Does market structure trigger efficiency? Evidence for the USA before and after the financial crisis

Halkos, George and Polemis, Michael

Department of Economics, University of Thessaly, Department of Economics, University of Piraeus

9 February 2018

Does market structure trigger efficiency?
Evidence for the USA before and after the financial crisis

George E. Halkos[a*] and Michael L. Polemis[b,c]

[a] Laboratory of Operations Research,
Department of Economics, University of Thessaly,

[b] Department of Economics, University of Piraeus
[c] Hellenic Competition Commission, Athens, Greece

Abstract

This paper investigates the relationship between efficiency and market structure for a sample of industrial facilities dispersed among the USA states. In order to measure the relevant efficiency scores, we use a Data Development Analysis (DEA) allowing for the inclusion of desirable and undesirable (toxic chemical releases) outputs in the production function. In the next stage, we utilise the bootstrapped quantile regression methodology to uncover possible non-linear relationships between efficiency and competition at the mean and at various quantiles before and after the global financial crisis (2002 and 2012). In this way, we impose no functional form constraints on parameter values over the conditional distribution of the dependent variable (efficiency). At the same time, we estimate at which part of its conditional distribution function, the efficiency is located and draw substantial conclusions about the range of policy measures obtained. The empirical findings, indicate that the relationship between efficiency and market concentration did not remain unchanged in the aftermath of the economic crisis. The empirical results survived robustness checks under the inclusion of an alternative market concentration indicator (CR8).

Keywords: Market concentration; Industrial Toxic Releases; Efficiency, Financial crisis; Bootstrapped quantile regression.

JEL codes: Q52; L1; C14

1. Introduction

Industrial pollution affects a broad spectrum of the efficient use of natural and environmental resources to economic activity (Harrington et al, 2014; Bi and Khanna, 2012). Over the last five years researchers have tried to shed light on this relationship by focusing on the well-functioning of market forces (competition) in reducing the level of environmental degradation (see for example Simon and Prince, 2016; Polemis and Stengos, 2017). It is well documented from prior theoretical studies that increased competition in an industry may result in lower levels of production per facility not allowing pollution to grow (Farber and Martin, 1986). On the other hand, recent theoretical work claims that increased competition triggers the incentives of a firm to reduce costs in order to reduce its final prices and thus the pollution control activities (Shleifer, 2004; Mansur, 2007).

Despite the profound interest by policy makers and government officials on the possible spillovers between market structure usually proxied by the level of concentration (the inverse of competition) and environmental efficiency the existing literature is still in its infancy with controversial results (Polemis, 2017). These can be justified on the grounds that many academic researchers acknowledge that competition may have positive as well as negative effects on environmental pollution (Polemis and Stengos, 2017; Simon and Prince, 2016; Branco and Villas-Boas, 2015; Shleifer, 2004). More specifically, in a recent study, Chen et al, (2018), investigate the causal link between market structure and industry performance using a micro panel data set of USA manufacturing industries over the period 1958-2007. They argue that there is a non-monotonic relationship between market structure and total-factor productivity. Similarly, Polemis, (2017) employed a semiparametric fixed effects estimator to examine the impact of market structure on the level of toxic chemical releases to a sample of four-digit industrial sectors among the USA states. Contrary to the parametric results, the study uncovers an inverted "*U-shaped*" relationship between industrial output and toxic chemical releases when market concentration is present.

Our approach is one of the very few attempts at modeling and quantifying the relationship between efficiency and market structure using facility level data. For this reason, we formulate a number of research questions including *inter alia* the following: In what way market structure affect environmental efficiency? What is the underlying mechanism for competition to affect the levels of (desirable and undesirable) output in the production process? In what way does the level of efficiency is determined before and after the global financial crisis? Why does the level of efficiency is lower (higher) under the presence of (non)-competitive market conditions? Lastly, what policy implications could be drawn in order to boost an efficient abatement mechanism?

To address these concerns, we rely on quantile regressions (QR) methodology. It is crucial to mention that quantile regressions impose no function form constraints on parameter values over the conditional distribution of the dependent variable (Apergis and Christou, 2015). In other words, quantile regression is more robust to outliers than ordinary least squares regression, and is semiparametric as it avoids assumptions about the parametric distribution of the error process. In contrast to standard linear regression techniques, which summarize the average relationship between a set of regressors and the outcome variable based on the conditional mean function, this type of regression describes the relationship at different points over the conditional distribution of the deponent variable (i.e efficiency scores). At the same time, we estimate at which part of its conditional distribution function, the efficiency is located and draw substantial conclusions about the range of policy measures obtained.

