ECONOMIC DISCUSSION PAPERS

Efficiency Series Paper 02/2007

Estimating Firm-Specific Market Power: A Composed Error Term Approach

Luis Orea

Departamento de Economía

Universidad de Oviedo

Available online at: www.uniovi.es/economia/edp.htm

UNIVERSIDAD DE OVIEDO

DEPARTAMENTO DE ECONOMÍA

PERMANENT SEMINAR ON EFFICIENCY AND PRODUCTIVITY

ESTIMATING FIRM-SPECIFIC MARKET POWER: A COMPOSED ERROR TERM APPROACH[*]

Luis Orea[♣]

Efficiency Series Paper 02/2007

Abstract: Measuring the degree of competition in oligopolistic markets is a key activity in empirical industrial organization. Earlier studies focused on estimating conduct parameters imposed some *structure* on the way the conduct parameter varies across firms and time. We bring market power and efficiency literatures together by treating firms' behaviour as an asymmetric random parameter, which allows estimating firm-specific and time-varying market power even when no panel data is available. We show that different oligopolistic equilibrium outcomes are associated to different types of skewness of the conduct random term. Hence, analyzing the skewness of the conduct random term not only allows us to identify collusive and competitive (maverick) firms, but also it provides useful information about the nature of the competition among firms in a particular market.

Key words: market power, random conduct parameter, composed error model, asymmetric distributions.

---

[*] The author thanks Andrés Gómez Lobo, Subal Kumbhakar, Albert Satorra and Julio del Corral for their comments and suggestions on various portions of early draft. The usual disclaimer applies.

[♣] Departamento de Economía, Universidad de Oviedo, Spain.

e-mail: lorea@uniovi.es

1. Introduction

Measuring the degree of competition in oligopolistic markets and finding the underlying determinants of such competition are key activities in empirical industrial organization. Earlier studies focused on estimating conduct parameters that distinguish collusive behaviours from non-collusive behaviours, using contemporary observations of outputs, costs, and prices. The literature on measuring oligopolistic conduct follows from original research by Iwata (1974), Gallop and Roberts (1979), and Appelbaum (1982), who estimate a static model of firm's behaviour. The traditional static market power model is designed to estimate the level of market competition in a one-shot game that is repeated over time. Since in these models the (firm or industry) degree of market power is measured by a parameter θ that is jointly estimated with other cost and demand parameters, it is called the Conduct Parameter Method.[1]

In static models, firms maximize individual profits each period without explicit consideration of the effect of behaviour in one period on the competitive environment in other periods. As the problem of repeated oligopoly interaction has received greater attention, the estimation of time-varying conduct parameters that are truly dynamic has become an issue. The Stigler's (1964) theory of collusive oligopoly implies that, in an uncertain environment, both collusive and price-war periods will be seen in the data. If data on cartelized or tacitly collusive industry should show both periods of successful cooperation and periods of outright competition, empirically, this will show up as a time-varying conduct parameter.[2] In particular, Green and Porter (1984) predict a procyclical behaviour pattern for mark-ups because of price reversion during a period of low demand. Hence the conduct parameter changes from collusive value to competitive value when there is an unanticipated negative demand shock. Meanwhile, Rotemberg and Saloner (1986) predict that prices and mark-ups are countercyclical. The incentive to deviate from collusive agreements is greater when demand is high, so the optimal price decreases during a boom to prevent a deviation from the collusion in this model. Hence the conduct parameter will decrease when demand is high.

---

[1] Another approach in the New Empirical Industrial Organization (NEIO) is to estimate the demand and pricing relationship under specific assumptions of market competition. This approach has been used for differentiated product markets with price competition. See, for example, Berry, Levinston and Pakes (1995), Nevo (2001) and Jaumandreu and Lorences (2002).

[2] Empirical studies that estimate time-varying conduct include Bresnahan (1987), Brander and Zhang (1993), Gallet and Schroeter (1995), and Gallet (1997).

The above theories have in common that they imply an empirical model where the conduct parameter changes over time. Since they predict different time patterns the time-series behaviour of conduct is hard to estimate. Moreover, Abreu et al. (1986) find that in complex cartel designs the length of price wars is (i.e. changes in conduct parameter are) random due to there are "triggers" both for beginning a price war and for ending one.

The conduct parameter not only varies over time, but also across firms. In a many treatments of oligopoly as a repeated game, firms expect deviations from the collusive outcome. Firms expect that if they deviate from the collusive arrangement, other will too. This expectation deters them from departing from their share of the collusive output. Since these deviations are unobserved in an uncertain environment, each firm might have their own expectation about what would happen if he deviates from collusive output. That is, "there is nothing in the logic of oligopoly theory to force all firms to have the same conduct" (Bresnahan, 1989, p. 1030).

Taking all of these theories at once would lead to complex firm and time-varying conduct parameters. Obviously, allowing the conduct parameter varies both by firm and observation results in an overparameterized model. To avoid this problem the empirical models proposed in the literature always put some *structure* on the way the conduct parameter varies across firms and time. The overparameterization is solved by aggregation (i.e. by estimating the average of the conduct parameters of the firms in the industry), reducing the time variation into a period of successful cartel cooperation and a period of price wars or similar breakdowns in cooperation, allowing different conduct parameters between two o more groups of firms.

Most of the structural econometrics models treat the firm's behaviour as a common parameter to be estimated altogether with other cost and demand parameters. A crucial issue using this approach is how the data identify market power (i.e. θ_{it} or an industry-aggregation of this parameter) from other cost and demand parameters. Bresnahan (1982) solved this by introducing variables that combine elements both of rotation and of vertical shifts in the demand curve.[3] Other methods to achieve this

[3] The economic intuition behind this result is quite straightforward. Suppose that the exogenous variables entering demand rotate the demand curve around the industry equilibrium point.

objective are based on comparative statics in costs, supply shocks and econometric estimation of marginal costs (see, for instance, Appelbaum, 1982, and Bresnahan, 1989).

Instead of treating firms' behaviour as a (restricted) parameter to be estimated, in the present paper we propose treating firms' behaviour (i.e. θ_{it}) as a random variable. Therefore, our approach relies on the estimation of a "composed error" model where the stochastic part is formed by two random variables, i.e. the traditional noise term, capturing random shocks, and a random conduct term, which measures market power. Once all parameters describing the structure of the pricing equation are estimated using appropriate econometric techniques, distributional assumptions can be invoked on either error component in order to obtain consistent estimates of the parameters describing the structure of the two error components. Conditional on these parameter estimates, market power is then estimated for each firm by decomposing the estimated residual into a noise component and a market-power component.

Our approach does not require deciding, in advance, which firms are behaving strategically and which are not. Instead, conditional on the distributional assumption on the one-sided error component, our approach allows us to detect firms which are likely involved in a partial (perfect) collusion or behaving (more) competitively (e.g. maverick firms). Note that other papers (see, e.g., Puller, 2007) allowed each firm to have different conduct parameters, but these point estimates are only asymptotically consistent when the time dimension of their panel data set is long (i.e. $T \rightarrow \infty$). Since, in our model, the firm-specific market power estimates relies on distributional assumptions on the two error components, they can be obtained, unlike in previous papers, when a simple cross-section data set is available.

It should be noted that the main contribution of the proposed approach is not the way that both parameters describing the structure of the pricing equation and an average level of market power are estimated, but the way the asymmetry of the composed error term (i.e. its skewness or kurtosis) is exploited in order to get firm-specific market

---

Under perfect competition, this will have no effect: supply and demand intersect at the same point before and after the rotation. Under either oligopoly or monopoly, changes in the slope (and thus the elasticity of demand) will shift the perceived marginal revenue of firms. Equilibrium price and quantity will respond. Hence, the market comparative statics of perfect competition are distinct from those of monopoly.

power estimates. Indeed, the key feature of the composed error term is that it is asymmetric in two senses. First, the economic theory suggests that θ_{it} always takes positive values, so it follows a one-sided distribution. Second, several well-known oligopolistic equilibrium outcomes yield skewed conduct random terms. This makes asymmetric the composed error term.[4]

If we take advantage of this asymmetry we can measure not only the average-industry market power, but also market power of each firm. Another advantage of the proposed approach is that the type of skewness or kurtosis of the conduct random term can be used to test different types of oligopolistic equilibrium outcomes. As we aware, skewness or kurtosis of conduct parameter (i.e. prices) in oligopolistic industry settings are not examined explicitly in most (if any) of the previous empirical papers.

We illustrate this approach by estimating a MLE specification of the proposed model using a balanced panel data set from the Spanish banking industry for the period 1993-1998. Our preliminary results show that imposing common conduct behaviour to all banks would yield biased results. In addition, we find a deterioration of competition during the period 1994-1998 since expected demand increase and expected decreases in costs.

2. Theoretical background and empirical specification

Homogeneous product

As it is customary, we first assume a static model, i.e. firms maximize individual profits each period without explicit consideration of the effect of behaviour in one period on the competitive environment in other periods.[5] Assume, for instance, that N symmetric firms simultaneously choose to supply individual quantities each period in a homogenous product market. When sellers are price takers, price equals marginal costs, and price is determined such that supply equals demand. However, when firms

[4] . See Section 3 for a discussion of some distributions and their associated equilibrium outcomes.

[5] See Section 5.1 for a dynamic version of the models outlined below.

are not price takers, perceived marginal revenue, and not price, will be equal to marginal cost.

Let's denote $P(Q_t,x_t)$ as inverse demand, Q_t as total industry quantity, $C(q_{it},w_t)$ is the firm i's cost function, q_{it} is the individual firm quantity in period t, and x_t and w_t are vectors of demand and cost shifters observable to all firms in period t, respectively. If firm chooses quantity of output to maximize profit, the FOCs are:

$$P(Q_t, x_t) - mc(q_{it}, w_t) + P'(Q_t, x_t)Q_t \cdot \theta_{it} = 0 \tag{1}$$

where $mc(q_{it},w_t)$ is marginal cost, and θ_{it} is a new parameter that parameterizes the MR=MC optimality condition. Equation (1) (and aggregations of this equation) is the standard model used to justify the Conduct Parameter Method. Under perfect competition, $\theta_{it}=0$ and price equals marginal cost. When $\theta_{it}=1$ we face a perfect cartel, and when $0<\theta_{it}<1$ various oligopoly regimes apply.[6]

The empirical specification of (1) can be obtained by adding a traditional *symmetric* noise term capturing random shocks, ε_{it}. That is:

$$P_t = mc(q_{it}, w_t, \alpha) - P'(Q_t, x_t, \beta)Q_t \cdot \theta_{it} + \varepsilon_{it} \tag{2}$$

where α and β are respectively firm i's marginal cost and demand parameters to be estimated.

Differentiated product

If firm's products are differentiated, firm i's profit function in period t can be written as:

$$\pi_{it} = p_{it} q_{it}(p_{it}, p_{jt}, x_t) - C(q_{it}(p_{it}, p_{jt}, x_t), w_t) \tag{3}$$

where p_{it} is firm i's price, and p_{jt} is the price for firm j, C is the firm i's cost function, and x_t and w_t are vectors of demand and cost shifters observable to all firms in period t,

[6] The estimated conduct parameter does not consistently estimate how rivals response to a change in firm i's behaviour, i.e. it cannot be interpreted as a conjectural variation. Rather it estimates behaviour in equilibrium. i.e. it only indicates that the result of that behaviour is as competitive "as-if" the firms were in fact playing a conjectural variation game with the estimated conjectural variation parameter. The estimated conduct parameter can be interpreted as a summary statistic measuring the degree of market power. See Bresnahan (1989) and Reiss and Wolak (2005) for a good discussion of (mis)interpreting an estimate of the conduct parameter θ_{it}.

respectively. We assume that firms charge different prices and their marginal cost varies across firms and over time. Finally, q_{it} represents firm *i*'s demand function at period *t*.

In a static model, i.e. firms maximize profits without explicit consideration of the effect on future competition, that is:

$$\max_{p_{it}} \quad p_{it}q_{it}(p_{it}, p_{jt}, x_t) - C(q_{it}(p_{it}, p_{jt}, x_t), w_t) \tag{4}$$

The static FOC's are:

$$q_{it} + [p_{it} - mc_{it}]\left\{\frac{\partial q_{it}}{\partial p_{it}} + \frac{\partial q_{it}}{\partial p_{jt}} \cdot \theta_{it}\right\} = 0 \tag{5}$$

where $\theta_{it}=\partial p_{jt}/\partial p_{it}$ is the traditional static conduct parameter. This equation captures, as special cases, several static solutions. If $\theta_{it}=0$ firm's conduct is consistent with one-shot Nash-Bertrand behaviour. If $\theta_{it}=1$, it is perfect collusion. An imperfect collusion arise when $0<\theta_{it}<1$.

The empirical specification of (5) can be obtained by adding again a *symmetric* noise term:

$$q_{it} = [p_{it} - mc_{it}(\alpha)]b_{it}(\beta) - [p_{it} - mc_{it}(\alpha)]d_{it}(\beta) \cdot \theta_{it} + \varepsilon_{it} \tag{6}$$

where $b_{it}(\beta) = -\frac{\partial q_{it}}{\partial p_{it}}$, and $d_{it}(\beta) = \frac{\partial q_{it}}{\partial p_{jt}}$.[7]

General specification of the empirical model

Whatever the product is homogeneous or differentiated, the model to be estimated is quite similar. In both settings the empirical model can be written as:

$$y_{it} = f_{it} + \varepsilon_{it} - g_{it} \cdot \theta_{it} \tag{7}$$

[7] For notational ease, the model above assumes that firm *i* only has one competitor (firm *j*). With N firms the random variable θ_{it} can be interpreted as the firm *i*'s weighted average conduct, and $\frac{1}{N-1}d_{it}$ as the average cross derivative of firm i's demand with respect to the N-1 competitor's prices.

where, depending on the model, f_{it} and g_{it} are functions of cost parameters (α), demand parameters (β), or both cost and demand parameters (α,β). If the product is homogeneous, $y_{it}=P_t$, $f_{it}=mc_{it}(\alpha)$, and $g_{it}=P_t'(\beta)Q_t$. If the product is heterogeneous $y_{it}=q_{it}$, $f_{it}=[p_{it}-mc_{it}(\alpha)]b_{it}(\beta)$, and $g_{it}=[p_{it}-mc_{it}(\alpha)]d_{it}(\beta)$.

