and a massive wholesale price reduction in 1996. The paper emphasizes the importance of allowing *conduct parameters* θ to vary over time and across firms when an industry is subject to important shocks. This view is consistent with interpretation of conduct parameters as endogenous objects in a broader dynamic game of the industry, as we have discussed above in this chapter. The authors are also concerned with finding powerful instruments to separately identify conduct and marginal costs parameters. They propose novel instruments that exploit information on firms' promotional activities.

The main data consists of consumer scanner data from the Dominick's Finer Food (DFF) between February 1991 and October 1996. It includes 58 supermarket stores located in the Chicago metropolitan area. The authors aggregate the data at the monthly level (69 months) and focus on 26 brands of cereal from the 6 nationwide manufacturers: Kellogg's, General Mills, Post, Nabisco, Quaker Oats, and Ralston Purina. Brands are classified into three groups: adult, family, and kids. Importantly for the purpose of this paper, the dataset contains information on wholesale prices and in-store promotional activities.

It is well-known that this is a highly concentrated industry. During this period and market, the leader (Kellogg's) had a market share of 45%, and the top-2 firms accounted for 75%. Firms market shares were relatively stable over the sample period, though there some changes after the 1993 merger between Post and Nabisco.

In the antitrust authority's evaluation of the proposed merger between Post and Nabisco, the main concern was the strong substitutability in the adult cereal segment between Post's and Nabisco's products. The merger did not lead to any product entry or exit or any changes in existing products. Following the merger, Post+Nabisco increased significantly its prices, and this price increase was followed by the rest of the firms. In principle, this response could be explained under Nash-Bertrand competition (before and after the merger), without the need of any change in conduct parameters.

On April 1996, Post decreased its wholesale prices by 20%. This was followed, a few weeks later, by significant price cuts by the other firms. The average decrease in the wholesale price between April and October 1996 was 9.66% (and 7.5% in retail price). The main purpose of this paper is explaining the role that different factors played in this price reductions – including potential changes in firms' conduct.

The authors estimate a random coefficients nested logit model for the demand system. This demand system is similar to the one in Nevo (2001), but it has an important distinguishing feature: it includes as a product characteristic the variable PRO_{jt} that represents the total (aggregated over stores and type of promotion) in-store promotions of product j during month t. The estimate this demand system using BLP-instruments (characteristics of other products) as instrumental variables. In particular, the authors exploit the substantial time variation in the promotion variables.

Given the estimated demand system, the authors then estimated the conduct parameters θ in a regression model very similar to the one in equation (4.51) but for six multi-product firms, instead of two single-product firms. That is, the authors estimate the whole matrix Θ of conduct parameters, together with marginal cost parameters, and allow this matrix to vary across three different subperiods: before Post+Nabisco merger in 1993; between 1993 and April 1996; after April 1996. To deal with the endogeneity of variables $\widetilde{p}$ and $\widetilde{\mathbf{x}}$ in the regression equation (4.51), the authors use promotional variables of other products as instruments. Demand elasticities are significantly affected by these variables. They have substantial variation across products, over time, and markets. Still, there is the concern that promotional variables are endogenous: they can be correlated with the unobservable component of the marginal cost. Promotions are chosen by firms: it is more profitable to make promotions when marginal costs are low. To deal with this endogeneity, the authors assume that the error term follows and AR(1) ($u_{jt} = \rho$ $u_{j,t-1} + v_{jt}$, where v_{jt} is i.i.d.), and they argue that promotions are negotiated between manufacturers and retailers at least one month in advance. Then, they take a quasi-first-difference of the regression equation (that is, a Cochrane-Orcutt transformation, $y_{jt} - \rho$ $y_{j,t-1}$). In this transformed equation, $PROMO_{kt}$ is not correlated with the i.i.d. shock v_{jt} because promotions are determined at least one month in advance.