We contribute to the literature in several ways. First, we build for the first time in the empirical literature a DEA framework in which the level of industrial toxic chemical releases constitute the undesirable output (local pollutant). Second, we try to analyse the possible correlations between the efficiency scores (drawn from the DEA model on each USA state) with the level of market structure. In this way, we argue that an industry needs to obtain a certain level of market concentration (competition) in order to achieve (technical) efficiency. Third, we make use of econometric techniques (quantile regressions) to determine possible non-linear interactions between the level of industrial efficiency and market concentration. In this way we are able to draw sharp inference on the adjustment mechanism underlying these non-linearities before and after the global financial crisis that hit the USA.

The rest of the paper is organized as follows. Section 2 describes the data. Section 3 outlines the DEA framework along with the quantile regression methodology, while Section 4 reports the empirical findings and provides the necessary robustness checks. Finally, Section 5 concludes the paper.

2. Data

The primary source for our data was drawn from the Toxics Release Inventory (TRI) which covers the period 1987-2015. The latter is a plant-level database that includes information about the industrial facility (e.g., name, state zip code, primary industry, etc), and releases of toxic chemicals to the air and water, as well as transfers to any kind of land disposal in the US territory. The reason for using industrial toxic chemical releases as a proxy for pollution and not the standard pollutants such as CO_2, NO_X or SO_2 is that the latter are only available at a state level. However, since the

Decision Making Units (DMU) used in the DEA are industrial facilities we decide to utilize local pollutants generated by each distinct industrial plant.

The structural variables such as market concentration, level of employment, value added that correspond to each 6-digit code were drawn from the National Bureau of Economic Research (NBER) and especially from Manufacturing Industry Database (CES). This database contains annual industry-level data from 1958-2011 on output, employment, payroll and other input costs, investment, capital stocks, and various industry-specific price indexes. Especially for the year 2012, and due to data restrictions concerning the level of market concentration as measured by certain indicators (i.e CR8, HHI), we used data directly from the US Census of Manufacturers. Similarly to Polemis and Stengos (2017) and in order to check for the robustness of our findings, we take two measures of market concentration: HHI is the Herfindahl-Hirschman index for the 50 largest firms in the industry, and CR8 which is the sum of the market shares of the eight largest firms of in the sector (concentration ratio). It is worth mentioning that our measures of market structure reveal the existence or the absence of effective competition in the industry since concentration is simply the inverse of competition (Polemis and Stengos, 2017; Simon and Prince, 2016).

Lastly, our sample consists of hundreds of observations. However, we excluded observations for facilities with missing values for toxic chemical releases. As stated before, we used the TRI database in order to incorporate the level of toxic chemical releases in our sample. In this way we merge chemical releases data with national industry concentration ratios drawn from the NBER database and the Census of Manufacturers (only for the year 2012). It is worth mentioning that the final date (2012), represents the last year for which data regarding the Census of Manufacturers were available at the time the research was conducted.

3. Proposed methodology

3.1 Efficiency measurement

Data Envelopment Analysis (hereafter DEA) method may be used for evaluating the efficiency of a decision making unit (DMU) in relation to other DMUs. DEA has been widely applied in evaluating relative efficiencies in many scientific applications. A problem demanding attention when using DEA stems when undesirable outputs are present. In this case as efficiency is achieved by minimizing inputs and maximizing outputs we may desire to retain same inputs with higher level of desirable and lower level of undesirable outputs. This implies that undesirable outputs require a special treatment in model specifications.

Fare et al. (1989) distinguished desirable and undesirable outputs and proposed a non-linear programming model in evaluating DMUs efficiencies when coping with both desirable and undesirable outputs. Various other efficiency measurements in the presence of undesirable outputs have been proposed. Likewise one approach to deal with this problem is shifting undesirable outputs into inputs and then usey DEA. Seiford and Zhu (2002) proposed radial measures presuming that efficiency may be achieved by increasing desirable and reduce undesirable outputs at the same time. In doing so a multiplication of undesirable outputs by -1 is proposed together with the application of a satisfactory translation vector to convert all negative undesirable outputs to positive. These two transformations of altering position and translation endow us with the same efficient frontiers (Scheel 2001) with the Seiford and Zhu technique to be valid when assuming variable returns to scale (VRS) and the two techniques to supply dissimilar inefficiency scores.