Most structural econometrics models treated firm's behaviour as a (common) parameter to be estimated altogether with other cost and demand parameters. In the present paper we propose treating firms' behaviour as a random variable, which follows a one-sided distribution once we incorporate the theoretical restriction that of $0\leq\theta_{it}$ or $\theta_{it}\leq 1$.[8] The distinctive feature of our model is that the stochastic part is formed by two random variables, i.e. the traditional *symmetric* noise term, ε_{it}, and a *one-sided* random conduct term, $g_{it}\cdot\theta_{it}$, that depends on the measure of market power. The one-sided restriction makes asymmetric the composed error term and allows us getting separate estimates of θ_{it} and ε_{it} from an estimate of the composed error term. [9]

While assuming that the noise term is *i.i.d.* and symmetric with zero mean is conventional, several simple distributions for the (one-sided) conduct random term can be estimated. Here, there is a trade-off between tractability and (economic) reasonability. On one hand, we need to choose a *simple* distribution for the asymmetric term in order to make estimable the empirical model. But, on the other hand, each distribution can be associated to a different oligopolistic equilibrium outcome or a specific family of (similar) oligopolistic equilibriums. For this reason, the selected distribution for the one-sided conduct term reflects the researcher's believes about the underlying oligopolistic equilibrium that generates the data. Although, in principle, this feature of the present approach can be viewed as a disadvantage, the types of skewness (or kurtosis) of the conduct random term that can be estimated can be used to test different types of oligopolistic equilibrium outcomes.

[8] Since it is not easy to impose at the same time the theoretical restriction that $0\leq\theta_{it}$ and $\theta_{it}\leq 1$ and allowing for *simple* asymmetric distributions, we will only impose $0\leq\theta_{it}$ or $\theta_{it}\leq 1$, and leave the opposite limit open as a way to capture more sophisticated (e.g. two modal) distributions or as a model's specification test. Note also that if firms are asymmetric in size and cost, the upper limit (i.e. in perfect collusion) in a homogeneous product market is $Q_t/N\cdot q_{it}$, and hence it does not take a particular value except for the case where all firms are of the same size. This is an additional reason to impose a unique restriction on θ_{it} when products are homogeneous.

[9] Note that in a homogeneous setting g_{it} is by construction negative, while in a differentiated setting it is positive. Hence, the one-sided error term is positive (i.e. $-g_{it}\cdot\theta_{it}\geq 0$) in the first case, but negative ($-g_{it}\cdot\theta_{it}\leq 0$) in the second. The asymmetry, however, still holds.

Unlike previous papers that focused on industry-average conduct scores, here we focus on the distribution of this random term and the implications of choosing a specific one-sided distribution in order to get firm-specific market power scores.[10] While some papers have obtained firm-specific conduct parameter using panel data set, they have not explicitly analyzed the skewness and kurtosis of the estimated conduct parameters. If we cannot take advantage of the panel structure of the data (because, for instance, the time dimension is not long enough), firm-specific market power scores can still be obtained analyzing the skewness of the conduct random term. This kind of analysis, in addition, can provide useful information about the nature of the competition among firms in a particular market.

In Section 3 we first review several distributions and their associated oligopolistic equilibriums, whereas in Section 4 we discuss the estimation strategies that we can follow in order to get both the parameter estimates of the pricing equation (7) and individual and industry-average market power scores.

3. Distributional assumptions and equilibrium outcomes

With the aim of estimating firm-specific market power scores, several simple distributions for the (one-sided) conduct random term can be used. All of them make asymmetric the composed error term, but the particular skewness of the conduct random term is associated to a different oligopolistic equilibrium outcome or a specific family of (similar) oligopolistic equilibriums.

For instance, it is common to assume that demand is supplied by few (and likely large) firms that face a fringe of firms. These firms are often assumed to be small and/or foreign firms that do not behave strategically (i.e. they are modelled as a competitive fringe), while the first firms have enough market power to fix prices over marginal cost or/and are involved in some kind of collusion agreement. [11] Unlike previous papers,

[10] It should be highlighted that, in practice, we can estimate the pricing equation (7) and the industry-average market power level without assuming any distribution assumptions for the two random terms (e.g. using some method of moments). This, however, does not allow as getting firm-specific market power scores. In contrast, if we are interested in firm-specific market power scores or we want to estimate all parameters using MLE techniques, we need to choose a distribution for both the one-sided conduct term and the noise term (see Section 4).

[11] See, for instance, Gallet (1997) and Gallet and Schroeter (1995).

here we do not split in advance the sample into firms that behave strategically and competitive firms. Instead, conditional on the distributional assumption on the one-sided error component, our approach allows us to detect firms with (high) market power (or involved in a cartel) or firms behaving (more) competitively. This *incomplete* collusion equilibrium is reasonable in markets with many firms where coordination among all firms is extremely difficult to maintain as the number of firms in the collusive scheme is too high or other market characteristics (e.g. markets with differentiated products) makes coordination too expensive. The most important characteristic of this equilibrium is that the modal value of the conduct random term is close to zero. This situation is illustrated in Figure 1, where we have also assumed that only a few number of firms are involved in almost a perfect collusion scheme. As illustrated in Figure 1, if increasing values of θ_{it} are less likely, several relative simple (i.e. with one-parameter densities) distributions widely employed in the stochastic frontier literature for the one-sided random term, such as half-normal or exponential, [12] fits quite well this market-equilibrium outcome.

[Insert Figure 1 here]

In these markets, the fringe of firms that remains outside the cartel has an incentive to price up to the cartel's level. This phenomenon is called the "umbrella pricing", which yields that some firms set prices higher than those fixed by *pure* competitive (i.e. Bertrand) firms. In this oligopolistic equilibrium, as it is illustrated in Figure 2, the modal value of the conduct random term might be higher than zero (but closer to zero than to one). This equilibrium outcome, and other oligopolistic equilibriums which yield a

[12] The half-normal distribution is obtained from the truncation below zero of a random variable which follows a normal distribution with zero mean and variance σ. In this case, the density function of $\theta_{it} \geq 0$ is given by

$$f(\theta_{it}) = \frac{2}{\sqrt{2\pi}\cdot\sigma}\cdot\exp\left\{-\frac{{\theta_{it}}^2}{2\sigma^2}\right\}$$

If $\theta_{it} \geq 0$ follows an exponential distribution, the density function is

$$f(\theta_{it}) = \frac{1}{\sigma}\cdot\exp\left\{\frac{-\theta_{it}}{\sigma}\right\}$$

Both half-normal and exponential are a single-parameter distributions. There is some evidence in the frontier literature that neither rankings of firms by their efficiency (here conduct) scores or the composition of the top and bottom scores deciles are particular sensitive to the single distribution assigned to the one-sided error term (see Kumbhakar and Lovell, 2000, p. 90).

similar distributional outcomes, can be modelled by allowing θ_{it} to follow a truncated normal distribution.[13]

[Insert Figure 2 here]

All firms in other markets are involved in perfect cartel scheme. In a cartel-equilibrium, firms usually fix prices by agreeing to sell at some “target” price (see, Connor, 2005). This target price is the monopoly price in perfect-collusion equilibrium and it is associated with the maximum conduct value (e.g. θ_{it} =1). This means that the modal value of the conduct random term in this equilibrium is one, with less values of θ_{it} increasingly less likely. As shown in Figure 3, a truncated (over *one*) one-mean normal distribution fits quite well this cartel-equilibrium outcome. It is well known that secret price cuts by cartel members are almost always a problem in cartels. For instance, Ellison (1994) finds that secret price cuts occurred during 25% of the cartel period and that the price discounts averaged about 20%.[14] Here, cheating behaviour explains why firm-conduct and prices are negative skewed compared to the same market in a perfect collusion.

[Insert Figure 3 here]

Firms involved in cartel equilibrium fix prices by agreeing to sell at some lower “floor” (minimum) price. In these industries, the cartel price is high, and it can be associated with a high conduct value (mark-up). This means that the modal value of the conduct random term in this equilibrium is, as shown in Figure 4, less than (but close to) one, with less and high values of θ_{it} increasingly less likely. Here, a truncated (over *one*) normal distribution with nonzero mean fits quite well this cartel-equilibrium outcome. Again, cheating behaviour explains why firm-conduct and prices to become negative skewed. Because the cartel price is not as high as in a perfect-collusion, some sales by cartel members would occur at supra-target prices or mark-ups.

[Insert Figure 4 here]

[13] The truncated normal distributions assumed for θ_{it} generalizes the half normal distribution, by allowing the pre-truncated normal distribution to have a nonzero mode, i.e. it is a two-parameter distribution (see below equation (14)). It should be noted that two-parameter distributions are much more difficult to estimate as shown by Ritter and Simar (1997).

[14] See also Borenstein and Rose (1994).

Two final situations are worthy to mention from an estimation point of view. First, all the previous distribution functions for the skewed conduct random term assume the existence of one mode, with less (high) values of the conduct value increasingly less likely. Many industries, however, may be formed by two groups of firms with two different behaviours, which imply the existence of two modes. In principle, the previous simple one-parameter and two-parameter distributions cannot handle this situation. In addition, the tractability principle prevents us using other (one-sided) distributions which are much more sophisticated and difficult to estimate.

Note, however, that we only impose $0 \leq \theta_{it}$ or $\theta_{it} \leq 1$, and leave open the opposite limit. This allows us to "capture" *indirectly* more sophisticated (e.g. two modal) distributions in the sense that is explained below. Imagine that the conduct variable is distributed as in Figure 5. Here, compared to Figure 1, there are accumulation points not only in zero but also in one. As illustrated in Figure 5, if we estimate a half-normal (i.e. one-mode) distribution allowing for values higher than one, the estimation method will try to estimate a distribution with the same "mass" close to one by enlarging the range of values for the conduct variable up to, say, A is equal to B. Hence, values higher than one might indicate that the model is not well specified, but also that an accumulation point is not *directly* captured by the selected distribution function.

[Insert Figure 5 here]

Second, in all previous situations, we have assumed that the conduct random term is skewed distributed. Since assuming that the noise term is symmetric with zero mean is conventional, this allows us getting firm-specific market power scores. If the conduct random term is distributed as in Figure 6, where both right hand side and left hand side tails are symmetric, we cannot estimate firm-specific market power scores. This is because we cannot get separate estimates of statistical noise and conduct value from estimates of the composed error for each firm. Note, however, that in this situation we can estimate the pricing equation (7) and the industry-average market power level.

[Insert Figure 6 here]

4. Estimation strategy

We now consider how to estimate the model. The traditional structural market power model to be estimated is formed by a demand function and a pricing equation. In order to focus on empirical issues regarding the estimation of industry and firm-specific market power scores, we only discuss the estimation of the pricing relationship (7), given a previous estimation of the demand parameters.[15]

Models of the form in (7) have been proposed and estimated in the literature of frontier production function literature (measuring of the efficiency and production).[16] Two estimation methods are possible: a method of moments approach and MLE. The method of moments approach involves three stages. In the first stage, least squares or GMM is used to generate consistent estimates of all parameters describing the structure of the pricing equation, apart from the variances of both random terms. This stage is thus independent of distributional assumptions on either error component. In the second stage of the estimation procedure, distributional assumptions are invoked in order to obtain consistent estimates of the parameter(s) describing the structure of the two error components, conditional on the first-stage estimated parameters describing the structure of the pricing equation. In the third stage, market power is estimated for each firm by decomposing the estimated residual into a noise component and a market-power component.

The MLE approach uses maximum likelihood techniques to obtain consistent second-stage estimates of the parameter(s) describing the structure of the two error components, conditional on the first-stage estimated parameters, describing the structure of the pricing equation. It can be also used to estimate simultaneously both types of parameters. In this case, the MLE approach merges the two first stages of the method of moments approach into one.[17]

---

[15] This is the strategy followed, for instance, by Nevo (2001) and Jaumandreu y Lorences (2002).

[16] See, in particular, Simar, Lovell and Vanden Eeckaut (1994), and the references in Kumbhakar and Lovell (2000).

[17] Base on the results of a Monte Carlo experiment, Olson, Schmidt, and Waldman (1980) concluded that the choice of estimator (MLE versus method of moments) depends on the relative values of the variance of both random terms and the sample size. When the sample size is small and the variance of the one-sided error component, compared to the variance of the noise term, is not large the method of moments outperforms MLE in a mean-squared error sense.

4.1. First stage

We start describing the method of moments approach. In the first stage, least squares or GMM is used to generate consistent estimates of all parameters describing the structure of the pricing equation, apart from the variances of both random terms. Let us rewrite the pricing equation (7) as:

$$y_{it} = f_{it}(\alpha,\hat{\beta}) - g_{it}(\alpha,\hat{\beta}) \cdot \overline{\theta} + v_{it} \tag{8}$$

where we define f_{it} and g_{it} in general terms as functions of both cost and demand parameters, α are parameters of the firm *i*'s marginal cost function to be estimated, $\hat{\beta}$ are demand parameters already estimated, $\overline{\theta} = E(\theta_{it})$, and

$$v_{it} = \varepsilon_{it} - g_{it}(\alpha,\hat{\beta}) \cdot \{\theta_{it} - \overline{\theta}\} \tag{9}$$

The parameters in equation (8) can be estimated by (non)linear least squares by means of

$$\left(\hat{\alpha}, \hat{\overline{\theta}}\right) = \arg\min \sum_i \sum_t \left[y_{it} - f_{it}(\alpha,\hat{\beta}) + g_{it}(\alpha,\hat{\beta}) \cdot \overline{\theta}\right]^2 \tag{10}$$

Note that the pricing relationship (10) which is estimated in this first stage is equivalent to the traditional specification of a structural market power model where an industry-average market power level is estimated altogether with cost (and, just in case, demand) parameters. Hence, our estimation procedure nets the traditional methods used in the literature. As mentioned in the introduction section, the main contribution of the proposed approach is not the way that both parameters describing the structure of the pricing equation and an average-industry market power are estimated, but the way the asymmetry of the composed error term (i.e. the skewness of the conduct random variable) is exploited in order to get firm-specific market power estimates in the second and third stages.[18]

[18] The pricing equation (8) can be extended allowing for firm-specific conduct parameters (see, for instance, Puller, 2007) as follows:

$$q_{it} = f_{it}(\alpha,\hat{\beta}) + g_{it}(\alpha,\hat{\beta}) \cdot \overline{\theta}_i + v_{it}$$

This model assumes that firm-specific conduct parameters are time-invariant and it is only consistent when long panel data sets are available (i.e. as T→∞). In addition, the incidental parameter problem appears as N→∞.