The estimation results show strong evidence for coordination between 1991-1992. On average the conduct parameter is 0.277: that is, a firm values \$1 of its rivals' profits as much as \$0.277 of its own profits. Because of this coordination, pre-merger price-cost margins are 25.6% higher than under multi-product Bertrand-Nash pricing. After the Post + Nabisco merger in 1993, the degree of coordination increased significantly, on average to 0.454. Towards year 1996, the degree of coordination becomes close to 0, consistent with multi-product Bertrand-Nash pricing. Counterfactual experiments show that if firms had competed à la Bertrand-Nash before 1996, consumer welfare would have increased by between \$1.6 – \$2.0 million per year, and the median wholesale price would have been 9.5% and 16.3% lower during the pre-merger and post-merger periods, respectively.

4.3.5 Conjectural variations with differentiated product

So far, in the model with differentiated product, we have incorporated the *nature of competition* by including parameters $\theta_{f,g}$ that represent to what extent firm f values the profit of firm g relative to its own profit. This is a reasonable way of modelling collusion. However, it seems quite different to the *conjectural variation* model that we studied for the homogeneous product model. In this section, we present the conjectural variation model in differentiated product industry with price competition. We show that the marginal conditions of optimality from this model have a similar form as those from the model with profit-weights $\theta_{f,g}$.

For simplicity, consider a differentiated product industry with two single-product firms: firm 1 and firm 2. The profit function of firm j is $\Pi_j = p_j\, q_j - C_j(q_j)$. Define the conjecture parameter CV_1 as firm 1's belief about how firm 2 will change its price when firm 1 changes marginally its own price. That is, CV_1 represents firm 1's belief about $\partial p_2 / \partial p_1$. Similarly, CV_2 represents firm 2 's belief about $\partial p_2 / \partial p_1$. Nash-Bertand competition implies $CV_j = 0$ for every firm j. Perfect collusion, implies $CV_j = 1$ for every firm j. Taking the conjecture CV as given, the marginal condition for profit maximization for firm 1 is:

$$q_1 + (p_1 - MC_1)\,\frac{\partial D_1}{\partial p_1} + CV_1\,(p_1 - MC_1)\,\frac{\partial D_1}{\partial p_2} \;=\; 0 \tag{4.52}$$

There are both similarities and differences between this equation and the marginal condition with profit-weights in equation (4.48). The two equations are equivalent when the firms have the same marginal costs and there is symmetric product differentiation.[7] However, there are quantitative differences between the predictions of the two models when firms are heterogeneous in marginal costs or product quality.

Example: Logit demand model with conjectural variations.
Suppose that the demand system has a logit structure where the average utility of product j is $\beta_j - \alpha\, p_j$, where β_j represents the qaulity of product j. This model implies the following equation for the marginal condition of product 1:

$$p_1 - MC_1 \;=\; \frac{1}{\alpha\,(1 - s_1 - s_2\,CV_1)} \tag{4.53}$$

where s_j is the market share of product j.

Suppose that the researcher does not know the magnitude of the marginal costs MC_1 and MC_2, but she knows that the two firms use the same production technology, the same type of variable inputs, and purchase these inputs in the same markets where they are price takers. Therefore, the researcher knows that $MC_1 = MC_2 = MC$, though she does not know the magnitude of MC. This information, together with the marginal conditions of optimality, imply the following equation for the difference between prices:

$$p_1 - p_2 \;=\; \frac{1}{\alpha\,(1 - s_1 - s_2\,CV_1)} - \frac{1}{\alpha\,(1 - s_2 - s_1\,CV_2)} \tag{4.54}$$

The researcher observes prices $p_1 = \$200$ and $p_2 = \$195$ and market shares $s_1 = 0.5$ and $s_2 = 0.2$. Firm 1 has both a larger price and a larger market share because its product has better quality.[8] The researcher has estimated the demand system and knows that $\alpha = 0.01$. Solving these data into the previous equation, we have:

$$\$200 - \$195 = \frac{100}{1 - 0.5 - 0.2\,CV_1} - \frac{100}{1 - 0.2 - 0.5\,CV_2} \tag{4.55}$$

