Quantile Regression in Lower Bound Estimation[1]

by
Maria Letizia Giorgetti
Sant'Anna School of Advanced Studies, Pisa Italy[2]
EPRIS Program, University of Siena, Italy

Contents:

Discussion paper
EI/29
September 2001

The Toyota Centre
Suntory and Toyota International Centres for
Economics and Related Disciplines
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
Tel: 020 7955 7719

[1] I am very grateful to John Sutton for comments and suggestions. I want to thank all the members of the Economics of Industry Programme, STICERD, LSE for helpful discussions. I am also grateful to Carlo Casarosa, Fabrizio Bulckaen, Giovanni Dosi, Alberto Pench and Alessandro Sembenelli who gave me helpful comments on my PhD dissertation, which was a preliminary step of this work. I want to thank Roberto Monducci, Director of SSI, ISTAT (National Institute of Statistics, Italy) who allowed me to perform these analyses using the SCT95 dataset. The support of the EPRIS Program is acknowledged.

[2] This work has been realized while the author was a visitor to the Economics of Industry Programme, STICERD, London School of Economics and Political Science.

Abstract

In this paper, I illustrate the additional information that can be provided in estimating the lower bound (Sutton 1991, 1998) by using quantile regression. Quantile regression allows us to investigate the influence of outliers. Previous lower bound estimates have been performed using the simplex method. In this paper, the lower bound estimates are obtained using both methods for sectors belonging to a “control group” and sectors belonging to an “experimental group” for Italian manufacturing sectors in 1995. The data employed are drawn from the ISTAT (National Institute of Statistics, Italy) dataset. The results suggest that Sutton’s predictions are robust.

Keywords: Lower bound, quantile regression, simplex

JEL Classification: C13, L11

1 Introduction

One of the most important theoretical predictions of Sutton's analysis (Sutton, 1991, 1998), is that there exists a lower bound to concentration, as market size goes to infinity, for sectors with high level of investments in R&D and advertising (endogenous sunk costs sectors), referred to as the 'experimental group' in what follows, while for sectors characterized by low investments in R&D and advertising (exogenous sunk costs sectors), i.e. the control group, the concentration level approaches zero as market size goes to infinity. This theoretical finding is developed inside the game theory framework, with a multistage game. In the first stages, according to the viability condition, firms decide to enter on the basis of irreversible investments they can sustain (the level of outlays in R&D and advertising, the number of products to introduce for each firm), on the basis of the viability condition, in the final stage some form of competition occurs.[1] The final stage allows us to define the profit function for each firm and to draw inferences about the market structure configuration. The lower bound approach allows us to recover the regularities tradition of the S-C-P paradigm, to introduce a two-way relationship between concentration and profitability, and to go beyond the standard paradigm which identified a single equilibrium and to deal with multiple equilibria[2]. The bigger theoretical importance of the lower bound approach is to fuse different important contributions: the S-C-P paradigm and the game theoretical framework, to set the basis for a new industrial organization.

One method employed in the previous literature following Smith (1985,1994) and used by Sutton (1991) and Robinson and Chiang (1996)[3], is a two-step procedure where the error distribution is a two or three parameters Weibull: the first step employs the simplex method to estimate the model parameters and the second step employs the pseudo maximum likelihood[4] to estimate the Weibull parameters. One interesting alternative approach, which has not as yet been used in this literature, is quantile regression. This method has two advantages; first, it is relatively straightfoward to implement, and second, it avoids the extreme sensitivity to outliers which arises in the simplex method.

By observing the scatter plots of both groups (see Figure 1), one can easily identify the presence of a couple of apparent outliers in the experimental group which could influence the lower bound estimations. For this reason the paper

[1] The viability condition tells us that firms decide to enter a market if and only if the profit level is higher or equal to zero. As regards the viability condition see chapter three Sutton, 1998.

[2] As regards the standard paradigm see chapter one, Sutton 2000. The bound's approach defines a set of possible equilibria, where the viability and stability conditions hold.

[3] It is important to remark that Davies and Lyons (1996) also did the lower bound estimation for all European manufacturing sectors for different types of sectors,(for more details, see footnote 8). In this paper, the cut-off level, to discriminate low and high investments, was fixed at 1% of R&D/sales and Advertising/sales. However the goal of this estimation is also to investigate the level of integration in the European Union.