Two estimation issues arise in this first stage. Since some regressors are endogenous, a GMM or IV method should be employed in order to get consistent estimates. Second, the resulting parameter estimates are consistent, but not efficient by construction since the v_{it}'s are independently but not identically distributed. Assuming that θ_{it} and ε_{it}, are distributed independently of each other, the second moment of the composed error term can be written as:

$$E(v_{it}^2) = \sigma_{vit}^2 = \sigma_\varepsilon^2 + g_{it}^2(\alpha,\beta)\cdot\sigma_\theta^2 \tag{11}$$

where $E(\varepsilon_{it}^2) = \sigma_\varepsilon^2$, and $V(\theta_{it}) = \sigma_\theta^2$. Equation (11) shows that the error in the regression indicated by (9) is heteroskedastic. Therefore generalized least squares would be needed to obtain estimates that are efficient. Efficient parameter estimates can be obtained using weighted least squares by means of

$$\left(\hat{\alpha},\hat{\bar{\theta}}\right) = \arg\min \sum_i \sum_t \left[\frac{y_{it} - f_{it}(\alpha,\beta) + g_{it}(\alpha,\beta)\cdot\bar{\theta}}{\sigma_{vit}}\right]^2 \tag{12}$$

However, since σ_{vit}^2 is unknown, it is necessary to construct a feasible least squares estimator. This is done in the second stage.

4.2. Second stage

Since the conduct random variable is likely skewed (see previous section), the first-stage residual is likely asymmetric, and then its third moment can be used to estimate σ_{vit}^2.[19] Assuming, as it is customary, that ε_{it} is symmetrically distributed, the third moment of the composed error term can be written as

$$E(v_{it}^3) = -g_{it}^3(\alpha,\beta)\cdot E\left[(\theta_{it} - \bar{\theta})^3\right] \tag{13}$$

Equation (14) shows that the third moment of v_{it} is simply the third moment of the random conduct term, adjusted by $-g_{it}^3(\alpha,\beta)$. Note also that the variance of the noise error term does not appear in (13). That is, while the second moment (11) provides

[19] If θ_{it}=0, then $v_{it}=\varepsilon_{it}$, the first-stage residuals are symmetric. However, if θ_{it}>0, then v_{it} is likely positively or negatively skewed. Hence, this suggests that a test that first-stage residuals are symmetric provide a simple test for the existence of non-Bertrand behaviour.

information about the parameter(s) describing the structure of the two error components (i.e. σ_ε^2 and σ_θ^2), the third moment (13) only provides information about the asymmetric random conduct term. Now, if we assume a specific distribution for this one-sided random term, we can estimate σ_θ^2 from the third moment of v_{it}, and then σ_ε^2 from its second moment. As shown below, this provides all the information required to estimate not only and σ_{vit}^2 but also to get firm-specific market power scores.

Assume, for instance, that our market equilibrium outcome can be modelled by allowing θ_{it} to follow a truncated normal distribution. [20] If we assume that $\theta_{it} \sim N^+(\mu,\sigma^2)$ then

$$f(\theta_{it}) = \frac{1}{\sqrt{2\pi}\cdot\sigma}\cdot\Phi^{-1}(\mu/\sigma)\cdot\exp\left\{-\frac{(\theta_{it}-\mu)^2}{2\sigma^2}\right\} \tag{14}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function (see Greene, 2003). Given this distributional assumption for the conduct term, the first three moments of θ_{it} can be written as (see Jawitz, 2004):

$$E(\theta_{it}) = \bar{\theta} = \mu + \frac{\phi(\mu/\sigma)}{\Phi(\mu/\sigma)}\cdot\sigma \tag{15}$$

$$E\left[(\theta_{it}-\bar{\theta})^2\right] = \sigma_\theta^2 = \sigma^2 - \frac{\phi(\mu/\sigma)}{\Phi(\mu/\sigma)}\cdot\sigma\cdot\bar{\theta} \tag{16}$$

$$E\left[(\theta_{it}-\bar{\theta})^3\right] = -\frac{\phi(\mu/\sigma)}{\Phi(\mu/\sigma)}\sigma^3 + \left[\frac{\phi(\mu/\sigma)}{\Phi(\mu/\sigma)}\right]^2\sigma^2\cdot\bar{\theta} + \frac{\phi(\mu/\sigma)}{\Phi(\mu/\sigma)}\cdot\sigma\cdot\bar{\theta}^2 \tag{17}$$

Note that, while the first stage is thus independent of distributional assumptions on either error component, in the second stage of the estimation procedure we invoke distributional assumptions in order to obtain consistent estimates of the parameter(s) describing the structure of the two error components, *conditional* on the first-stage estimated parameters describing the structure of the pricing equation, which includes

[20] We have chosen this distribution because it is a generalization of the one-parameter half-normal distribution and it is one of the most employed in the production frontier literature. See Jawitz (2004) for other distributions and the moments required to estimate then using the method of moments.

an estimation of $\bar{\theta}$. Hence, equation (15) can be viewed as a nonlinear constraint between μ and σ.

Given the assumed distribution function for θ_{it}, the second and third moments of the composed error term can be rewritten as:

$$E(v_{it}^2) = \sigma_{vit}^2 = \sigma_\varepsilon^2 + g_{it}^2(\alpha,\beta)\cdot\left(\sigma^2 - \frac{\phi(\mu/\sigma)}{\Phi(\mu/\sigma)}\cdot\sigma\cdot\bar{\theta}\right) \tag{18}$$

$$E(v_{it}^3) = -g_{it}^3(\alpha,\beta)\cdot\left\{-\frac{\phi(\mu/\sigma)}{\Phi(\mu/\sigma)}\cdot\sigma^3 + \left[\frac{\phi(\mu/\sigma)}{\Phi(\mu/\sigma)}\right]^2\cdot\sigma^2\cdot\bar{\theta} + \frac{\phi(\mu/\sigma)}{\Phi(\mu/\sigma)}\cdot\sigma\cdot\bar{\theta}^2\right\} \tag{19}$$

Next, using the first-stage residuals, the two equations formed by the nonlinear constraint and the third moment of the composed error term

$$\hat{\theta} = \mu + \frac{\phi(\mu/\sigma)}{\Phi(\mu/\sigma)}\cdot\sigma \tag{20}$$

$$\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[\frac{\hat{v}_{it}}{-g_{it}(\hat{\alpha},\hat{\beta})}\right]^3 = -\frac{\phi(\mu/\sigma)}{\Phi(\mu/\sigma)}\sigma^3 + \left[\frac{\phi(\mu/\sigma)}{\Phi(\mu/\sigma)}\right]^2\sigma^2\cdot\hat{\theta} + \frac{\phi(\mu/\sigma)}{\Phi(\mu/\sigma)}\sigma\cdot\hat{\theta}^2 \tag{21}$$

provide estimates of μ and σ, which yield an estimate of $\phi(\mu/\sigma)\cdot\Phi^{-1}(\mu/\sigma)$. Using the second moment of the composed error term, these estimates together yield an estimate of σ_ε^2 by means of

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left\{\hat{v}_{it}^2 - g_{it}^2(\hat{\alpha},\hat{\beta})\cdot\left(\hat{\sigma}^2 - \frac{\phi(\hat{\mu}/\hat{\sigma})}{\Phi(\hat{\mu}/\hat{\sigma})}\cdot\hat{\sigma}\cdot\hat{\theta}\right)\right\} \tag{22}$$

This provides all the information required to estimate σ_{vit}^2, which is used in least squares in equation (12).

The procedure is much more simple if we assume that θ_{it} follows a half-normal distribution, that is, $\theta_{it} \sim N^+(0,\sigma^2)$. In this case θ_{it} comes from a truncation below zero of a normal distribution with zero mean. Note that if μ=0, $\phi(\mu/\sigma)\cdot\Phi^{-1}(\mu/\sigma) = (2/\pi)^{1/2}$, $\hat{\sigma} = \hat{\theta}\cdot(\pi/2)^{1/2}$, and $\hat{\sigma}_\theta^2 = \hat{\theta}\cdot(\hat{\theta}\pi/2 - 1)$. Hence

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left\{\hat{v}_{it}^2 - g_{it}^2(\hat{\alpha},\hat{\beta})\cdot\left(\hat{\bar{\theta}}^2\cdot(\pi/2)-\hat{\bar{\theta}}\right)\right\} \tag{23}$$

These estimates provide all the information required to estimate σ_{vit}^2 and, hence, to carry out equation (12).[21]

Given that we have assumed a particular distribution for the conduct term, the structure of the two error components can be estimated by MLE. Consider, for instance, that θ_{it} ~$N^+(\mu,\sigma^2)$ and ε_{it}~$N(0,\sigma_\varepsilon^2)$. The density function of θ_{it} in (8) is given in equation (14). Hence, it follows that $\tilde{\theta}_{it} \equiv g_{it}(\alpha,\hat{\beta})\cdot\theta_{it} \sim N^+(\mu_{it},\sigma_{it})$ where $\mu_{it} \equiv g_{it}(\alpha,\hat{\beta})\cdot\mu$, and $\sigma_{it} \equiv g_{it}(\alpha,\hat{\beta})\cdot\sigma$. Given the independence assumption, the joint density function of $\tilde{\theta}_{it}$ and ε_{it} is the product of their individual density functions

$$f(\tilde{\theta}_{it},\varepsilon_{it}) = \frac{1}{2\pi\cdot\sigma_\varepsilon\cdot\sigma_{it}}\cdot\Phi^{-1}(\mu/\sigma)\cdot\exp\left\{-\frac{(\theta_{it}-\mu)^2}{2\sigma^2}-\frac{\varepsilon_{it}^2}{2\sigma_\varepsilon^2}\right\} \tag{24}$$

The joint density of $\tilde{\theta}_{it}$ and $\tilde{v}_{it} = \varepsilon_{it} - \tilde{\theta}_{it}$ is

$$f(\tilde{\theta}_{it},\tilde{v}_{it}) = \frac{1}{2\pi\cdot\sigma_\varepsilon\cdot\sigma_{it}}\cdot\Phi^{-1}(\mu/\sigma)\cdot\exp\left\{-\frac{(\theta_{it}-\mu)^2}{2\sigma^2}-\frac{(\tilde{v}_{it}+\tilde{\theta}_{it})^2}{2\sigma_\varepsilon^2}\right\} \tag{25}$$

The marginal density of $\tilde{v}_{it}$ is obtained by integrating $\tilde{\theta}_{it}$ out of $f(\tilde{\theta}_{it},\tilde{v}_{it})$, which yields

[21] So far we have assumed that θ_{it} is positive skewed as in Figures 1 and 2. If, on the other hand, θ_{it} is negative skewed as in Figures 3 and 4, the procedure is roughly the same. Both cartel-equilibrium outcomes represented by Figures 3 and 4 can be modelled as the truncation over *one* of a normal distribution with mean equal to one (Figure 3) or less than one (Figure 4). The truncation over one can be converted into a (more traditional) truncation below zero if we redefine the random conduct term as $\theta_{it} = 1-\theta_{it}^*$, where θ_{it}^*~$N^+(1-\mu,\sigma)$. This implies that the first-stage residual is $v_{it} = \varepsilon_{it} + g_{it}(\alpha,\beta)\cdot\left\{\theta_{it}^* - \bar{\theta}^*\right\}$, where $E(\theta^*) = \bar{\theta}^* = 1-\bar{\theta}$. Unlike (11), a plus sign appears before $g_{it}(\cdot)$ and θ_{it}^* comes from a truncation below zero of a normal distribution with mean equal to $\mu^*=1-\mu$. Here, μ^* and σ can be estimated from (20) and (21) by replacing μ by μ^*, substituting the estimate of $\bar{\theta}$ by the estimate of $\bar{\theta}^*$, and changing the sign of right-hand side of (21). An estimate of σ_ε^2 can be obtained from (22) by replacing μ by μ^*.

$$f(\tilde{v}_{it}) = \frac{1}{\sqrt{2\pi}\cdot\tilde{\sigma}_{it}\Phi(\mu/\sigma)}\cdot\Phi\left(\frac{\mu_{it}}{\tilde{\sigma}_{it}\lambda_{it}} - \frac{\tilde{v}_{it}\lambda_{it}}{\tilde{\sigma}_{it}}\right)\exp\left\{-\frac{(\tilde{v}_{it}+\mu_{it})^2}{2\tilde{\sigma}_{it}^2}\right\}$$

$$= \frac{1}{\tilde{\sigma}_{it}}\cdot\phi\left(\frac{\tilde{v}_{it}+\mu_{it}}{\tilde{\sigma}_{it}}\right)\cdot\Phi^{-1}(\mu/\sigma)\cdot\Phi\left(\frac{\mu_{it}}{\tilde{\sigma}_{it}\lambda_{it}} - \frac{\tilde{v}_{it}\lambda_{it}}{\tilde{\sigma}_{it}}\right) \qquad (26)$$

where $\tilde{\sigma}_{it} \equiv \left(\sigma_{it}^2 + \sigma_\varepsilon^2\right)^{1/2}$ and $\lambda_{it} \equiv \sigma_{it}/\sigma_\varepsilon$. In this case, the likelihood to be maximized is

$$\ln f(\tilde{v}_{it}) = -\ln(\tilde{\sigma}_{it}) + \ln\phi\left(\frac{\tilde{v}_{it}+\mu_{it}}{\tilde{\sigma}_{it}}\right) - \ln\Phi(\mu/\sigma) + \ln\Phi\left(\frac{\mu_{it}}{\tilde{\sigma}_{it}\lambda_{it}} - \frac{\tilde{v}_{it}\lambda_{it}}{\tilde{\sigma}_{it}}\right)$$

or,

$$\ln f(\tilde{v}_{it}) = -\frac{1}{2}\ln\left(g_{it}^2(\hat{\alpha},\hat{\beta})\cdot\sigma^2 + \sigma_\varepsilon^2\right) + \ln\phi\left(\frac{y_{it} - f_{it}(\hat{\alpha},\hat{\beta}) + g_{it}(\hat{\alpha},\hat{\beta})\cdot\mu}{g_{it}^2(\hat{\alpha},\hat{\beta})\cdot\sigma^2 + \sigma_\varepsilon^2}\right) - \ln\Phi(\mu/\sigma)$$
$$+\ln\Phi\left(\left(g_{it}^2(\hat{\alpha},\hat{\beta})\cdot\sigma^2 + \sigma_\varepsilon^2\right)^{-1/2}\cdot\left(\frac{\sigma_\varepsilon\mu}{\sigma\cdot g_{it}(\hat{\alpha},\hat{\beta})} - \frac{g_{it}(\hat{\alpha},\hat{\beta})\cdot\sigma}{\sigma_\varepsilon}\cdot\left(y_{it} - f_{it}(\hat{\alpha},\hat{\beta})\right)\right)\right) \qquad (27)$$

Given the first-stage estimates $(\hat{\alpha},\hat{\beta})$ of the parameters describing the pricing equation structure, maximizing the likelihood function $\ln L = \sum_i \sum_t \ln f(\tilde{v}_{it})$ with respect to the parameters generates estimates of (σ, σ_ε ,μ). If we do not condition the maximization on the first-stage parameter estimates (i.e. firm *i*'s marginal cost, α), we can estimate both the parameters describing the pricing equation structure and the structure of the two error components simultaneously, merging the two previous first stages of the method of moments approach into one.