This is a condition that the parameters CV_1 and CV_2 should satisfy. Using this equation we can show that the hypothesis of Nash-Bertrand competition (that requires $CV_1 = CV_2 = 0$) implies a prediction about the price difference $p_1 - p_2$ that is substantially larger than the price difference that we observe in the data. The hypothesis of Nash-Bertrand competition, $CV_1 = CV_2 = 0$, implies that the right hand side of equation (4.55) is:

$$\frac{100}{0.5} - \frac{100}{0.8} = 200 - 125 = \$75 \tag{4.56}$$

That is, Nash-Bertrand implies a price difference of \$75 but the price difference in the data is only \$5. The hypothesis of Collusion, $CV_1 = CV_2 = 1$, implies that the right hand side of the equation in (4.55) is:

$$\frac{100}{0.5 - 0.2} - \frac{100}{0.8 - 0.5} = \$0 \tag{4.57}$$

[7] We have symmetric product differentiation if $\partial D_j / \partial p_j$ is the same for every product j, and $\partial D_j / \partial p_k = \partial D_k / \partial p_j$ for every pair of products j,k. For instance, this is the case in a logit demand model where all the products have the same quality, or in Hotelling (1929) linear-city and Salop (1979) circle-city models when firms are equidistant from each other.

[8] In this industry, higher product quality requires a larger fixed cost but it does not affect marginal cost.

That is, Collusion implies a price difference of \$0, which is closer to the price difference of \$5 that we observe in the data. Under the restriction $CV_1 = CV_2$, we can use equation (4.55) to obtain the value of the conjecture parameter. It implies a quadratic equation in CV, and the positive root is $CV = 0.984$.

4.4 Incomplete information

In this chapter, we have considered different factors that can affect price and quantity competition and market power in an industry. Heterogeneity in marginal costs, product differentiation, multi-product firms, or conduct/nature of competition are among the most important features that we have considered so far. All the models that we have considered assume that firms have perfect knowledge about demand, their own costs, and the costs of their competitors. In game theory, this type of model is a *game of complete information*. This assumption can be quite unrealistic in some industries. Firms have uncertainty about current and future realizations of demand, costs, market regulations, or the behavior of competitors. This uncertainty can have substantial implications for their decisions and profits, and for the efficiency of the market. For example, some firms may be more efficient in gathering and processing information, and they can use this information in their pricing or production strategies to improve their profits.

The assumption of firms' complete information has been the status quo in empirical models of Cournot or Bertrand competition. In reality, firms often face significant uncertainty about demand and about their rivals costs and strategies. Firms are different in their ability and their costs for collecting and processing information, for similar reasons as they are heterogeneous in their costs of production or investment. In this section, we study models of price and quantity competition that allow for firms' incomplete and asymmetric information. Our main purpose is to study how limited information affects competition and market outcomes.

4.4.1 Cournot competition with private information

Vives (2002) studies theoretically the importance of firms' private information as a determinant of prices, market power, and consumer welfare. He considers a market in which firms compete à la Cournot and have private information. Then, he studies the relative contribution of private information and market power in accounting for the welfare losses. He shows that in large enough markets, abstracting from market power provides a much better approximation than abstracting from private information. If M represents market size, then the effect of market power is of the order of $1/M$ for prices and $1/M^2$ for per-capita deadweight loss, while the effect of private information is of the order of $1/\sqrt{M}$ for prices and $1/M$ for per-capita deadweight loss. Numerical simulations of the model show that there is a critical value for market size M^* (that depends on the values of structural parameters) such that the effect of private information dominates the effect of market power when market size is greater than this threshold value.

(i) Demand, costs, and information structure

Consider the market for a homogeneous product where firms compete à la Cournot and there is free market entry. A firm's marginal cost is subject to idiosyncratic shocks that are private information to the firm. The demand function and the marginal cost functions are linear such that the model is linear-quadratic. This feature facilitates substantially the characterization of a Bayesian Nash equilibrium in this model with incomplete information.