[4] Smith (1985) has studied asymptotic properties for a class of non regular models in which the range of distribution depends on unknown parameters. In this paper the Weibull's parameters, α and s are unknown.

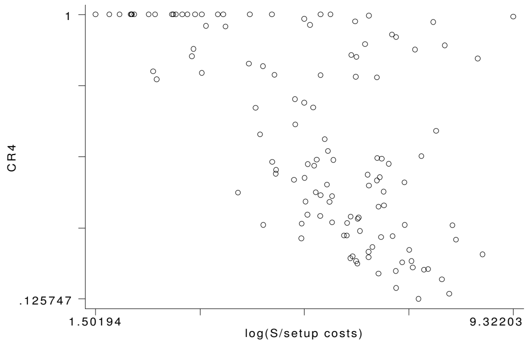
is organized as follows: first, we estimate the lower bound with the simplex method by considering all the observations in both groups, second we use the quantile regression to detect "apparent outliers", third we try to run the lower bound estimation, without detected outliers, with the simplex method.

I believe that the quantile regression method could help us to obtain better estimates of the lower bound, as it allows us to investigate the impact of outliers in the estimation.

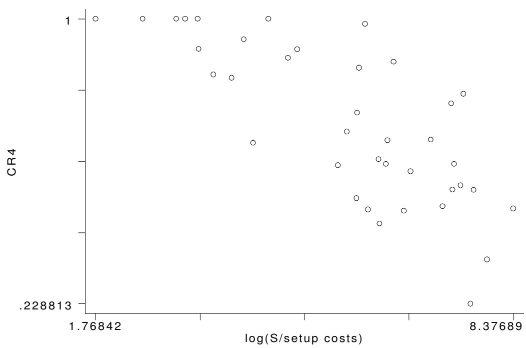
The analysis is performed on Italian manufacturing sectors in 1995 by employing one ISTAT (Italian National Institute of Statistics) dataset.[5] The data are collected by the SCI'95 survey, one ISTAT survey on budget data for firms with more than twenty employees; the industry's level classification is a high level: five digits, ATECO '91.[6] The ATECO'91 classification is easily comparable to SIC classification, employed in United States.

[5] The elaborations on raw SCI data and the Concentration ratio index have been completed inside the ISTAT, SSI Department (Firms Structural Statistics) with the financial support of an ISTAT scholarship in the period 1998/1999.

[6] The ATECO'91 is the classification used by Italian National Institute of Statistics. One of the advantages of this classification is that it is very similar to the NACE.Rev classification, employed by Eurostat. Therefore the ATECO'91 classification allows us to do easy and objective comparisons with classifications employed in other countries.



Control group, CR4 to log(S/σ)



Experimental group, CR4 to log(S/σ)

Figure 1: Scatter diagrams of the 4-firm concentration ratio versus a measure of market size for the control group (upper panel) and the experimental group (lower panel).

2 Sutton's Bound Approach in Italian manufacturing sectors

Sutton's model is estimated for two groups of sectors in Italy: the control group, with low R&D and advertising expenses and the experimental group, with high R&D and advertising expenses. I want to show that, as the market size grows, the concentration is higher for the experimental group than for the control group. As market size increases, firms belonging to the experimental group increase their expenses in R&D and advertising in order to capture a larger market share. The presence of higher profit opportunities induces an "escalation process" (Sutton, 1998), which prevents the concentration level from falling too much and, in particular, from approaching zero. This doesn't happen for the sectors in the control group.

The first step is to identify sectors belonging to the two groups. The indicator employed to discriminate between the two groups is: RA= (R&D + Advertising) /sales. This indicator has been calculated by summing up the expenditure on R&D and advertising of all the firms in each industry and by dividing the sum for the total sales of all the firms in the industry.

The cut-off to define the control group is equal to 1% and the cut-off to define the experimental group is equal to 4%: that is if RA $\leq$ 1%, then the sector belongs to control group and if RA $\geq$ 4%, then the sector is included in the experimental group. Two different cut-off points have been fixed to avoid problems posed by measurement errors in RA[7], so that all the sectors in the intermediate range are eliminated from the analysis. This process yields a control group, which includes 124 sectors, and an experimental group which includes 38 sectors[8].