4.3. Third stage

The third stage is to obtain estimates of market power estimates of each firm. We have estimates of $\tilde{v}_{it} = y_{it} - f_{it}(\alpha,\hat{\beta}) = \varepsilon_{it} - \tilde{\theta}_{it}$, which obviously contains information on θ_{it}. The problem is to extract the information that $\tilde{v}_{it}$ contains on $\tilde{\theta}_{it}$, and given $g_{it}(\cdot)$, on θ_{it}. Jondrow *et. al* (1982) face the same problem in the frontier production function literature and propose using the conditional distribution of the one-sided random term (here $\tilde{\theta}_{it}$) given the composed error term (here $\tilde{v}_{it}$). The conditional distribution f($\tilde{\theta}_{it} \,|\, \tilde{v}_{it}$) is given by

$$\ln f(\tilde{\theta}_{it} \mid \tilde{v}_{it}) = \frac{f(\tilde{\theta}_{it}, \tilde{v}_{it})}{f(\tilde{v}_{it})} = \frac{1}{\sqrt{2\pi} \cdot \sigma_*} \Phi^{-1}\left(\frac{\mu_*}{\sigma_*}\right) \exp\left\{-\frac{(\tilde{\theta}_{it} - \mu_*)^2}{2\sigma_*^2}\right\} \tag{28}$$

Note that $f(\tilde{\theta}_{it} \mid \tilde{v}_{it})$ is distributed as $N^+(\mu_*, \sigma_*^2)$, where $\mu_* \equiv \left(-\sigma_{it}^2 \tilde{v}_{it} + \mu_{it}\sigma_\varepsilon^2\right)/\tilde{\sigma}_{it}$, and $\sigma_*^2 \equiv \left(\sigma_{it}^2 \cdot \sigma_\varepsilon^2\right)/\tilde{\sigma}_{it}^2 = \left(\sigma_{it}^2 \cdot \sigma_\varepsilon^2\right)/\left(\sigma_{it}^2 + \sigma_\varepsilon^2\right)$. Thus the mean of $f(\tilde{\theta}_{it} \mid \tilde{v}_{it})$ can be used to get firm-specific estimates of $\tilde{\theta}_{it}$. The mean is given by

$$E(\tilde{\theta}_{it} \mid \tilde{v}_{it}) = \mu_* + \sigma_* \cdot \frac{\phi(\mu_*/\sigma_*)}{\Phi(\mu_*/\sigma_*)} \tag{29}$$

The mode of this distribution can also be used as a point estimator for $\tilde{\theta}_{it}$. However, the mean is, by far, the most employed in the frontier literature.[22] Once we have got a point estimator for $\tilde{\theta}_{it}$, a conduct score θ_{it} can be obtained using the identity

$\theta_{it} \equiv \tilde{\theta}_{it} / g_{it}(\alpha, \hat{\beta})$.[23]

5. Other econometric issues

5.1. Dynamic pricing

Corts (1999) found in a homogeneous product framework that traditional approaches to estimate the competitive conduct in an oligopoly market can yield inconsistent estimates of the conduct parameter if firms are engaged in an efficient tacit collusion, i.e. firms try to maximize joint profits subject to the constraint that no firm has an incentive to deviate in order to earn higher one-time profits at the risk of starting a "price war". [24]

---

[22] It is also possible to obtain confidence intervals for the point estimates of $\tilde{\theta}_{it}$, by exploiting the fact that the density of is known (see, Horrace and Schmidt, 1996).

[23] Note that $\lambda \equiv \sigma/\sigma_\varepsilon = 0$ implies that $\theta_{it} = 0$. Hence, we can use this parameter to test Bertrand behaviour. Remember that our model allows for firm-specific conduct parameters, so we can identify those firms which are more (less) competitive. To do that, we compute a firm-specific competitiveness level measure by comparing firm's performance with a hypothetical Bertrand equilibrium. In particular, this measure uses Bertrand as "the" competitive framework, and can be calculated as $CL_{it} = \exp(-\theta_{it})$. If $CL=1$ firm's conduct is consistent with static Nash-Bertrand behaviour. If $CL<1$, its value indicates how far the conduct is from a "competitive" outcome.

[24] See e.g. Green and Porter (1984), Rotemberg and Saloner (1986), Haltiwanger and Harrington (1991) and Staiger and Wolak (1992).

The origin of this problem and its possible solutions can be discussed if we extend our static models outlined in Section 2 in a dynamic pricing game framework.[25] Under efficient tacit collusion, each firm solves a dynamic profit maximization problem by comparing the benefit of a deviation from the collusion with the future loss caused by retaliation. In this framework, deviation from the collusive quantity is punished by permanent reversion to a lower profit "punishment" outcome such as Cournot or Nash-Bertrand price. Hereafter, we assume that firms revert to Nash-Bertrand competition during the retaliation period.

Let's us assume a differentiated product market. Let $\pi_{it}^{b}(p,x_t,w_t)$ denote the firm's best response profit at time t and represents the individual profit to any firm that unilaterally deviates from the collusive regime by choosing its one-shot best response to the collusive prices of the other firms. Let $\pi_{is}^{nb}(x_s,w_s)$ denote the punishment profit in period $s>t$. If all other firms play their Nash-Bertrand equilibrium strategy in every period, the best that a single firm can do is to play its Nash-Bertrand equilibrium strategy in each period s. Let $\pi_{it}(p^*,x_t,w_t)$ denote a firm's profit that is obtained when collusion is sustained and p^* is the optimal collusive price in state x_t and w_t.

We can then write the FOC's as follows:

$$\begin{aligned} & p^*(x_t, w_t, \delta) = \arg\max \pi(p, x_t, w_t) \\ & s.t. \\ & \pi_{it}^{b}(p, x_t, w_t) + \sum_{s=t+1}^{\infty} \delta^{s-t} E_t\left[\pi_{it}^{nb}(x_s, w_s)\right] \le \pi_{it}(p, x_t, w_t) \\ & \qquad + \sum_{s=t+1}^{\infty} \delta^{s-t} E_t\left[\pi_{it}(p^*(x_s, w_s, \delta), x_s, w_s)\right] \quad \forall i \end{aligned} \tag{30}$$

where $E_t[\cdot]$ denotes expectations of future profits conditional on information known in period t, and δ is the firms' common discount rate. Following this, the dynamic FOC's are

$$(1+\psi)\left[q_{it} + \left[p_{it} - mc_{it}\right]\left\{\frac{\partial q_{it}}{\partial p_{it}} + \frac{\partial q_{it}}{\partial p_{jt}}\frac{\partial p_{jt}}{\partial p_{it}}\right\}\right] - \psi \cdot \frac{\partial \pi^{b}{}_{it}(p, x_t, w_t)}{\partial p_{it}} = 0 \tag{31}$$

---

[25] See the dynamic model developed by Kim (2006) in a differentiated product setting, and Corts (1999) and Puller (2006) for a similar model developed in a homogeneous product setting.

where ψ is the Lagrange multiplier on the incentive compatibility constraint (ICC). Adding a pure stochastic term that does not affect firm's pricing behaviour, η_{it}, and rearranging terms; this FOC can be written as:

$$q_{it} + \left[p_{it} - mc_{it}\right]\left\{\frac{\partial q_{it}}{\partial p_{it}} + \frac{\partial q_{it}}{\partial p_{jt}} \cdot \theta_{it}\right\} = \varepsilon_{it} \tag{32}$$

where now the random term ε_{it} captures not only errors in optimization (i.e. η_{it}), but also the effect on optimal pricing of the firm in a collusive regime with the binding ICC, that is:

$$\varepsilon_{it} = \eta_{it} + \frac{\psi}{1+\psi} \cdot \frac{\partial \pi_{it}^{b}(p, x_t, w_t)}{\partial p_{it}} \tag{33}$$

Note that θ_{it} is the traditional conduct parameter, but it is different than the static conduct parameter due to in (31) the conduct depends not only on the conduct parameter θ_{it}, but also on whether the incentive compatibility condition binds, i.e. ψ>0. This equation captures, as special cases, several static (i.e. ψ=0) and dynamic solutions (i.e. ψ>0). If θ_i=0 firm's conduct is consistent with Nash-Bertrand behaviour. If, in addition, ψ=0, this outcome is consistent with the static one-shot Nash-Bertrand competition. If θ_{it}=1 and ψ=0, it is perfect collusion. Two imperfect collusions arise. In a static solution when 0<θ_{it}<1. When ψ>0 and θ_{it}=1, conduct is consistent with the dynamic and efficient tacit collusion. Under efficient tacit collusion, firms jointly adjust prices so that no firm has an incentive to deviate from joint profit maximization.

Most market power papers estimate a similar pricing equation to (32) assuming a common conduct parameter to all firms. Corts (1999) demonstrated that the estimates of conduct parameter are biased and inconsistent if the ICC is binding (ψ>0) and the best-response profits are non-linear. If the static model is correct, the error term in an econometric model for (32) is a pure stochastic term and therefore should not affect a firm's pricing behaviour. However, if the ICC is ever binding (i.e. ψ>0) the static conduct parameters are biased and inconsistent.

Since both expected demand and cost shocks affect the ICC, consistent estimates can be obtained by replacing $\{\psi/(1+\psi)\}\cdot\partial\pi^{b}(\cdot)/\partial p_{it}$ by a function of expected demand and cost shocks and estimating the following extended pricing equation:

$$q_{it} + [p_{it} - mc_{it}]\left\{\frac{\partial q_{it}}{\partial p_{it}} + \frac{\partial q_{it}}{\partial p_{jt}} \cdot \theta_{it}\right\} - G(x_t, w_t) = \eta_{it} \tag{34}$$

Based on (34), Puller (2006) pointed out that, under the assumption that both expected demand and cost shocks are common to all firms, Corts' critique can be avoided by estimating the pricing equation (34), but replacing $G(x_t,w_t)$ by a set of time-dummy variables.

Kim (2006) proposes a similar solution to address Corts' critique. Firm's behaviour in (31) depends not only on the conduct parameter θ_{it}, but also on whether the ICC binds, i.e. ψ>0. Corts showed that when the latter is not modelled, the conduct parameter θ_{it} is biased *and* the bias depends on expected future demand and costs. Since firm's dynamic behaviour is influenced by contemporary demand levels, expected future demand, and expected future costs, he suggests modelling the time-varying conduct parameter as a *core* time-invariant conduct parameter, θ_i*, plus a function of $\{\psi/(1+\psi)\}\cdot\partial\pi_{it}^b(\cdot)/\partial p_{it}$. This function is in turns modelled as a linear function of their determinants, i.e. demand and cost shocks. That is:

$$\theta_{it} = \theta_i * + G(x_t, w_t) = \theta_i * + (\varphi_1 \cdot x_t + \varphi_2 \cdot w_t) \tag{35}$$

In equation (34), the first term, the core conduct parameter, measures the firm-specific average level of collusion over time while the second linear term captures the deviation from the average level. Kim mentions two advantages of the above specification. First, by specifying a time-varying conduct parameter we can test the relationship between the firm's conduct and both demand shocks and cost shocks. If x_t has a negative sign, this implies countercyclical firm conduct and mark-up as in Rotemberg and Saloner (1986). If x_t is positively associated with θ, this implies procyclical firm conduct and mark-ups as in Green and Porter (1984). Second, we can shed light on the source of bias that distinguishes the core conduct parameter θ_i* and the static conduct parameter θ when the ICC is no binding, i.e. ψ=0. If we wrongly assume a static game wrongly, what θ measures is not θ_i* but θ_i* plus a bias term. The bias term is a function of demand shock and cost shocks. In this case, market power will be underestimated or overestimated.

Kim assumed that θ_i* is common to all firms, that is θ_i*=θ*, and treated θ* as a common parameter to be estimated. Instead, we do not impose a common behaviour and treat firms' behaviour as a random variable that always takes positive values as it is suggested in oligopoly theory. Kim's suggestion can be incorporated into our model by allowing the random conduct variable θ_{it} be a function of some determinants. This is explored in the next section.

5.2. Conduct determinants

So far we have assumed several basic distributions for the asymmetric random term, viz. half-normal, truncated-normal, etc. We can also develop composed error models that include determinants of (the shape and magnitude) the one-sided random term. This allows us to analyze, for instance, the cyclical behaviour of firm conduct or evaluate bias in static market-power measures (see above), or identify clusters of firms with different strategic behaviour.

Let's us assume that the conduct random term follows a truncated-normal distribution. A general specification including a vector of conduct determinants, z_{it}, can be written as:

$$\theta_{it} \rightarrow N^{+}(\mu(z_{it}), \sigma(z_{it})) \tag{36}$$

where, in addition to other determinants of firms' behaviour, z_{it} might include, as suggested by theory, expected future demand and expected future costs. In this general specification the conduct determinants determine both the shape and magnitude of the one-sided random term, and their coefficients cannot be estimated using a method of moment. They can be estimated by generalizing the maximum likelihood techniques developed in Section 3.

The model can be still estimated using a method of moments, however, if θ_{it} satisfies the so-called *scaling property* (see Wang and Schmidt, 2002). In this case θ_{it} can be written as a scaling function $h(z_{it}, \varphi)$ times a random variable u_{it} that does not depend on z_{it}, that is

$$\theta_{it} \rightarrow h(z_{it}, \varphi) \cdot u_{it} \quad , \quad u_{it} \rightarrow N^{+}(\mu, \sigma) \tag{37}$$

This property implies that changes in z_{it} affect the scale but not the shape of u_{it}. This type of models has a similar economic interpretation than that proposed by Kim (2006). Dynamic models predict that a firm's behaviour is influenced by expected future demand and expected future costs. This effect is capture by the scaling function $h(z_{it}, \varphi)$.[26]

Although it is an empirical question whether or not the scaling property should hold, it has some features that we find attractive. The defining feature of models with the scaling property is that firms differ in their core collusive conduct, but not in the shape of the distribution of the dynamic conduct parameter θ_{it}. The question of whether the effects of the z_{it} on firm behaviour are monotonic can be handled easily by the choice of scaling function. If one wishes to impose monotonicity, simply use a monotonic scaling function, such as the exponential scaling function $exp(z_{it}'\varphi)$. If not, use a non-monotonic scaling function. As noted by Wang and Schmidt (2002), the interpretation of φ does not depend on the distribution of u_{it}, and simple scaling functions yield simple expressions for the effect of the z_{it} on the dynamic conduct parameter θ_{it}. For example, if we use the exponential scaling function, so that $\theta_{it}=exp(z_{it}'\varphi)\cdot u_{it}$, then the coefficients φ are just the derivatives of $\ln(\theta_{it})$ with respect to the z_{it}.