There are M consumers in the market and each consumer has an indirect utility function $U(x) = \alpha\ x - \beta\ x^2/2 - p\ x$, where x is the consumption of the good, p is the market price, and $\alpha > 0$ and $\beta > 0$ are parameters. This utility function implies the market level inverse demand function, $p = P(Q) = \alpha - \beta_M\ Q$, where $\beta_M \equiv \beta/M$. Firms are indexed by i. If firm i is actively producing in the market, its cost function is $C(q_i, \theta_i) = \theta_i\ q_i + (\gamma/2)\ q_i^2$ such that its marginal cost is $MC_i = \theta_i + \gamma\ q_i$. Variable θ_i is private information to firm i. In games of incomplete information, θ_i is denoted as player i's (in this case, firm i's) *type*. Firms' types are random variables which are i.i.d. with mean μ_θ and variance σ_θ^2. This distribution is *common knowledge* to all firms.[9] Every active firm producing in the market should pay a fixed cost $F > 0$.

(ii) Bayesian Nash equilibrium

The model is a two-stage game. In the first stage, firms decide whether to enter the market or not. If a firm decides to enter, it pays a fixed cost $F > 0$. When a firm makes its entry decision it does not know yet the realization of its type θ_i. Therefore, the entry decision is based on the maximization of expected profits. At the second stage, each active firm i that has decided to enter observes its own θ_i but not the θ's of the other active firms, and competes according to a Bayesian Nash-Cournot equilibrium. This equilibrium concept is a version of Nash equilibrium for games of incomplete information, and we describe it below.

We now recursively solve the equilibrium of the model starting at the second stage. For the moment, suppose that there are n firms active in the market: we later obtain the equilibrium value of n. The expected profit of firm i is:

$$\begin{aligned}\pi_i(\theta_i) &= \mathbb{E}\left[P(Q) \mid \theta_i\right]\ q_i - \theta_i\ q_i - \frac{\gamma}{2} q_i^2 \\ &= \left(\alpha - \beta_M \left(q_i + \mathbb{E}\left[\sum_{j \neq i} q_j\right]\right)\right)\ q_i - \theta_i\ q_i - \frac{\gamma}{2} q_i^2,\end{aligned} \tag{4.58}$$

where the expectation $\mathbb{E}\,[.]$ is over the distribution of the variables θ_j for $j \neq i$, which are not known to firm i. A *Bayesian Nash Equilibrium (BNE)* is an n-tuple of strategy functions, $[\sigma_1(\theta_1), \sigma_2(\theta_n), ..., \sigma_n(\theta_n)]$, such that for every firm's strategy maximizes its own expected profit taking as given other firms' strategies. That is, for every firm i:

$$\sigma_i(\theta_i) = \arg\max_{q_i}\ \mathbb{E}\left[P(Q) \mid \theta_i, \sigma_j \text{ for } j \neq i\right]\ q_i - \theta_i\ q_i - \frac{\gamma}{2} q_i^2 \tag{4.59}$$

The first order condition of optimality for the best response of firm i implies:

$$q_i = \sigma_i(\theta_i) = [\gamma + 2\beta_M\,]^{-1} \left[\alpha - \theta_i - \beta_M \sum_{j \neq i} \mathbb{E}\left(\sigma_j(\theta_j)\right)\right] \tag{4.60}$$

[9] In game theory, an object or event is *common knowledge* if everybody knows that everybody knows that ... knows it.

Since firms are identical up to the private information θ_i, it seems reasonable to focus on a symmetric BNE such that $\sigma_i(\theta_i) = \sigma(\theta_i)$ for every firm i. Imposing this restriction in the best response condition (4.60), taking expectations over the distribution of θ_i, and solving for $\sigma^e \equiv \mathbb{E}(\sigma(\theta_i))$, we obtain that:

$$\sigma^e \equiv \mathbb{E}(\sigma(\theta_i)) = \frac{\alpha - \mu_\theta}{\gamma + \beta_M\ (n+1)} \tag{4.61}$$

Solving this expression in (4.60), we obtain the following closed-form expression for the equilibrium strategy function under BNE, in the second stage of the game:

$$q_i = \sigma(\theta_i) = \frac{\alpha - \mu_\theta}{\gamma + \beta_M\ (n+1)} - \frac{\theta_i - \mu_\theta}{\gamma + 2\beta_M} \tag{4.62}$$