The model estimated is the Sutton model (1991,1998):

$$log(\frac{CR4}{1-CR4}) = a + b\frac{1}{\log(S/\sigma)} + \varepsilon \tag{1}$$

where CR4 is the concentration ratio for the first four biggest firms in each market (sector)[9], $\log(\frac{CR4}{1-CR4})$[10]is a logit transformation of CR4, and S is the market size, measured by summing up the sales of all the firms in each sector. The measure of market size is normalized by dividing by σ, where σ is a measure of setup costs, that is the "minimal level of sunk costs that must be incurred by each entrant to the industry prior to commencing production"[11] and, finally, ε is the error term.

As to "market definition", the five digits ATECO'91 classification has been adopted[12]. For setup costs, I have followed Sutton by setting: $\sigma = \mu K$, where

[7]Previous work on lower bound estimation includes: Davies and Lyons (1996) for Europe, Robinson and Chiang(1996) for North America, Giorgetti (1999) for Italy. All these papers employ a single cut-off point equal to 1% for two different indicators:

a) IA= Advertising expenses/sales.

b) IRD=R&D expenses/sales.

If IA and IRD $\leq$ 1% the specific sector belonged to Type1, if IA >1% and IRD $\leq$1% the sector belonged to Type2A, if IA $\leq$ 1% and IRD >1% then that sector is included in Type2R and finally if both IA and IRD >1%, the sector is included in Type2AR.

[8]The initial total number of sectors is equal to 361 five-digits sectors. It is important to be precise, once the five digits segmentation level was not present, we choose the level immediately less segmented, i.e. four digits and, if not present, three digits.

[9]The concentration ratio is the sum of the market shares owned by the first four firms in each sector, ranked on the basis of sales. The previous works on lower bound estimations employ the CR4 (Sutton 1991), the CR3 (Robinson and Chiang, 1996) and the Herfindahl index, (Davies and Lyons, 1996 and Giorgetti, 1999). The Herfindahl index is obtained by summing up the squared market shares of all the firms in each sector.

[10]Because of the logit transformation, CR4's maximum value is equal to 0.99.

[11](Sutton, 1991, page 94).

[12] "In principle, one could try to split an industry into smaller and smaller subindustries, until a break in the chain of substitutes is found." Luigi Buzzacchi and Tommaso (1999).

K is the total fixed capital in each sector and μ is the ratio of median firm sales to total sales of the sector.

3 Model estimation

The model is estimated using two different approaches: a two step method (simplex and pseudo-maximum likelihood) and the quantile regression.

In the following sub-sections, the estimations obtained with both the simplex and the quantile regression approaches, are presented.

We first run the estimations by including all the observations.

The scatter plots (See Figure 1) seem to suggest the presence of a "couple of apparent outliers". When running the quantile regression the presence of these outliers is confirmed. I then run the simplex method excluding outliers identified through the quantile regression.