If we assume an exponential scaling function, i.e. $h(z_{it}, \varphi)=exp(z_{it}'\varphi)$, the pricing equation to be estimated is:

$$y_{it} = f_{it}(\alpha, \hat{\beta}) - \tilde{g}_{it}(\alpha, \varphi, \hat{\beta}) \cdot u_{it} + \varepsilon_{it} \tag{38}$$

where $\tilde{g}_{it}(\alpha, \varphi, \hat{\beta}) = g_{it}(\alpha, \hat{\beta}) \cdot \exp(z_{it}'\varphi)$. Hence, except for the new vector of parameter φ, the model to be estimated is the same than (7), and both method of moments and MLE can be used.

5.3. Heteroskedastic noise term

If the symmetric noise term is heteroskedastic we still can get unbiased estimates of all parameters of the pricing equation (8), even if heteroskedasticity is ignored. However, it may bias the firm-specific conduct estimates. An unwarranted assumption of homoskedasticity in the noise-term causes a wrong application of the conditional expectation (29). Since this formula is developed under the assumption of constant variance of the noise-term, heteroskedasticity, if ignored, is improperly attributed to

[26] See Álvarez, Amsler, Orea and Schmidt (2006) for a survey of several models with and without the scaling property in a frontier framework.

either θ_{it}, or ε_{it} when we try to decompose an estimate of the composed error term into these two components.

The procedure outlined in Section 3 can be generalized in order to accommodate for a heteroskedastic noise term. This is achieved by modelling the variance of the noise term as a function of firm-specific size-related variables. The generalized model can be estimated using either maximum likelihood techniques or a method of moments.

We can make the distributional assumption that $\varepsilon_{it} \sim N(0, l(z_{it},\tau)^2\sigma_\varepsilon^2)$. The likelihood function (27) can be generalized in the case in which the noise term is heteroskedastic to

$$\begin{aligned}
&\ln f(\tilde{v}_{it}) = -\frac{1}{2}\ln\left(g_{it}^2(\alpha,\hat{\beta})\cdot\sigma^2 + l(z_{it},\tau)^2\sigma_\varepsilon^2\right) \\
&+\ln\phi\left(\frac{y_{it} - f_{it}(\alpha,\hat{\beta}) + g_{it}(\alpha,\hat{\beta})\cdot\mu}{g_{it}^2(\alpha,\hat{\beta})\cdot\sigma^2 + l(z_{it},\tau)^2\sigma_\varepsilon^2}\right) - \ln\Phi(\mu/\sigma) \\
&+\ln\Phi\left(\left(g_{it}^2(\alpha,\hat{\beta})\cdot\sigma^2 + l(z_{it},\tau)^2\sigma_\varepsilon^2\right)^{-1/2}\cdot\left(\frac{l(z_{it},\tau)\sigma_\varepsilon\mu}{\sigma\cdot g_{it}(\alpha,\hat{\beta})} - \frac{g_{it}(\alpha,\hat{\beta})\cdot\sigma}{l(z_{it},\tau)\sigma_\varepsilon}\cdot\left(y_{it} - f_{it}(\alpha,\hat{\beta})\right)\right)\right)
\end{aligned} \tag{39}$$

Maximizing the likelihood function $\ln L = \sum_i \sum_t \ln f(\tilde{v}_{it})$ with respect to the parameters generates estimates of $(\alpha, \sigma, \sigma_\varepsilon, \mu, \tau)$.

The method of moments' procedure is quite similar to that outlined in Section 3. In the first stage, equation (8) is estimated by linear squares or GMM. This provides an estimate of v_{it}. Given that the distributional assumptions for the conduct term are not modified, the first three moments of θ_{it} are the same than in Section 3. For the truncated normal, they are given by equations (15), (16) and (17). Given the assumed distribution function for θ_{it}, the second moment of the composed error term can be rewritten as:

$$E(v_{it}^2) = \sigma_{vit}^2 = l(z_{it},\tau)^2\sigma_\varepsilon^2 + g_{it}^2(\alpha,\hat{\beta})\cdot\left(\sigma^2 - \frac{\phi(\mu/\sigma)}{\Phi(\mu/\sigma)}\cdot\sigma\cdot\bar{\theta}\right) \tag{40}$$

The third moment is exactly the same than in Section 3. Hence, using the first-stage residuals, the two equations formed by the nonlinear constraint (20) and the third moment of the composed error term (21) still provide estimates of μ and σ. Using the second moment of the composed error term, these estimates together yield an estimate of τ and σ_ε^2 by regressing

$$\ln\left[\hat{v}_{it}^2 - g_{it}^2(\hat{\alpha},\hat{\beta})\cdot\left(\hat{\sigma}^2 - \frac{\phi(\hat{\mu}/\hat{\sigma})}{\Phi(\hat{\mu}/\hat{\sigma})}\cdot\hat{\sigma}\cdot\hat{\bar{\theta}}\right)\right] \tag{41}$$

on the log of $l(z_{it},\tau)^2\sigma_\varepsilon^2$. This provides all the information required to estimate σ_{vit}^2, which is used in NLWLS in equation (12). This NLWLS estimates are then used estimate firm-specific conduct

As in (28), the conditional distribution f($\tilde{\theta}_{it} \mid \tilde{v}_{it}$) is distributed as $N^+(\mu_*, \sigma_*^2)$, but now $\mu_* \equiv \left(-\sigma_{it}^2\tilde{v}_{it} + \mu_{it}\sigma_{\varepsilon it}^2\right)/\tilde{\sigma}_{it}$, and $\sigma_*^2 \equiv \left(\sigma_{it}^2\cdot\sigma_{\varepsilon it}^2\right)/\tilde{\sigma}_{it}^2 = \left(\sigma_{it}^2\cdot\sigma_{\varepsilon it}^2\right)/\left(\sigma_{it}^2+\sigma_{\varepsilon it}^2\right)$ where $\sigma_{\varepsilon it}^2 = l(z_{it},\tau)^2\sigma_\varepsilon^2$. Thus the mean of $f(\tilde{\theta}_{it} \mid \tilde{v}_{it})$ used to get firm-specific estimates of $\tilde{\theta}_{it}$ is given by

$$E(\tilde{\theta}_{it} \mid \tilde{v}_{it}) = \left(-\sigma_{it}^2\tilde{v}_{it} + \mu_{it}\sigma_{\varepsilon it}^2\right)/\tilde{\sigma}_{it} + \frac{\sigma_{it}\sigma_{\varepsilon it}}{\tilde{\sigma}_{it}}\left[\frac{\phi\left(\left(\sigma_{it}^2\tilde{v}_{it}+\mu_{it}\sigma_{\varepsilon it}^2\right)/\sigma_{it}\sigma_{\varepsilon it}\right)}{\Phi\left(\left(\sigma_{it}^2\tilde{v}_{it}+\mu_{it}\sigma_{\varepsilon it}^2\right)/\sigma_{it}\sigma_{\varepsilon it}\right)}\right] \tag{42}$$

If the conduct term follows a half-normal distribution $\mu=\mu_{it}=0$, the expectation (42) can be simplified as

$$E(\tilde{\theta}_{it} \mid \tilde{v}_{it}) = \frac{\sigma_{it}\sigma_{\varepsilon it}}{\tilde{\sigma}_{it}}\left[-\frac{\sigma_{it}\tilde{v}_{it}}{\sigma_{\varepsilon it}} + \frac{\phi\left(-\sigma_{it}\tilde{v}_{it}/\sigma_{\varepsilon it}\right)}{1-\Phi\left(-\sigma_{it}\tilde{v}_{it}/\sigma_{\varepsilon it}\right)}\right] \tag{43}$$

In contrast to (28), now there is a new source of variation in estimated market power among firms, i.e. the noise term with non-constant variance. Hence, if two firms have the same residual, their estimated market power score will differ unless they also have the same noise component variance.

5.4. Panel data specification

So far we have assumed that the θ_{it} are independent (conditional on the z_{it}) over time. This is widely recognized as an unrealistic assumption. For example, we expect market power behaviour to correlate positively over time, because firms that are involved in a collusive scheme this time period will probably also be colluding in other time periods.

Hence, the most plausible departures from independence would involve positive correlations over t, for a given i.

Although independence is likely an unrealistic assumption, it is not clear how to relax it, i.e. how to allow correlation over time in the dynamic conduct parameter. However, under scaling we have the possibility of the following alternative model:

$$\theta_{it} = h_{it}(z_{it}, \varphi) \cdot u_i \tag{44}$$

where u_i is a time-invariant individual effect. Several comments are in order. First, this specification is a restricted version of (37) where we added the restriction $u_{it}=u_i$. Second, the specification in (44) can be viewed as a multiplicative version of the additive conduct decomposition of θ_{it} suggested by Kim (2006). In fact, the term u_i and $h_{it}(z_{it}, \varphi)$ can be viewed as the so-called (time-invariant) core conduct parameter θ_i* and the function of $\{\psi/(1+\psi)\} \cdot \partial\pi_{it}^{b}(\cdot)/\partial p_{it}$ that is modelled as a function of their determinants, i.e. demand and cost shocks respectively. [27]

Third, a distinctive feature of the model is the interaction between the time-varying function $h_{it}(z_{it}, \varphi)$ and the individual effect u_i. As Han, Orea and Schmidt (2005) shown, a "fixed-effects" estimation of this type of models is not trivial due to the incidental parameters problem" i.e. the number of parameters grows with sample size (see, for example, Chamberlain, 1980). Models of this form have been proposed in the literature of production frontier functions. Orea and Kumbhakar (2005) have estimated a model with a specification of one-sided random term (the efficiency of production) equivalent to (44). Their model is in turn a slight generalization of those introduced by Kumbhakar (1990) and Battese and Coelli (1992) where z_{it}=t. All of these papers considered a "random- effects" treatment and proposed specific (truncated normal) distributions for the u_i, with estimation by maximum likelihood.

A related point is that maximum likelihood estimates based on the assumption of independent observations are consistent even if the observations are not independent,

[27] The scaling property in (44) corresponds to a multiplicative decomposition of θ_{it}. An alternative that has sometimes been proposed in the literature on frontier production functions (Huang and Liu (1994), Battese and Coelli (1995), Simar and Wilson (2003)) is an additive decomposition of the form $\theta_{it}(z_{it},\varphi) = h(z_{it},\varphi) + \tau_{it}$. However, this can never actually be a decomposition into independent parts, because $\theta_{it}(z_{it},\varphi) \geq 0$ requires $\tau_{it} \leq h(z_{it},\varphi)$.

so long as the (marginal) distribution of each observation is correctly specified. Thus, for example, MLE estimates from (27) will be consistent even if the conduct term θ_{it} is not independent over time, so long as the model is otherwise correctly specified. However, the estimated variances (or standard errors) of the estimated parameters, calculated under the assumption of independence, will not be correct if independence does not hold. It is possible to calculate asymptotically valid "corrected" estimated variances that allow for non-independence of unspecified form. These points are known in the econometric literature. For example, see Hayashi (2000) and Álvarez, Amsler, Orea and Schmidt (2006).

An important feature of all previous models that assume that θ_{it} is independent over time is that firm observed in two periods is treated as two different firms. This assumption does not allow us to estimate a firm-specific conduct score consistently since its variance does not vanish as the sample size increases. However, since the random conduct term in (44) is developed in a panel data framework (i.e. $\text{cov}(\theta_{it}, \theta_{it-1})\neq 0$), it allows us to estimate market power levels consistently when $T\rightarrow\infty$. A detailed discussion of this issue can be found in Schmidt and Sickles (1984) and Greene (1993).

5.5. Uncaptured differences

The proposed approach can be viewed as belonging to the same family of Porter (1983), Brander and Zhang (1993), and Gallet and Schroeter (1995) who estimate a regime-switching model where market power enters in the model as a supply shock. As in our model, the identification of market power in these papers relies on making simple assumption about a specific component in the error term which is unobservable.

To finish up, it should be noted that since the inference is based on a error component, uncaptured differences among firms and over time (e.g. unobserved changes in factor prices or shocks in technology) might wrongly interpreted as differences or changes in conduct due to both phenomena shift the supply relationship. Hence, estimating market power with a composed error model requires capturing most of the variables which enters in the supply relationship.

6. Empirical illustration: the Spanish banking sector

In this section we illustrate the proposed approach using a balanced panel data set from the Spanish banking industry for the period 1993-1998. To accomplish this, we specify a model where banks sell differentiated products and choose price so as to maximize their profits comparing the benefit of a deviation from collusion with the expected future loss from the deviation. The basic empirical model to be estimated is provided by the equation (5), that is:

$$q_{it} = [p_{it} - mc_{it}(\alpha)]b_{it}(\beta) - [p_{it} - mc_{it}(\alpha)]d_{it}(\beta)\cdot\theta_{it} + \varepsilon_{it} \tag{45}$$

Since there are more than one competitor, θ_{it} can be interpreted as the average conduct of bank *i* with respect to all its rivals and $d_{ij,t}$ is now an aggregate measure of all cross derivatives of bank *i* demand. Let's us assume that θ_{it} satisfies the *scaling property*. In particular, here we assume that θ_{it} can be written as an exponential scaling function, i.e. $h(z_{it},\varphi)=exp(z_{it}'\varphi)$, times a random variable u_{it} that follows a half-normal distribution, that is

$$\theta_{it} \rightarrow \exp(z_{it}'\varphi)\cdot u_{it} \quad , \quad u_{it} \rightarrow N^{+}(0,\sigma) \tag{46}$$

This type of models has a similar economic interpretation than that proposed by Kim (2006). Dynamic models predict that a firm's behaviour is influenced by expected future demand and expected future costs. This effect is capture by the scaling function. Note that this model combines the traditional structural econometric market power models and the traditional frontier approach to measure firms' inefficiency.

The structural model to be estimated is formed by a demand function and the pricing equation (45)-(46). Following Nevo (2001) and Jaumandreu y Lorences (2002), we first estimate the demand side parameters and, given the estimated demand surface, we then estimate the pricing relationship. Estimating the demand equation separately from the pricing equation (i.e. the supply side) does not affect the consistency of the estimates.