Now, we proceed to the first stage of the game to obtain the equilibrium number of active firms in the market. Under the BNE in the second stage, the expected profit of an active firm, before knowing the realization of its own θ_i is:

$$\mathbb{E}\left[\pi(\theta_i)\right] = \left[\beta_M + \gamma/2\right]\ \mathbb{E}\left[\sigma(\theta_i)^2\right] = \frac{\left[\alpha - \mu_\theta^2\right]}{\left[\gamma + \beta_M\ (n+1)\right]^2} + \frac{\sigma_\theta^2}{\left[\gamma + 2\beta_M\right]^2} \tag{4.63}$$

Given this expected profit, we can obtain the the equilibrium number of entrants in the first stage of the game. Given a market of size M, the free-entry number of firms $n^*(M)$ is approximated by the solution to $\mathbb{E}\left[\pi(\theta_i)\right] - F = 0$. Given the expression for the equilibrium profit, it is simple to verify that $n^*(M)$ is of the same order as market size M. That is, the ratio $n^*(M)/M$ of the firms per consumer is bounded away from zero and infinity.

(iii) Welfare analysis

From the point of view of a social planner, the optimal allocation in this industry can be achieved if firms share all their information and behave as price takers. Let us label this equilibrium as $CI - PT$: *complete information* with *price taking* behavior. If p and W are the price and the total welfare, respectively, under the "true" model (with both Cournot conduct and private information), then the differences $p - p_{CI-PT}$ and $W - W_{CI-PT}$ represent the combined effect of incomplete information and Cournot behavior on prices and on welfare.

To measure the separate effects of incomplete information and Cournot behavior, it is convenient to define other two models: a model of Cournot competition with complete information, that we label as CI; and a model of incomplete information that assumes that firms are price takers, that we label as PT.[10] Using these models, we can make the following decomposition:

$$\begin{aligned} p - p_{CI-PT} &= \left[p - p_{PT}\right] + \left[p_{PT} - p_{CI-PT}\right] \\ W_{CI-PT} - W &= \left[W_{CI-PT} - W_{PT}\right] + \left[W_{PT} - W\right] \end{aligned} \tag{4.64}$$

[10] In the complete information Cournot model, equilibrium output is: $q_i^{CI} = \dfrac{\alpha - \widetilde{\theta}_n}{\gamma + \beta_M\ (n+1)} - \dfrac{\theta_i - \widetilde{\theta}_n}{\gamma + 2\beta_M}$, where $\widetilde{\theta}_n \equiv (n-1)^{-1} \sum_{j \neq i} \theta_j$.

The term $p - p_{PT}$ captures the effect of Cournot behavior (market power) on prices, and the term $p_{PT} - p_{CI-PT}$ captures the effect of incomplete information. Similarly, $W_{CI-PT} - W$ is the total deadweight loss, $[W_{CI-PT} - W_{PT}]$ is the contribution of incomplete information, and $[W_{PT} - W]$ is the contribution of Cournot competition.[11]

Vives (2002) shows that as market size M (and therefore n) goes to infinity, market price and welfare per capita converge to the optimal allocation: that is, $[p - p_{CI-PT}] \longrightarrow 0$ and $[W_{CI-PT} - W]/M \longrightarrow 0$. Private information and Cournot behavior have an effect only when the market is not too large. Vives shows also that there is a critical value for market size, M^* (that depends on the values of structural parameters), such that if market size is greater this threshold value, then the effect of private information on prices and consumer welfare dominates the effect of market power.

This result has interesting policy implications. Antitrust authorities look with suspicion at the information exchanges between firms because they can help collusive agreements. The collusion concern is most important in the presence of a few players because collusion is easier to be sustained in this case (repeated game). Vives (2002)'s results show that with few firms, market power (Cournot) has the most important contribution to the welfare loss, so it seems reasonable to control these information exchanges. When market size and the number of firms increase, information asymmetry becomes a more important factor in welfare loss and it is optimal to allow for some information sharing between firms.

(iv) An empirical application

Armantier and Richard (2003) study empirically how asymmetric information on marginal costs affects competition and outcomes in the US airline industry. They investigate how marketing alliances between American Airlines and United Airlines facilitate information sharing and how this affects market outcomes. The authors find that such information exchanges would benefit airlines with a very moderate cost in terms of consumer welfare.