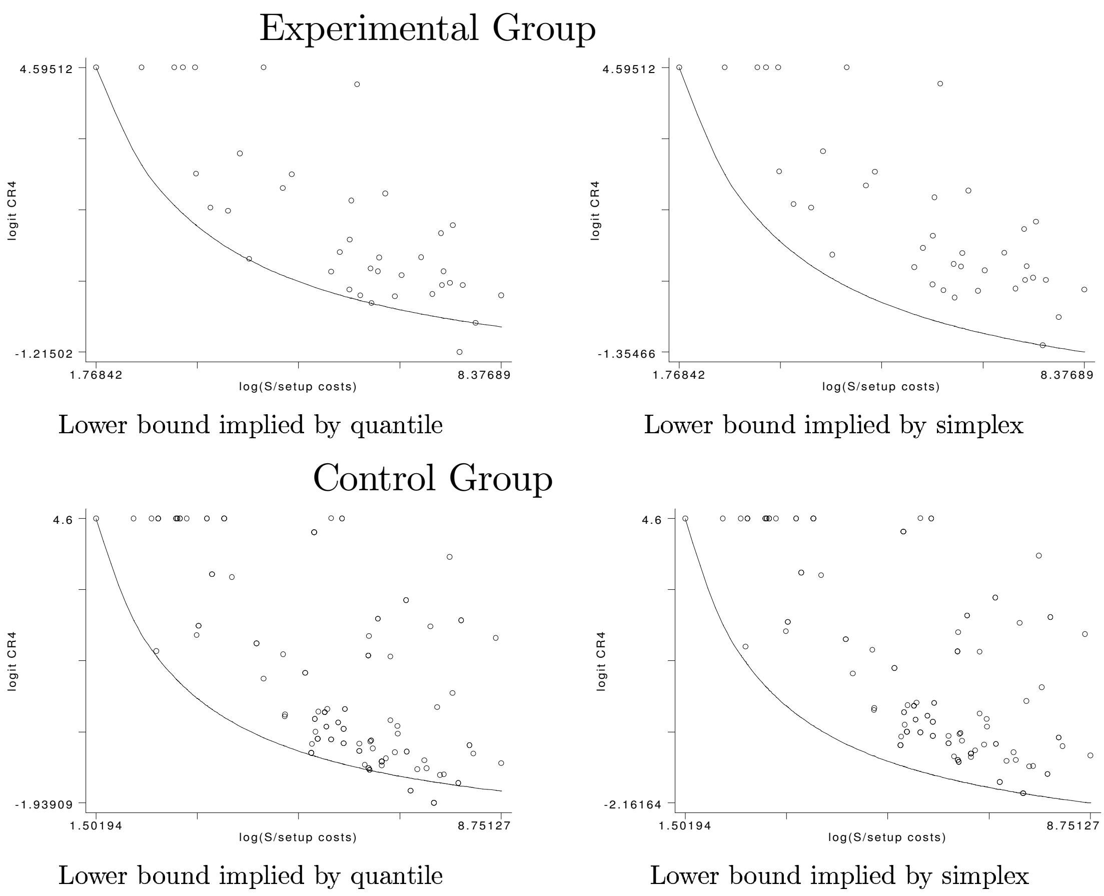


Figure 2:Lower bound estimation

3.1 The simplex approach

The first method follows the procedure employed by Sutton (1991) and Robinson and Chiang (1996); it is a two step procedure based on the papers by Smith (1985, 1994). The first step involves the use of the simplex method in order to estimate parameters a and b, and the second step estimates the Weibull parameters α and s. By solving the following linear programming problem, I obtain the parameters a and b parameters in equation (1):

$$min_{a,b} \sum_{i=1}^{n} \left[\log \frac{CR4_i}{(1 - CR4_i)} - \left(a + \frac{b}{\log \frac{S_i}{\sigma_i}} \right) \right] \tag{2}$$

s.t.

$$log \frac{CR4_i}{(1 - CR4_i)} \geq \left[a + \frac{b}{\log \frac{S_i}{\sigma_i}} \right]$$

I assume that ε is distributed as a two parameters Weibull ($\varepsilon \backsim$Weibull (α, s)) and the likelihood is $\Pi_{i=1}^{n} \ f \ (\varepsilon_i, \alpha, s)$.

By substituting the value of estimated errors obtained at the first step it can be shown that the maximum pseudo-likelihood[13] is:

$$\Pi_{i=1}^{n} f(\hat{\varepsilon}, \alpha, s) = \Pi_{i=1}^{n} f \left(\left[\log \frac{CR4_i}{(1 - CR4_i)} - \left(\hat{a} + \hat{b} \frac{1}{\log \frac{S_i}{\sigma_i}} \right) \right], \alpha, s \right) \tag{3}$$

The log pseudo maximum likelihood, knowing the Weibull density function, can be written as:

$$\sum_{i=n}^{n} \log \left[\frac{\alpha}{s} \hat{\varepsilon}_i^{(\alpha-1)} \exp \left(- \frac{\hat{\varepsilon}_i^{\alpha}}{s} \right) \right] \tag{4}$$

Maximizing this formula with respect to α and s, one can find the estimators for the two parameters α and s. The procedure studied by Smith focusses in particular on the non regularity of the maximum likelihood inference when $\alpha \leq 2$. In this case, the Fisher information for a, b (corresponding to β, in Smith's paper) could show differences compared to the regular case: "for $1 < \alpha \leq 2$ a local maximum of the likelihood function exists but it does not have the usual asymptotic properties, the estimator $\hat{\beta}$ being consistent at rate $O(n^{1/\alpha})$ when $\alpha < 2$ and $O(n^{1/2} \log n)$ when $\alpha = 2$".[14]

The results obtained with this procedure are presented in Table 1[15].