6.1. Empirical specification of the demand function

The functional forms for the own-demand derivatives and cross-demand derivatives appearing in the pricing equation are derived from a *nested logit model*. This model is based on the discrete choice random utility framework outlined in Berry (1994).[28] This framework enables us to estimate demand for differentiated products using firm-level data on sales, prices and other attributes.

The approach based on discrete choice models was employed in the banking industry in several papers. For instance, Dick (2002) estimates a structural demand model for commercial bank deposit services for the US. Based on *logit* model estimates, the results indicate that consumers respond to deposit rates, and to a lesser extent, to account fees, when choosing their deposit's bank. Also with bank-level data, Nakane, Alencar, and Kanczuk (2005) model the demand for time deposits, for an aggregate of demand and passbook savings deposits and for loans in Brazil with a multinomial *logit* specification. Gollier and Ivaldi (2005) estimate a *logit* model altogether the pricing equation in order to measure the unilateral effects of merger in the Portuguese Insurance Market.

To get the own-demand derivatives and cross-demand derivatives we estimate a *nested logit model*. For historical reasons, we assume two different types of banks: private banks and saving banks. Loans within the same group are closer substitutes than loans from other groups. In this framework, consumers have a choice of purchasing a loan to one of the sample (private or saving) banks or purchasing to a bank outside the sample.

Given the set of available banks, consumers are assumed to select the bank that gives them the highest utility. Given the distributional assumptions on consumer tastes and functional form for utility, we can aggregate over individual consumer purchases to obtain predicted aggregate market share s_i for bank i:

$$s_{it} = \frac{e^{\delta_{it}/(1-\rho)}}{D_{gt}} \cdot \frac{{D_{gt}}^{1-\rho}}{\sum_{g} {D_{gt}}^{1-\rho}} \quad where \quad D_{gt} = \sum_{i \in g} e^{\delta_{it}/(1-\rho)} \tag{47}$$

The first term in this expression is bank *i*'s market share in its market segment, while the second is the market share of a market segment g in the overall loans market. δ_{it} is the utility level that product *i* yields to consumers, which depends on the price of the

[28] In the industrial organization literature, Berry, Levinsohn, and Pakes (1995), Nevo (2001), Irwin and Pavcnik (2004) and Ivaldi and Verboven (2005) applied discrete choice models to the analysis of market power using firm-level data.

product *i*, a vector of observed characteristics of bank *i*, and an error term ξ_{it} reflecting unobserved characteristics:

$$\delta_{it} = x_{it}\beta - \alpha p_{it} + \xi_{it} \tag{48}$$

The parameter ρ lies between 0 and 1 and measures the substitutability between products within a group. The higher the ρ, the more correlated the consumer tastes are for product within the same market segment and the competition among products is stronger within than across market segments. When ρ=0, consumer tastes are independent across all banks' loans (i.e. there is no market segmentation between private banks and saving banks). In this case, we have the standard logit model in which products compete symmetrically.

Since it is assumed that the outside product yields zero utility, δ_{Ot}=0 and D_{Ot}=1. Berry (1994) showed that rearranging the above equation yields the following demand equation:

$$\ln S_{it} - \ln S_{Ot} = x_{it}\beta - \alpha p_{it} + \rho \ln S_{i|g,t} + \xi_{it} \tag{49}$$

where S_{it} is observed market share of product *i*, S_{Ot} is the observed market share of the outside product, and $S_{i|g}$ is the observed market share of product *i* within its market segment *g*. For the model to be consistent with (random) utility maximization, α has to be positive and ρ has to lie between 0 and 1. The total quantity sold of product *i* q_{it} is simply given by the observed market share of product *i*, S_{it} , times the total market size, M.

The expressions for own and cross-price elasticities of demand derived from the market shares are as follows:

$$\varepsilon_{i,t} = \frac{\partial S_{it}}{\partial p_{it}}\frac{p_{it}}{S_{it}} = -\alpha p_{it} S_{it} + \alpha p_{it}\left(\frac{1}{1-\rho} - \frac{\rho}{1-\rho} S_{i|g,t}\right) \tag{50}$$

$$\varepsilon_{ij,t} = \frac{\partial S_{it}}{\partial p_{jt}}\frac{p_{jt}}{S_{it}} = -\alpha p_{jt} S_{jt} \quad if \quad j \notin g, i \in g \tag{51}$$

$$\varepsilon_{ij,t} = \frac{\partial S_{it}}{\partial p_{jt}} \frac{p_{jt}}{S_{it}} = -\alpha p_{jt} S_{jt} \left[\frac{\rho}{1-\rho} \frac{S_{j|g,t}}{S_{jt}} + 1 \right] \quad if \quad i, j \in g \tag{52}$$

So the expressions for own and cross-price derivatives of demand are as follows:

$$\frac{\partial q_{it}}{\partial p_{it}} = M \cdot \frac{\partial S_{it}}{\partial p_{it}} = M \cdot \varepsilon_{i,t} \frac{S_{it}}{p_{it}} = M \cdot \left\{ -\alpha S_{it}{}^{2} + \alpha S_{it} \left(\frac{1}{1-\rho} - \frac{\rho}{1-\rho} S_{i|g,t} \right) \right\} \tag{53}$$

$$\frac{\partial q_{it}}{\partial p_{jt}} = M \cdot \frac{\partial S_{it}}{\partial p_{jt}} = M \cdot \varepsilon_{ij,t} \frac{S_{it}}{p_{jt}} = -\alpha S_{jt} S_{it} M \quad if \quad j \notin g, i \in g \tag{54}$$

$$\frac{\partial q_{it}}{\partial p_{jt}} = M \cdot \frac{\partial S_{it}}{\partial p_{jt}} = M \cdot \varepsilon_{ij,t} \frac{S_{it}}{p_{jt}} = -\alpha_{jt} S_{jt} S_{it} \left[\frac{\rho}{1-\rho} \frac{S_{j|g,t}}{S_{jt}} + 1 \right] \cdot M \quad if \quad i, j \in g \tag{55}$$

6.2. Empirical Results

The data cover loan prices, sales, and characteristics by the most important private and savings banks from 1993 to 1998 in Spain. Table 1 presents the descriptive statistics of the data.

First we estimate the demand equation using bank level data. The price is the real effective loan interest rate, which is the preference interest rate that each bank communicates to the Bank of Spain, adjusted by the IPC index, plus a commission fee ratio. As a measure of proximity we also include the size of the network of branches measured by their number. There are three issues in estimating the market share equation (49). First, although the econometrician does not observe loans quality ξ_{it}, the banks likely set the price to reflect the product quality. The bank prices are, therefore, likely correlated with unobserved quality. Second, the within-group market share $S_{i|g,t}$ are also likely correlated with ξ_{it}. We instrument for the two variables with two types of instruments: cost-shifters (such as the interbank rates and managerial cost) and the shares and prices lagged one period. Cost shifters affect product prices, but are uncorrelated with loans *i*'s unobserved quality.[29] The demand equation is linear in all parameters and the error term, so it was estimated by two-stage least squares (IV).

---

[29] Similarly, rival banks' characteristics (e.g. branch network) influence the market share and prices of rival bank, and through strategic interaction, also affect the pricing decisions and

Table 2 presents the estimation results. The first two columns report the OLS estimates of the demand parameters and the last two columns report two-stage least squares estimates (IV). Accounting for the endogeneity of price and within market segment market share affects the estimated parameters. For example, the OLS estimate of the price coefficient is -0.0186, while the magnitude of coefficient on price increases (in absolute value) in the IV regression (-0.2723). The estimated value of ρ is 0.43, which suggests that within private banks' loans market segment are better substitutes for each other than loans across the private-saving market segments.

This substitutability of products is quantified in Table 3 that presents the means of the own and cross price elasticities of demand. The average demand elasticity decreases in absolute value over time, averaging about -3.47 in 1993 to-2.36 in 1998. These estimates suggest that a 1% increase in the price lowers a bank's market share by 3.47% (2.36%) in 1993 (1998). Thus, the Spanish loans market appears to have become less price sensitive over time. Within a year, the own-price elasticities also differ across banks, for example, ranging from -3.17 for bank 153 to -1.43 for bank 124 in 1998. In addition, the estimates of the cross-price elasticities reported in column 2 (for products in the same market segment) and 3 (for product in different market segments) suggest that loans within each market segment are closer substitutes for each other than loans across the segments. For example, the average cross-price elasticity 1998 suggests that a 1% increase in the price of a bank leads on average to 0.11% increase in the market share of the banks' loans in the same segment and only 0.03% increase in the market share of the banks' loans in a different market segment.

Once we have estimated the demand parameters, we computed each bank own and cross derivatives using both (44) and (45). Next, assuming that u_{it} follows a half-normal distribution, the pricing equation is estimated assuming, as it is customary in this literature, that banks' marginal costs are constant and independent of the level of operation. We specify a bank's marginal cost, $mc_{it}(\alpha)$, as the sum of both financial cost and managerial cost. On the one hand, interbank rates are the right variable for the marginal financial cost of funds under the common assumption of separability of loans and deposits (see, for example, Freixas and Rochet, 1997). On the other hand, we

market shares of the bank *i* in question. Since, however, they are not econometrically correlated with bank *i*'s unobserved quality ξ_{it}. We have used both rival bank branches and prices as instruments. But they did not work well.

assume that most managerial cost are fixed and are function of the size of the network of branches. We include also a private-bank dummy variable to account for differences between private banks and savings banks.

Regarding the dynamic conduct parameter, θ_{it}, we modelled the scaling function $\exp(\varphi_1 x_t+\varphi_2 w_t)$ as a function of demand shock, x_t, and cost shocks, w_t. To serve as a demand shock, x_t, we include current industry output divided by expected future output. As a proxy for future output, industry output at t+1 is used. For the cost shock, w_t, we use expected future cost rather than contemporary cost. The future cost shock is approximated by the interbank rate at t+1. If only static profit maximization matters, the parameters φ_1 and φ_2 should be equal to zero. We specify x_t and w_t in a mean deviation form.

The pricing equation, which contains 9 parameters, is estimated by ML. It is fitted on a data set covering the period 1994-1998. As the real loan interest rate is endogenous we have previously instrumented this variable.[30] Estimating results are gathered in Table 4. Most parameters are significant. The parameter of the interbank rate is positive, but less than one, which is the value we expected. [31] The constant term for private banks is lower than that for savings banks. The number of branches held by each bank has a negative effect on marginal cost, as found by Gollier and Ivaldi (2005) for Portuguese insurance firms.

Regarding both random variables, note that $\lambda=\sigma_u/\sigma_v$ is statistically significant. Since λ=0 implies that θ_{it}=0 we can use this parameter to test Bertrand behaviour. Our results indicate thus that we can reject that bank's conduct is consistent with Nash-Bertrand behaviour. Remember that, unlike previous empirical models, our model allows for firm-specific conduct parameters, so we can identify those banks which are more (less) competitive. To do that, we compute a bank-specific competitiveness level measure by comparing bank's performance with a hypothetical Bertrand equilibrium. In particular, this measure uses Bertrand as "the" competitive framework, and can be calculated as:

[30] As instruments we have used the number of branches of other competitors in the same market segment and the interest rate of all the other competitors.

[31] This means that the interbank rate is not a good proxy for financial marginal cost during the lifetime of all actual loans. For new loans one lead seems a sensible specification for this cost during a period in which variable interest loans have generalized (Jaumandreu, Lorences and Orea, 2005). Jaumandreu and Lorences (2002) use, however, a polynomial in anticipations for a period in which fixed rates loans were prevalent.

$$CL_{it} = \exp(-\hat{\theta}_{it}) = \exp(-\exp(z_{it}'\hat{\varphi})\cdot\hat{u}_{it}) \qquad (56)$$

where u_{it} is estimated from the overall error term using a modified version of Jondrow et. al. (1982).

If CL_{it}=1 firm's conduct is consistent with static Nash-Bertrand behaviour. If CL_{it}<1, its value indicates how far the conduct is from a "competitive" outcome. The estimated CL measure for each bank in 1996 (the mid of the sample period) are shown in Figure 7. We see that while the performance of most of banks is quite close to the competitive outcome (i.e. their CL measure is close to 1), there are a few banks that their CL measure are about 0.6. This suggests that imposing common conduct behaviour to all banks would yield biased results.

[Insert Figure 7 here]

Regarding the time path of banks' behaviour, both φ_1 and φ_2 parameters are negative and statistically different from zero. This result indicates that the dynamic conduct parameter θ_{it} is countercyclical to current demand shock and expected future cost increase. Figure 8 shows that loans market is increasing while interbank rate (a measure of financial cost) is decreasing. If expected future cost decreases, then the expected loss from the deviation will increase.

[Insert Figure 8 here]

This provides firms with an incentive to converge towards full collusion and, hence, the collusive market price must not be lowered to prevent the deviation. Figure 9 seems to confirm this conclusion. Since expected demand (cost) shocks are positive (negative), there is a deterioration of competition from 1994 to 1998.

[Insert Figure 9 here]

7. Summary and main conclusions

Measuring the degree of competition in oligopolistic markets is a key activity in empirical industrial organization. Earlier studies focused on estimating conduct parameters that distinguish collusive behaviours from non-collusive behaviours. All the structural econometrics models treat the firm's behaviour as a parameter to be estimated altogether with other cost and demand parameters. Dynamic models suggest the conduct parameter is random, changes over time and varies across firms. Obviously, allowing conduct parameters vary both by firm and observation results in an overparameterized model. The empirical models proposed in the literature avoid this problem by putting some *structure* on the way the conduct parameter varies across firms and time.

Instead of treating firms' behaviour as a (restricted) parameter to be estimated, we propose treating firms' behaviour as an asymmetric random variable. Therefore, our approach relies on the estimation of a "composed error" model where the stochastic part is formed by two random variables, i.e. the traditional noise term, capturing random shocks, and a random conduct term, which measures market power. This model allows us to get firm-specific market power scores even though when we cannot take advantage of the panel structure of the data (e.g. the time dimension is not long enough).

While some papers have obtained firm-specific conduct parameter using panel data set, they have not explicitly analyzed the skewness and kurtosis of the conduct random term. In addition to provide firm-specific market power estimates, this kind of analysis provides useful information about the nature of the competition among firms in a particular market. As we aware, skewness or kurtosis of conduct parameter in oligopolistic industry settings are not examined explicitly in most of the previous empirical papers.