[11]Note that this is one of different ways we can decompose these effects. For instance, we could also consider the decomposition, $p - p_{CI-PT} = [p - p_{CI}] + [p_{CI} - p_{CI-PT}]$ and $W_{CI-PT} - W = [W_{CI-PT} - W_{CI}] + [W_{CI} - W]$. The main results are the same regardless of the decomposition chosen.

4.5 Exercises

4.5.1 Exercise 1

Consider an industry with a differentiated product. There are two firms in this industry, firms 1 and 2. Each firm produces and sells only one brand of the differentiated product: brand 1 is produced by firm 1, and brand 2 by firm 2. The demand system is a logit demand model, where consumers choose between three different alternatives: $j = 0$, represents the consumer decision of no purchasing any product; and $j = 1$ and $j = 2$ represent the consumer purchase of product 1 and 2, respectively. The utility of no purchase ($j = 0$) is zero. The utility of purchasing product $j \in \{1,2\}$ is $\beta\, x_j - \alpha\, p_j + \varepsilon_j$, where the variables and parameters have the interpretation that we have seen in class. Variable x_j is a measure of the quality of product j, for instance, the number of stars of the product according to consumer ratings. Therefore, we have that $\beta > 0$. The random variables ε_1 and ε_2 are independently and identically distributed over consumers with a type I extreme value distribution, that is, Logit model of demand. Let H be the number of consumers in the market. Let s_0, s_1, and s_2 be the market shares of the three choice alternatives, such that s_j represents the proportion of consumers choosing alternative j and $s_0 + s_1 + s_2 = 1$.

Question 1.1. Based on this model, write the equation for the market share s_1 as a function of the prices and the qualities x's of all the products.

Question 1.2. Obtain the expression for the derivatives: (a) $\dfrac{\partial s_1}{\partial p_1}$; (b) $\dfrac{\partial s_1}{\partial p_2}$; (c) $\dfrac{\partial s_1}{\partial x_1}$; and (d) $\dfrac{\partial s_1}{\partial x_2}$. Write the expression for these derivatives in terms only of the market shares s_1 and s_2 and the parameters of the model.

The profit function of firm $j \in \{0,1\}$ is $\pi_j = p_j\, q_j - c_j\, q_j - FC(x_j)$, where: q_j is the quantity sold by firm j (that is, $q_j = H\, s_j$); c_j is firm $j's$ marginal cost, that is assumed constant, that is, linear cost function; and $FC(x_j)$ is a fixed cost that depends on the level of quality of the firm.

Question 1.3. Suppose that firms take their qualities x_1 and x_2 as given and compete in prices ala Bertrand.
(a) Obtain the equation that describes the marginal condition of profit maximization of firm 1 in this Bertrand game. Write this equation taking into account the specific form of $\dfrac{\partial s_1}{\partial p_1}$ in the Logit model.
(b) Given this equation, write the expression for the equilibrium price-cost margin $p_1 - c_1$ as a function of s_1 and the demand parameter α.

Now, suppose that the researcher is not willing to impose the assumption of Bertrand competition and considers a conjectural variations model. Define the conjecture parameter CV_1 as the belief or conjecture that firm 1 has about how firm 2 will change its price when firm 1 changes marginally its price. That is, CV_1 represents the belief or conjecture of firm 1 about $\dfrac{\partial p_2}{\partial p_1}$. Similarly, CV_2 represents the belief or conjecture of firm 2 about $\dfrac{\partial p_2}{\partial p_1}$.

Question 1.4. Suppose that firm 1 has a conjectural variation CV_1.
(a) Obtain the equation that describes the marginal condition of profit maximization of firm 1 under this conjectural variation. Write this equation taking into account the specific form of $\frac{\partial s_1}{\partial p_1}$ in the Logit model. [Hint: Now, we have that: $\frac{dq_1}{dp_1} = \frac{\partial q_1}{\partial p_1} + \frac{\partial q_1}{\partial p_2}\frac{\partial p_2}{\partial p_1}$, where $\frac{\partial q_1}{\partial p_1}$ and $\frac{\partial q_1}{\partial p_2}$ are the expressions you have derived in Q1.2].
(b) Given this equation, write the expression for the equilibrium price-cost margin $p_1 - c_1$ as a function of the market shares s_1 and s_2, and the parameters α and CV_1.