The estimation of CR4 as market size goes to infinity is obtained from the parameter a. From equation 1, it is evident that $(\log \frac{CR4}{1-CR4}) = a$ as S goes to

[13] Smith (1985) calls this procedure pseudo maximum likelihood, only the residuals strictly greater than zero are included.

[14] Smith 1994, pag.471.

[15] The estimations in this paragraph have been obtained using the Gauss software.

infinity. It is therefore possible to find out the asymptotic value of CR4 by calculating the antilog of a. This calculation is performed for both procedures.

The estimates for α supply values lower than 2, thus justifying the adoption of Smith's procedure. As market size goes to infinity, the estimates indicate a concentration level of 2.8% for the control group, and a concentration level of 5.0%, for the experimental group.

The results about a confirm Sutton's predictions about the lower bound, though the difference between the groups is not very large.

The lack of a big difference between the groups may depend on the presence of a "couple of apparent outliers" in the experimental group (see Figure 1, lower panel) as they could influence and underestimate the asymptotic value of concentration ratio. This motivates the use of quantile regression in the next section.

The standard errors of α and s are obtained using White's correction, in order to deal with the presence of heteroscedasticity[16], while the two-tailed (95%) confidence intervals for a and b are calculated by employing the tabulated values from the paper by Smith (1994, page 178).

As regards the parameter b[17], previous works in this literature predict a higher slope for the control group than for the experimental group. This is because in the experimental group, as market size increases, firms are pushed to increase their level of expenses in R&D and advertising; as a consequence the concentration level doesn't decrease as quickly as in the control group. "Thus, the existence of a shallower slope in Type 2 industries (i.e. experimental group) would constitute evidence that these expenditures are active competitive weapons subject to escalation, and not a simple fixed cost of competing"[18]. On the other hand, in the control group, the concentration level declines very fast because, as market size increases, firms don't have an incentive to outspend rivals and sustain only the minimal level of sunk cost in order to enter the market.

Table 1, Estimates from simplex approach

control group	a	b	α	s
estimations	-3.56	12.26	1.35	3.17
	(-3.57)	(12.27)	(0.03)	(0.17)
experimental group	a	b	α	s
estimations	-2.94	13.23	1.42	3.20
	(-2.96)	(13.03)	(0.05)	(0.03)

figures in parentheses indicate (one-sided) CI for a and b and s.e for α, s

3.2 Quantile regression

The other approach to estimate the lower bound is quantile regression. The interest in this method has grown because of its robustness. The quantile regression is a general case of the minimum absolute deviations method which

[16] See Greene, W.H. 2000, page 463.
[17] The theory's predictions are relative to the a parameter.
[18] Davies and Lyons (1996), page 97.

appeared in the literature before the least square estimation but which, because of difficulties in calculations, had been set aside for a long time. "Many illustrious figures (Gauss, Laplace and Legendre, to name only three) suggested that the minimization of absolute deviations might be preferable to least squares when some sample observations are of dubious reliability".[19] The outliers influence in lower bound estimation could lead to mistaken conclusions about the theory. Their presence in the experimental group could, in fact, shift down the lower bound by reducing the difference in the concentration ratio between the two groups, as market size goes to infinity. This is a good reason to investigate alternative and more robust methods. The most important works in this field are those by Koenker R. and Bassett JR. (1978, 1982). The central idea in quantile regression is to minimize the absolute residuals sum by giving different weight according to the quantile investigated. The quantile regression is, therefore, solved by turning the LAD, the least absolute deviation, into a linear programming problem. The θth regression quantile, $0 < \theta < 1$, is defined as any solution to the minimization problem:

$$min_{b \epsilon R} \left[\sum_{t \epsilon \{t: y_t \geq x_t b\}} \theta \left| y_t - x_t b \right| + \sum_{t \epsilon \{t: y_t < x_t b\}} (1-\theta) \left| y_t - x_t b \right| \right] \tag{5}$$

where $|y_t - x_t b|$ are the absolute residuals, y_t,$\{y_t : t = 1, ..T\}$ is a random sample on a random variable Y having distribution function F, and θ is the chosen quantile.

If one is interested in analyzing the median, θ will take the value 0.5 and the weight for negative and positive residuals will be the same. If one is interested in finding out what happens in the first percentile of distribution, θ will be equal to 0.01 and the higher weight will be assigned to negative residuals, i.e. the observations lying under the fitted quantile line. If one tries to employ the quantile regression to estimate the lower bound, surely one should be interested in the first percentiles or the first decile of the distribution. In fact in order to find robust estimates of the parameters a and b, in this case, the fifth percentile and the first decile of the distribution will be investigated.