We outlined also how to estimate the model using a method of moments approach and MLE, and discuss some econometric issues regarding estimation under dynamic behavior, conduct determinants, noise heteroskedasticity and panel data specifications. The outlined methods roots on the frontier production function literature. The main contribution of the proposed approach is not the way that a traditional pricing equation is estimated, but the way the asymmetry of the composed error term is exploited in order to get firm-specific market power estimates. In particular, once all parameters describing the structure of the pricing equation and an average-industry market power are estimated using well-known econometric techniques, distributional assumptions are invoked to obtain consistent estimates of the parameters describing the structure of the two error components. Conditional on these estimates, market power is estimated for each firm by decomposing the estimated residual into a noise component and a market-power component.

We estimated this model using a balanced panel data set from the Spanish banking industry for the period 1993-1998. Our preliminary results show that imposing common conduct behaviour to all banks would yield biased results. In addition, we find a deterioration of competition during this period since expected demand increase and expected decreases in costs.

References

Abreu, D. Pearce, D and Stacchetti, E. (1986), “Optimal cartel equilibria with imperfect monitoring”, *Journal of Economic Theory*, 39, 251-269.

Álvarez, A., Amsler, C., Orea, L., y Schmidt, P. (2006), "Interpreting and testing the scaling property in models where inefficiency depends on firm characteristics", *Journal of Productivity Analysis*, 25, 201-212.

Appelbaum, E., (1982), “The Estimation of the Degree of Oligopoly Power,” *Journal of Econometrics*, 19, 287-299.

Battese, G.E. and T.J. Coelli (1992), “Frontier Production Functions, Technical Efficiency and Panel Data: With Applications to Paddy Farmers in India,” *Journal of Productivity Analysis*, 3, 153-169.

Battese, G.E. and T.J. Coelli (1995), “A Model for Technical Efficiency for Panel Data”, *Empirical Economics*, 20, 325-332.

Berry, S. T., J. Levinsohn and A. Pakes (1995), Automobile Prices in Market Equilibrium, *Econometrica*, 841-890.

Berry. S. T. (1994). Estimating Discrete-choice Models of Product Differentiation. *Rand Journal of Economics*. 25. 242-262.

Borenstein, S., and N.L. Rose (1994), “Competition and price dispersion in the US airline industry”, *Journal of Political Economy*, 102,653-683.

Brander, A. J. and A. Zhang, (1993), “Dynamic Oligopoly Behavior in the Airline Industry”, *International Journal of Industrial Organization*, 11, 407-435.

Bresnahan, T. (1982), “The oligopoly solution concept is identify”, *Economic Letters*, 10, 87-92.

Bresnahan, T. (1987), “Competition and Collusion in the American Automobile Industry: The 1955 Price War”, *Journal of Industrial Economics*, June, 457-482.

Bresnahan, T. (1989), “Empirical Studies of industries with market power”, in Schmalensee, R. and Willig, R. (eds), *Handbook of Industrial Organization*, Vol. 2, chapter 17, 1011-1057.

Chamberlain, G.C. (1980), “Analysis of covariance with qualitative data”, Review of Economics Studies, 47, 225-238.

Connor (2005).”Collusion and price dispersion”, *Applied Economic Letters*, 12, 335-338.

Corts, K. (1999), “Conduct Parameters and the Measurement of Market Power,” *Journal of Econometrics*, 88, 227-250.

Dick, A. (2002),”Demand estimation and consumer welfare in the banking industry", Discussion Paper 2002-58.

Ellison (1994) “Theories of cartel stability and the Joint Executive Committee”, Rand *Journal of Economics*, 25(1), 37-57.

Freixas, X. and J. Rochet (1997), Microeconomics of Banking, MIT Press, Cambridge MA.

Gallet, A.G. (1997), “Cyclical fluctuations and coordination in the US steel industry”, *Applied Economics*, 29, 279-285.

Gallet, A.G. and John R. Schroeter, (1995), “The Effects of the Business Cycle on Oligopoly Coordination: Evidence from the U.S. Rayon Industry”, *Review of Industrial Organization*, 181-196.

Gallop, D., and Roberts, M., (1979), “Firm Interdependence in Oligopolistic Markets,” *Journal of Econometrics*, 10, 313-331.

Gollier, C. And M. Ivaldi (2005), “Study on competition policy in the Portuguese insurance sector: econometric measurement of unilateral effects in the CAIXA/BCP merger case”, Discussion Paper, Autoridade da Concurréncia do Portugal.

Green, E.J., Porter, R.H. (1984), “Noncooperative Collusion Under Imperfect Price Information”, *Econometrica*, 52, 87-100.

Greene, W.H. (1993), The Econometric Approach to Efficiency Analysis. In: Fried, H., Lovell, C.A.K. and Schmidt, S. (eds.) *The Measurement of Productive Efficiency: Techniques and Applications*, Oxford University Press, New York.

Greene, W.H. (2003), *Econometric Analysis*, Prentice Hall International Editors.

Haltiwanger, J., Harrington, J.E. (1991), “The Impact of Cyclical Demand Movements on Collusive Behavior”, *RAND Journal of Economics*, 22, 89-106.

Han, C., L. Orea, and P. Schmidt (2005),”Estimation of a Panel Data Model with Parametric Temporal Variation in Individual Effects”, *Journal of Econometrics*, 126, 241-267.

Hayashi, F. (2000), *Econometrics*, Princeton University Press.

Horrace, W.C., and P. Schmidt (1996), “Confidence Statements for Efficiency Estimates from Stochastic Frontier Models”, *Journal of Productivity Analysis*, 7(2/3), 257-282.

Huang, C.J. and J.T. Liu (1994),”Estimation of Non-neutral Stochastic Frontier production Function”, *Journal of Productivity Analysis*, 5(2), 171-180.

Irwin, D. and N. Pavcnik (2004), “Airbus versus Boeing revised: international competititon in the aircraft market”, *Journal of International Economics*, 64, 223-245.

Ivaldi, M., and F. Verboven (2005): “Quantifying the Effects from Horizontal Mergers in European Competition Policy", *International Journal of Industrial Organization*, 23(9), 699-702.

Iwata, G., (1974), “Measurement of Conjectural Variations in Oligopoly,” *Econometrica* 42, 949-966.

Jaumandreu, J. and J. Lorences (2002), “Modelling price competition across many markets (An application to the Spanish loans market),” *European Economic Review*, 46, 93-115.

Jaumandreu, J., J. Lorences, and L. Orea (2004) “Growing to become European or European alibi for market power? An analysis of the Spanish banking mergers”, Discussion Paper.

Jawitz, J.W. (2004), “Moments of truncated continuous univariate distributions”, *Advances in Water Resources*, 27, 269-281.

Jondrow, J., C.A.K. Lovell, S. Materov, and P. Schmidt (1982), “On the estimation of technical efficiency in the stochastic frontier production function model”, Journal of Econometrics, 19(2/3), 233-238.

Kim, D. (2006), “Measuring Market Power in a Dynamic Oligopoly Model: The Dallas-Forth Worth Milk Market Case”, *Economic Development & Policy Series* EDP06-4, GSIR Working Papers.

Kumbhakar, S.C. (1990), “Production frontiers, Panel Data, and Time-Varying Technical Efficiency”, *Journal of Econometrics*, 46(1/2), 201-212.

Kumbhakar, S.C. and Lovell, C.A.K. (2000), *Stochastic Frontier Analysis*. Cambridge University Press.

Nevo, A. (2001): “Measuring Market Power in the Ready-to-Eat Cereal Industry", *Econometrica*, 69(2), 307-342.

Olson, J.A., P. Schmidt, and D.M. Waldman (1980), “A Monte Carlo Study of Estimators of Stochastic Frontier Production Functions”, *Journal of Econometrics*, 13(1), 67-82.

Orea, L., y Kumbhakar, S (2004): “Efficiency Measurement using a Latent Class Stochastic Frontier Model”, *Empirical Economics*, 29, 169-183.

Porter, R.H. (1983), "Optimal Cartel Trigger Price Strategies." *Journal of Economic Theory*, 29, 313-338.

Puller, S. (2006), “Estimation of Competitive Conduct when Firms are Efficiently Colluding: Addressing the Corts Critique” *Applied Economics Letters*, forthcoming.

Puller, S. (2007), “Pricing and Firm Conduct in California’s Deregulated Electricity Market”, *Review of Economic and Statistics*, forthcoming.

Reiss, P.C. and F.A. Wolak (2005), “Structural Econometric Modeling: Rationales and Examples from Industrial Organization”, Mimeo, Stanford.

Rotemberg, J.J., Saloner, G. A (1986), “Supergame-Theoretic Model of Price Wars During Booms”, American Economic Review, 76, 390-407.

Schmidt, P. and R. Sickles, (1984), “Production Frontiers and Panel Data”, *Journal of Business and Economic Statistics* 2(4), 367-374

Simar, L., CAK Lovell, and P. Vanden-Eeckaut (1994), “Stochastic Frontiers Incorporating Exogenous Influences on Efficiency”, Discussion Paper Nº 9403, Institut de Statistique, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.

Staiger, R.W., Wolak, F.A. (1992), “Collusive Pricing with Capacity Constraints in the Presence of Demand Uncertainty”, *RAND Journal of Economics,* XX.

Stigler, G.J. (1964), “A Theory of Oligopoly”, Journal of Political Economy, 72, 44-61.

Wang, H.J. and P. Schmidt (2002), “One-Step and Two-Step Estimation of the Effects of Exogenous Variables on Technical Efficiency Levels,” *Journal of Productivity Analysis*, 18, 129-144.

Table 1. Descriptive statistics

Variable	Definition	Units	Mean	Std.	Min	Max
Preference interest rate[1]	r_{it}	%	8.52	2.07	4.00	12.80
Commission fee ratio[2]	f_{it}	%	0.70	0.36	0.16	1.90
Consumption price index rate[3]	ipc_{it}	%	3.37	1.29	1.41	4.93
Real effective loan interest rate	p_{it}=-r_{it}-f_{it}-ipc_{it}	%	5.86	1.17	3.11	8.46
Within-group market share[5]	$S_{i\|g,t}$	%	7.1	7.0	1.0	26.9
Outside Loans share[5]	S_{Ot}	%	37.0	1.8	33.8	39.5
Branches[5]	Br	Thousand of offices	0.80	0.77	0.06	3.69
Interbank rate[4]	i_{it}	%	7.65	2.46	4.00	10.90
Real interbank rate	ri_{it} =i_{it}- ipc_{it}	%	4.28	1.16	2.59	5.97
Industry output[5]	Q_t	10000 of million Euros of 1998	3.47	0.43	3.07	4.25

Notes:

[1] Bank of Spain.

[2] f_{it}=100·(Commission Fees)/(Loans+Deposits), Banks' Annual Reports

[3] National Statistics Institute of Spain

[4] Interbank market for non-transferable deposits, 1 year, Bank of Spain

[5] Banks' Annual Reports

Table 2. Demand parameter estimates

Variable	Parameter	Robust-t	Parameter	Robust-t
Const	-0.9811	-7.3439	-0.9234	-1.7040
p_{it}	-0.0186	-1.5043	-0.2723	-2.2769
$lnS_{i\|g,t}$	0.7595	23.1923	0.4317	3.5582
br_{it}	0.2470	6.0099	0.6949	3.7679
	Method: OLS		Method: IV	
	Dep. Var.: lnS_{it}-lnS_{Oit}		Dep. Var.: lnS_{it}-lnS_{Oit}	
			Sargan Test: 1.64 (1)	
			Instruments: Const, i_t, $lnS_{i\|g,t-1}$, p_{it-1}, $mcost_{it}$	

Table 3. Estimated own and cross demand elasticities

Year	Price elasticities				
	Own price	Cross price across segments		Cross price same segment	
		Savings Banks	Private Banks	Savings Banks	Private Banks
1993	-3.470	0.037	0.052	0.157	0.160
1994	-2.292	0.029	0.033	0.118	0.100
1995	-2.619	0.032	0.037	0.129	0.116
1996	-2.778	0.033	0.045	0.130	0.136
1997	-2.552	0.030	0.042	0.116	0.126
1998	-2.365	0.031	0.039	0.115	0.117

Table 4. Parameter estimates of the pricing relationship

	Parameters	Estimates	Est./s.e.
Marginal cost			
Constant	α_0	1.104***	15.768
Branches	α_1	-0.076***	-5.352
Private Bank	α_2	-0.099***	-4.433
Interbank real rate	α_3	0.446***	37.313
Dynamic conduct determinants			
Current demand shock/Expected demand	φ_1	-8.261*	-1.286
Expected cost shocks	φ_2	-0.414***	-4.520
Private Bank	φ_3	0.264**	1.688
Random variable variances			
$\sigma=\sigma_u+\sigma_v$		0.088***	4.610
$\lambda=\sigma_u/\sigma_v$		4.033**	1.889