Question 1.5. Suppose that the researcher does not know the magnitude of the marginal costs c_1 and c_2, but she knows that the two firms use the same production technology, they use the same type of variable inputs, and they purchase these inputs in the same markets where they are price takers. Under these conditions, the researcher knows that $c_1 = c_2 = c$, though she does not know the magnitude of the marginal cost c.
(a) The marginal conditions for profit maximization in Q1.4(b), for the two firms, together with the condition $c_1 = c_2 = c$, imply that price difference between these two firms, $p_1 - p_2$, is a particular function of their markets shares and their conjectural variations. Derive the equation that represents this condition.
(b) The researcher observes prices $p_1 = \$200$ and $p_2 = \$195$ and market shares $s_1 = 0.5$ and $s_2 = 0.2$. Firm 1 has both a larger price and a larger market share because its product has better quality, that is, $x_1 > x_2$. The researcher has estimated the demand system and knows that $\alpha = 0.01$. Plug in these data into the equation in Q1.5(a) to obtain a condition that the parameters CV_1 and CV_2 should satisfy in this market.
(c) Using the equation in Q1.5(b), show that the hypothesis of Nash-Bertrand competition (that requires $CV_1 = CV_2 = 0$) implies a prediction about the price difference $p_1 - p_2$ that is substantially larger than the price difference that we observe in the data.
(d) Using the equation in Q1.5(b), show that the hypothesis of Collusion (that requires $CV_1 = CV_2 = 1$) implies a prediction about the price difference $p_1 - p_2$ that is much closer to the price difference that we observe in the data.

4.5.2 Exercise 2

To answer the questions in this part of the problem set you need to use the dataset `verboven_cars.dta` Use this dataset to implement the estimations describe below. Please, provide the STATA code that you use to obtain the results. For all the models that you estimate below, impose the following conditions:

- For market size (number of consumers), use Population/4, that is, `pop/4`
- Use prices measured in euros (`eurpr`).
- For the product characteristics in the demand system, include the characteristics: `hp, li, wi, cy, le,` and `he.`
- Include also as explanatory variables the market characteristics: ln(pop) and log(gdp).
- In all the OLS estimations include fixed effects for market (`ma`), year (`ye`), and brand (`brd`).
- Include the price in logarithms, that is, ln(`eurpr`).

- Allow the coefficient for log-price to be different for different markets (countries). That is, include as explanatory variables the log price, but also the log price interacting (multiplying) each of the market (country) dummies except one country dummy (say the dummy for Germany) that you use as a benchmark.

Question 2.1.
(a) Obtain the OLS-Fixed effects estimator of the Standard logit model. Interpret the results.
(b) Test the null hypothesis that all countries have the same price coefficient.
(c) Based on the estimated model, obtain the average price elasticity of demand for each country evaluated at the mean values of prices and market shares for that country.

Question 2.2. Consider the equilibrium condition (first order conditions of profit maximization) under the assumption that each product is produced by only one firm.
(a) Write the equation for this equilibrium condition. Write this equilibrium condition as an equation for the Lerner Index, $\frac{p_j - MC_j}{p_j}$.
(b) Using the previous equation in Q2.2(a) and the estimated demand in Q2.1, calculate the Lerner index for every car-market-year observation in the data.
(c) Report the mean values of the Lerner Index for each of the counties/markets. Comment the results.
(d) Report the mean values of the Lerner Index for each of the top five car manufacturers (that is, the five car manufacturers with largest total aggregate sales over these markets and sample period). Comment the results.

Question 2.3.
(a) Using the equilibrium condition and the estimated demand, obtain an estimate of the marginal cost for every car-market-year observation in the data.
(b) Run an OLS-Fixed effects regression where the dependent variable is the estimated value of the marginal cost, and the explanatory variables (regressors) are the product characteristics `hp`, `li`, `wi`, `cy`, `le`, and `he`. Interpret the results.