The previous general formula in this specific case becomes:

$$min_{b \epsilon R} \left[\sum_{i \epsilon \left\{ i: \log \frac{CR4}{1-CR4}_t \geq a + \frac{b}{\log \frac{S}{\sigma}} \right\}} \theta \left| \log \frac{CR4_i}{1-CR4_i} - (a + \frac{b}{\log \frac{S_i}{\sigma_i}}) \right| \right.$$
$$\left. + \sum_{i \epsilon \left\{ i: \log \frac{CR4}{1-CR4}_t < a + \frac{b}{\log \frac{S}{\sigma}} \right\}} (1-\theta) \left| \log \frac{CR4_i}{1-CR4_i} - (a + \frac{b}{\log \frac{S_i}{\sigma_i}}) \right| \right] \tag{6}$$

[19] Koenker R and Bassett J.(1978)

Equation 6 clarifies that higher weight (because θ is equal to 0.05 or to 0.10) is on the observations lying under the fitted quantile line (the outliers); this means that, with the quantile regression, one minimizes the absolute residuals by taking into greater account the higher influence of the outliers.

The confidence intervals of the parameters are obtained with the boostrap method by randomly resampling the data.

Table 2 presents the quantile regression results for the lower bound[20]: the concentration ratio, as market size goes to infinity, is around 5% in the control group and around 11% in the experimental group[21], i.e. relative to the results reported in the previous section, the lower bound for the experimental group is higher, and the distance between the two groups is greater. The estimates obtained with the fifth or tenth percentiles are very similar, suggesting a robust fit for the lower bound. The fifth and tenth percentiles identify the "apparent outliers"[22].

The most striking feature of Table 2 is all the estimates[23] obtained with the quantile regression have very large confidence intervals, so that it is difficult to discriminate among results. The statistical advantage of quantile regression is insensitivity to outliers, the disadvantage is the lack of power relative to the maximum likelihood (i.e. simplex method).

---

[20] It is important to say that in order to have objective comparisons the quantile regression has also been run for the control group. The fifth percentile and first decile individuate 6 and 12 observations respectively, by shifting up the lower bound even if the observations detected by the quantile regression, in this group, do not constitute outliers, as well-known by the theory. The monotonic inverse relationship between concentration and market size, where concentration approaches zero as market size increases, is well-known in the Industrial Organization literature. The new important contribution of Sutton is to show how, for endogenous sunk cost sectors, the asymptotic concentration level reaches a high value, and this is the reason why the emphasis is put on the outliers in the experimental group and not in the control group.

[21] It is important to say that in order to have objective comparisons the quantile regression has also been run for the control group. The fifth percentile and first decile individuate 6 and 12 observations respectively, by shifting up the lower bound even if the observations detected by the quantile regression, in this group, do not constitute outliers, as well-known by the theory. The monotonic inverse relationship between concentration and market size, where concentration approaches zero as market size increases, is well-known in the Industrial Organization literature. The new important contribution of Sutton is to show how, for endogenous sunk cost sectors, the asymptotic concentration level reaches a high value, and this is the reason why the emphasis is put on the outliers in the experimental group and not in the control group.

[22] As regards the parameter b, the estimations with the quantile regression do not confirm the results observed in other works (Chiang and Robinson, Davies et al.).

[23] Estimations in this paragraph was carried out using STATA.

Table 2, Estimates using Quantile Regression

control group	a	b	**Asymptotic CR4**
quantile 5%	-2.96	11.35	4.9%
Confidence Interval 95%	(-3.62;-2.30)	(7.61;15.08)	
quantile 10%	-2.90	11.35	5.2%
Confidence Interval 95%	(-3.22;-2.58)	(9.42;13.28)	
experimental group	a	b	**Asymptotic CR4**
quantile 5%	-2.11	11.86	10.8%
Confidence Interval 95%	(-3.25;-0.98)	(5.94;17.7)	
quantile 10%	-2.08	11.81	11.1%
Confidence Interval 95%	(-3.96;-0.21)	(1.23;22.4)	

figures in parenthesis indicate CI for a and b

3.3 The sensitivity of the simplex method

In the light of the apparent identification of two 'outliers' under the quantile regression, it is of interest to examine what would happen under the simplex method if these two points were eliminated[24]. A re-estimation with these points deleted indicates that the lower bound for the experimental group now rises to 10.8%, which corresponds closely to the estimate obtained under quantile regression (11%). Table 3 shows all the estimates obtained with both procedures, for the simplex method with and without outliers (for the experimental group), in order to allow easy comparisons. The confidence intervals, for the simplex procedure, have been calculated following the table in the paper by Smith (1994)[25]: they are now very small and allow us to be more confident about the parameters estimations