Mean Log-likelihood: 0.957562
Number of observations: 140

*** (**) (*) indicates that the parameter is statistically significant at 1% (5%) (10%)

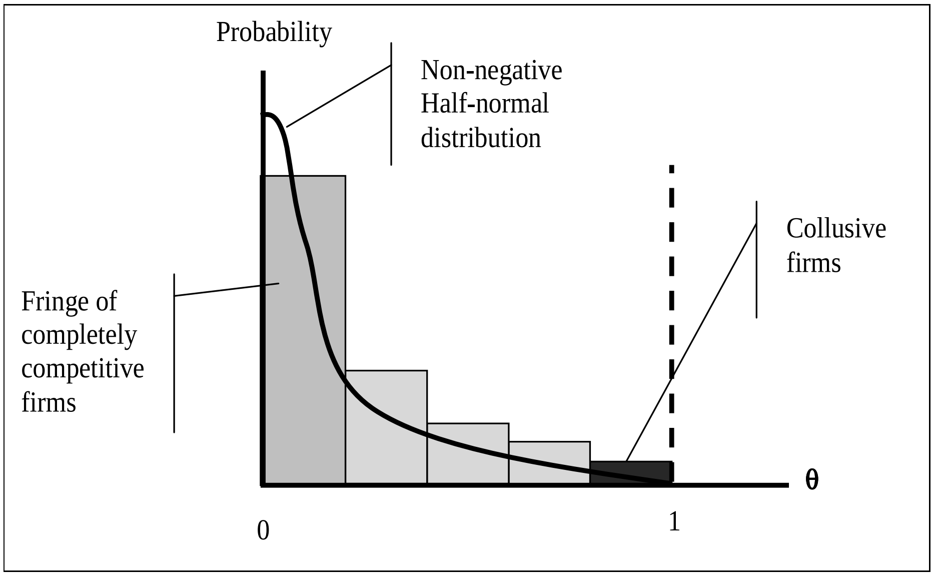

Figure 1. Fringe of *completely* competitive firms

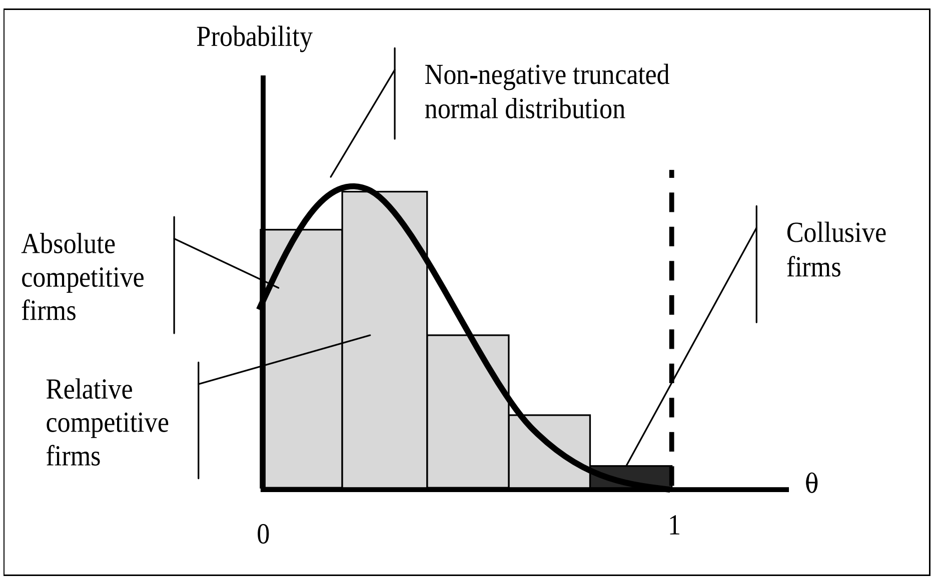

Figure 2. Fringe of *relative* competitive firms: umbrella pricing

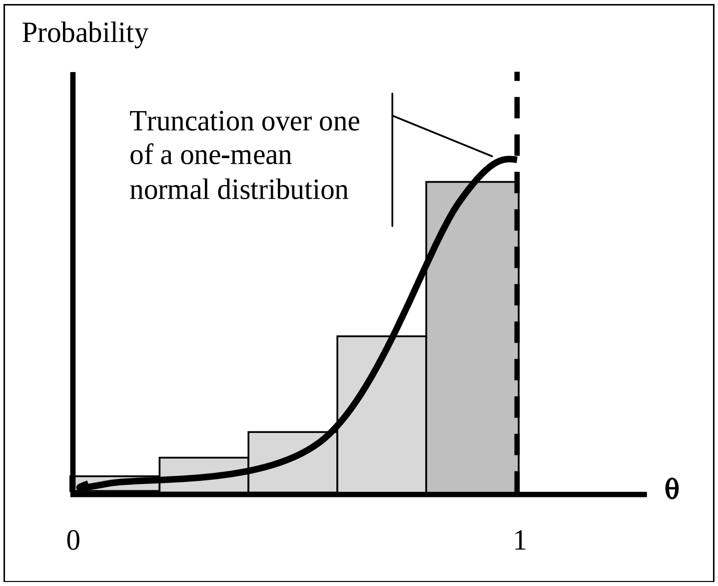

Figure 3. Collusion with *cheating* firms

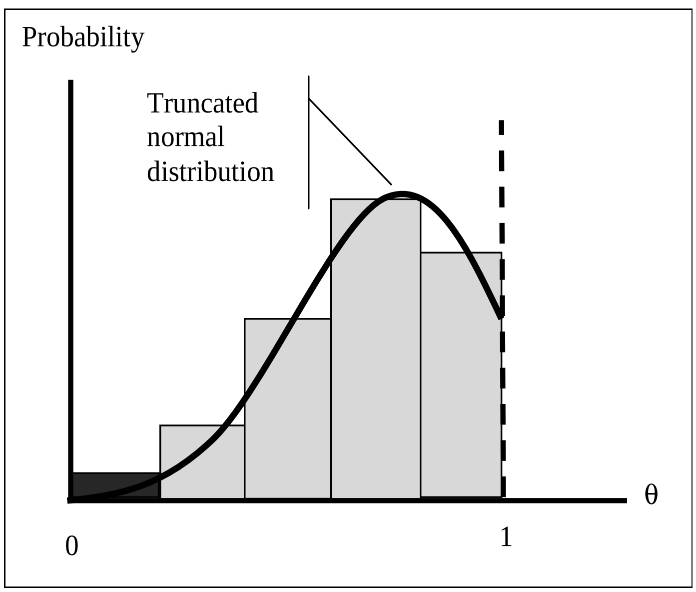

Figure 4. Collusion with a large fringe of *cheating* firms

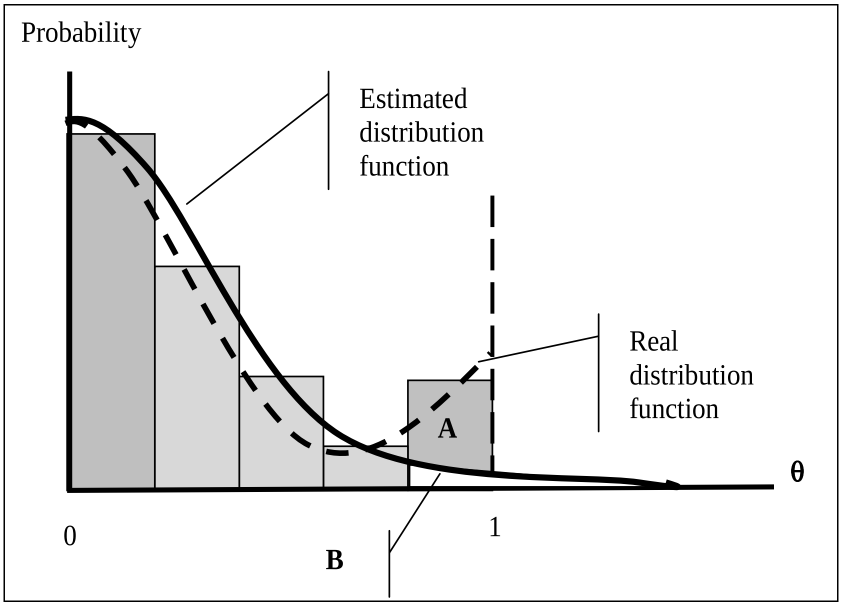

Figure 5. The two modes case and estimated distribution

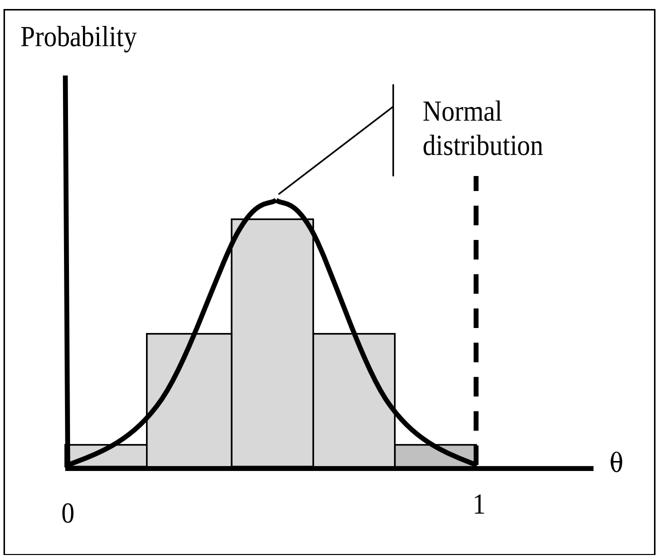

Figure 6. The symmetric case

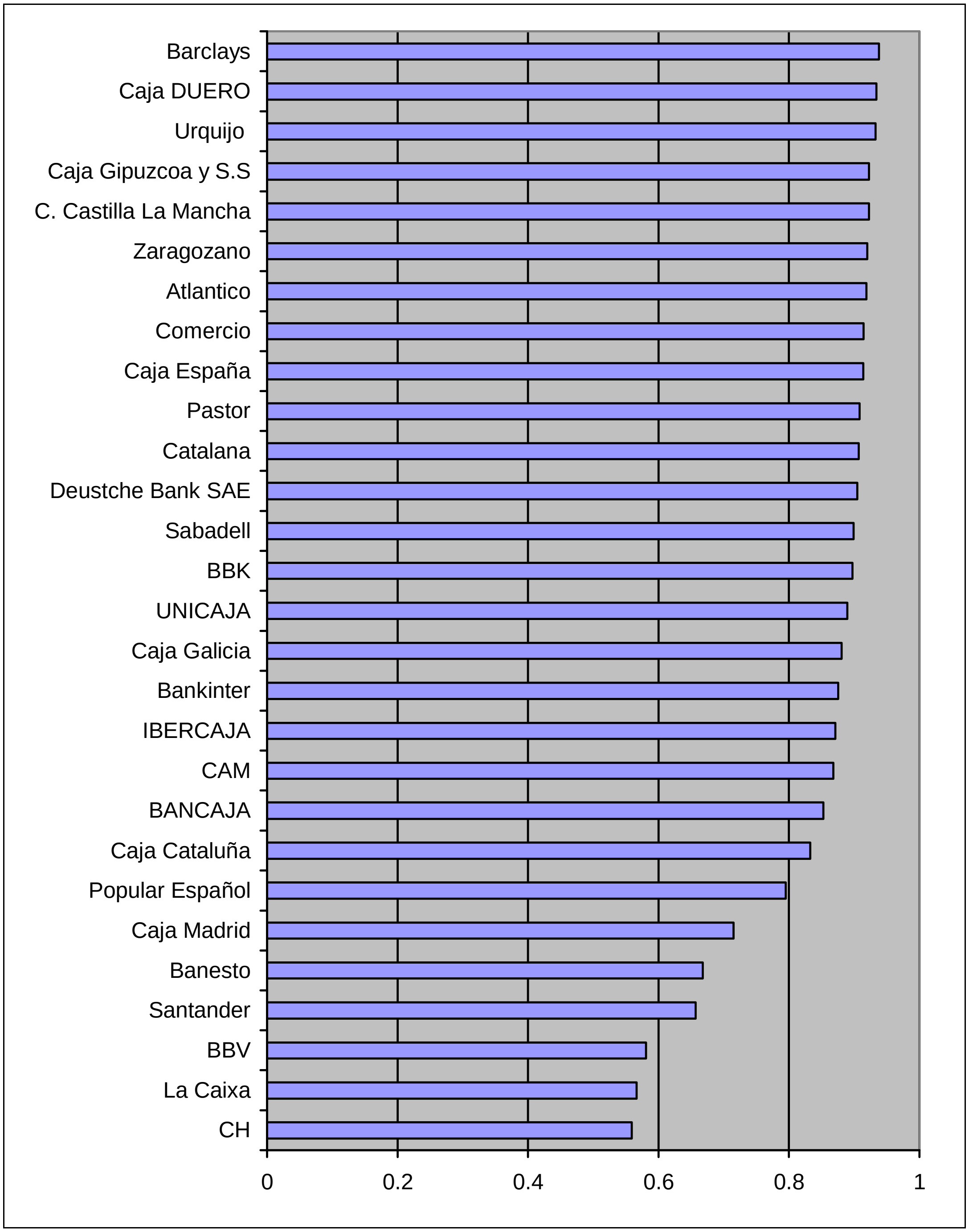

Figure 7. Banks' competitiveness level (Year: 1996)

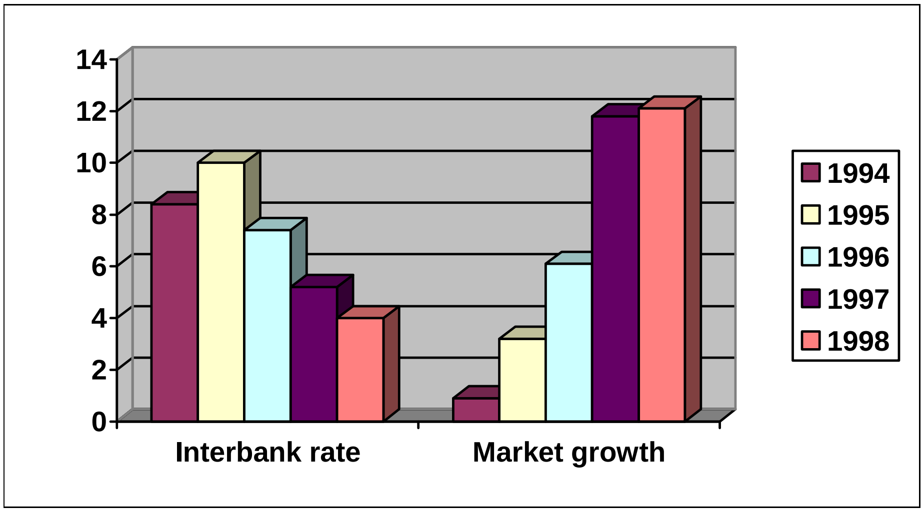

Figure 8. Interbank rate and Market growth (%)

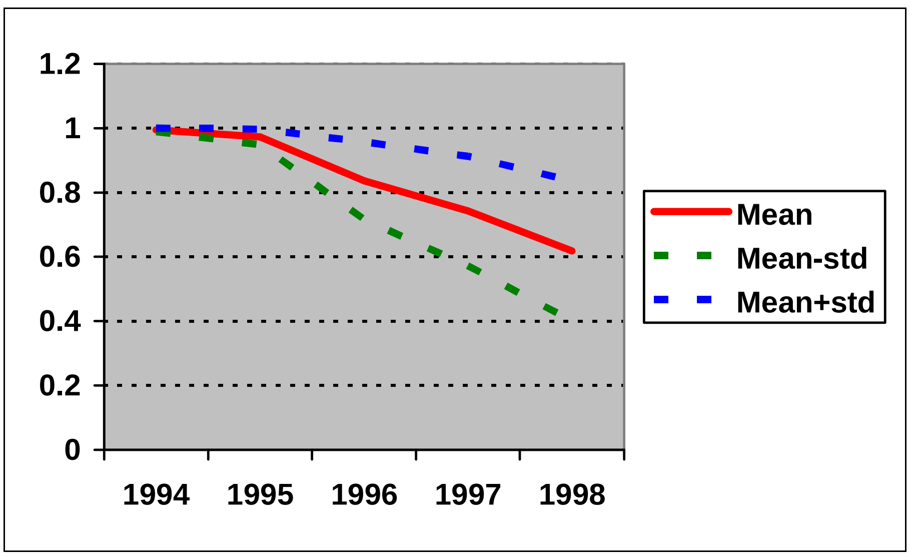

Figure 9. Evolution of the average competitiveness level