Table 3 all the model estimates

The control group	a **parameter**	b **parameter**	**As.CR4%**
simplex	$\frac{\mathbf{-3.56}}{(-3.57)}$	$\frac{\mathbf{12.26}}{(12.27)}$	**2.8**
quantile 5%	$\frac{-2.96}{(-3.62:-2.30)}$	$\frac{11.35}{(7.61:15.08)}$	4.9
quantile 10%	$\frac{-2.90}{(-3.22:-2.58)}$	$\frac{11.35}{(9.42:13.28)}$	5.2
The experimental group	a **parameter**	b **parameter**	**As.CR4%**
simplex without outliers	$\frac{\mathbf{-2.11}}{(-2.08)}$	$\frac{\mathbf{11.86}}{(11.88)}$	**10.8**
simplex with outliers	$\frac{-2.94}{(-2.96)}$	$\frac{13.23}{(13.03)}$	5.0
quantile 5%	$\frac{-2.11}{(-3.25:-0.98)}$	$\frac{11.86}{(5.94:17.7)}$	10.8
quantile 10%	$\frac{-2.08}{(-3.96:-0.21)}$	$\frac{11.81}{(1.23:22.4)}$	11.1

Figures in parentheses indicate CI (one-sided) for simplex and for quantile regression

[24] The estimation for the experimental group without outliers were performed with Gauss, by employing the simplex and the pseudo-maximum likelihood.

[25] In the experimental group without outliers the estimates for the two Weibull parameters are:

$\alpha = 1.56$ std.error$= 0.05$,

$s = 3.43$ std.error$=0.03$.

4 Conclusion

The analysis suggests that quantile regression is a robust method to detect the outliers.

In the analysis of the experimental group with 38 observations, the fifth and tenth of percentiles identify the couple of our initial "apparent outliers", in the light of the apparent identification of two outliers it has been of interest to examine what would happen under the simplex without these two points.

The estimates obtained with the simplex method excluding identified outliers confirm all Sutton's predictions: the distance of the asymptotic concentration ratio between the two groups is remarkably increased (the value is around 8%),the experimental group shows a higher asymptotic concentration level and a shallower slope of the lower bound.

The parameter b[26] now has a lower value in the experimental group compared to the control group. It supplies information about the lower bound slope and tells us that, for sectors included in the control group, the lower bound reaches a low concentration ratio very quickly while for sectors belonging to the experimental group, a low level of concentration is reached more slowly[27].

The comparisons of my lower bound estimation with the results obtained in North America and Europe must be interpreted with caution since the lower bound varies with the level of aggregation (Sutton 1998), and this varies somewhat between these studies[28].

In the present dataset, the effect of presence of two 'apparent outliers' is to induce a substantial shift in the asymptotic lower bound to concentration.

[26] As already mentioned, the theory supplies predictions only about the a parameter, in this paper the estimates for the b parameter are presented to allow comparisons with other works in this literature

[27] This depends on the fact, that in sectors with exogenous sunk costs a minimal amount of expenditures in R&D and advertising is required to enter the market and after that, firms do not sustain anymore outlays as the market size increases. On the other hand, in endogenous sunk cost sectors where the advertising and R&D outlays are strategic weapons, the increase in market size induces firms to spend further and that leads to a shallow slope for the bound to concentration.

[28] If I compare my results with the bound estimation in North America, it is possible to see that the asymptotic CR4 for the experimental group is around 11% for Italy while the result obtained by Robinson and Chiang(1996) in North America, for sectors with high investments in R&D and advertising and with the linear model, is equal to 15.2%. The American value is higher if one also considers that they calculated the CR3 while I have calculated the CR4. However for Italy, the level of concentration, found for experimental group, is very high in comparison to the average level of CR5 for Italian manufacturing sectors calculated by ISTAT in 1996, on the basis of ASIA (The Dynamic Register of the Active Firms), which is around 3%-4%. As a consequence a level of 10% is a very remarkable level as regards Italy and this suggests that Sutton's predictions are robust. Also the asymptotic concentration levels around 3%, found in exogenous sunk costs sectors are very similar for Italy and North America: the lower bound approaches zero as market size increases. As regards to the results obtained by Davies and Lyons (1996) for Europe, the comparison is not so immediate because they have employed the Herfindahl index for different types of industries, and they calculate the Equivalent Number as asymptotic concentration level. As regards the control group the equivalent number is equal to 5000, while for the experimental group it is 32.

References

[1] Buschinsky M.(1998),“Recent Advances in Quantile Regression Models, A practical Guideline for Empirical Research” *The Journal of Human Resources*, pages 88-126.

[2] Buzzacchi, L. and Valletti, T., (1999), ”Firm Size Distribution: Testing the ’Indipendent Submarket Model’ in Italina Motor Insurance Industry,” Economics of Industry Discussion Paper No.EI/24, STICERD, London School of Economics and Political Science.

[3] Davies S. and Lyons B. (1996), *Industrial Organization in the European Union: Structure, Strategy and the Competitive Mechanism,* Oxford University Press, Oxford.

[4] Efron B. (1982), “The Jackknife, the Bootstrap and Other resampling Plans”. Philadelphia: *Society for Industrial and Applied Mathematics.*

[5] Giorgetti M.L .(1999),“Concorrenza e Concentrazione: Il modello di Sutton,” *Working Paper n.54, Dipartimento di Scienze Economiche,* Università di Pisa, Italy.

[6] Giorgetti M.L. (2000), “La verifica empirica del modello di Sutton per i settori manifatturieri italiani”, in *L'Industria*, il Mulino, Bologna, Italia.

[7] Gould W.W. (1992), sg11.1: Quantile regression with boostrapped standard errors. *Stata Technical Bullettin* 9:19-21.Reprinted in Stata Technical Bullettin Reprints, vol.2,pp. 137-139.

[8] Greene W.H. (2000), *Econometric Analysis*, Prentice Hall, International. Inc.

[9] Koenker R. and Bassett G. JR (1978), Regression Quantiles, *Econometrica* 46:33-50.

[10] Koenker R. and Bassett G.JR.(1982), Robust tests for heteroscedasticity based on regression quantiles. *Econometrica* 50:43-61.

[11] Koenker R.(2000), “Galton, Edgeworth, Frisch, and prospectus for quantile regression in econometrics”, *Journal of Econometrics* 95, 347-374.

[12] Robinson W.T. and Chiang J. (1996),“Are Sutton’s Predictions Robust? Empirical Insights into Advertising, R&D, and Concentration”, *Journal of Industrial Economics*, VOLUME XLIV

[13] Rogers W.H.(1992), sg11: Quantile regression standard errors. *Stata Technical Bullettin* 9:16-19. Reprinted in Stata Technical Bullettin Reprints, vol.2,pp. 133-137.

[14] Rogers W.H.(1993) sg11.2: Calculation of quantile regression standard errors.*Stata Technical Bullettin 13*:18-19.Reprinted in Stata Technical Bullettin Reprints, vol.3,pp. 77-78.

[15] Philps L.(1998), *Applied Industrial Economics*, Cambridge, Cambridge University Press.

[16] Smith R.(1985), "Maximum likelihood estimation in a class of non regular cases", *Biometrika*,Volume 72, pp.67-90.

[17] Smith R.(1994), "Non Regular Regression", *Biometrika*, Volume 81,1, pp.173-83.

[18] Sutton J.(1991), *Sunk Costs and Market Structure*, MIT; Cambridge Press.

[19] Sutton J.(1997), "Game-theoretic Models of Market Structure", in D.Kreps and K.Wallis, *Advances in Economics and Econometrics:Theory and Applications*, (Proceedings of the World Congress of the Econometric Society, Tokyo,1995), Cambridge: Cambridge University Press, pp.66-86.

[20] Sutton J.(1998), *Technology and Market Structure*, MIT, Cambridge Press.

[21] Sutton J.(2000), *Marshall's Tendencies. What can economists know?*, MIT, Cambridge Press.