In a different way undesirable output may be considered as a desirable output. In a production function setup, Fare et al. (1989) considered desirable and undesirable outputs asymmetrically and computed environmental technology using distance functions non-parametrically. Simultaneously imposing strong and weak disposability environmental performance indicators were computed. But Cooper et al. (2007) indicate a disadvantage of radial models when coping with undesirable outputs as they ignore slacks in the efficiency measurement.

In our model specification let the input vector be denoted as $x \in \Re_{+}^{N}$ generating both undesirable $u \in \Re_{+}^{J}$ and desirable $\tau \in \Re_{+}^{M}$ outputs. For the environmental technology and following Shephard (1970) and Färe and Primont (1995) output sets are assumed closed and bounded with inputs freely disposable. In addition, environmental output set $P(x)$ is identified when outputs are weakly disposable, that is $(\tau,u) \in P(x)$ with $0 \leq \delta \leq 1$ and $(\delta\tau,\delta u) \in P(x)$. This weak disposability hypothesis indicates that reducing undesirable outputs is costly taking place with a comparable reduction in desirable outputs. An additional hypothesis refers to the null jointness postulation of desirable and undesirable outputs entailing that undesirable outputs are by-products when generating desirable outputs, namely $(\tau,u) \in P(x)$ and $u = 0$ implying that $\tau = 0$.

The DDF approach is used to compute DMUs' environmental efficiency levels utilizing a direction vector $\gamma = (\gamma_{\tau}, -\gamma_{u})$ to help decreasing undesirable and increasing desirable outputs. In this way, the environmental efficiency score for a DMU k' (for DMUs $k = 1,...,K$) can be attained from:

$$\begin{aligned} &D\left(x^{k'},\tau^{k'},u^{k'};\gamma_{\tau},\gamma_{u}\right) = \max \beta \\ &s.t.\left(\tau^{k'} + \beta\gamma_{\tau}, u^{k'} - \beta\gamma_{u}\right) \in P(x), \end{aligned} \tag{1}$$

where the intensity variables are not negative and with constant return to scale implying a more appropriate assumption when analysing environmental problems

(Picazo-Tadeo et al., 2012). Environmental efficiency is signified when $D\left(x^{k'},\tau^{k'},u^{k'};\gamma_\tau,\gamma_u\right)=0$ and environmental inefficiency when $D\left(x^{k'},\tau^{k'},u^{k'};\gamma_\tau,\gamma_u\right)>0$.

Specifically, we use the value of shipments (SHIP) as a proxy for desirable output, while one undesirable output accounting for the toxic chemical releases (REL) is incorporated in our analysis. The inputs in the production process are total real capital stock (CAP), as a proxy for capital, total employment (EMP), as a proxy for labor and finally cost of electricity and fuels (ENER) as a proxy for energy. Moreover, we assume that the three inputs affect the desirable output in a separable way since neither capital stock and employment nor energy cost of an industrial facility are linked with its production process. In contrast, the production of the desirable output generates the toxic chemical pollutants distorting the environmental conditions in a non-separable way (Halkos and Polemis, 2018).

3.2. Econometric framework

Quantile regression is used in modelling the calculated efficiencies. OLS relies on a strong simplification estimating the influence of the independent variables on the mean of the conditional distribution of the dependent variable. This is due to the fact that explanatory variables may not only verify the mean but may well affect other parameters of the conditional distribution of the dependent variable.

On the other hand, quantile regressions permit the consideration of the full conditional distribution of the dependent variable and it is less uncertain compared to the OLS (mean) regression allowing the estimated parameters to fluctuate at different points of the dependent variable’s conditional distribution. As a nonparametric method, quantile regression requires no functional form and is not susceptible to extreme values. This is due to minimization of residuals and not their squares as in

OLS. At the same time, considering various different quantile regressions allows us a more comprehensive depiction of the fundamental conditional distribution.

Quantile regressions are especially useful when dealing with non-identically distributed data (Distante et, al, 2018). In these situations one should expect to observe significant discrepancies in the estimated 'slopes' at different quantiles with respect a given set of covariates (Machado and Mata, 2000).

The estimated slopes of quantile regressions assess the change in a specific quantile of the dependent variable caused by a unitary change in the explanatory variable. This helps comparisons among quantiles of how much they are affected from explicit characteristics relative to other quantiles. In this way, quantile regressions are useful in the presence of heteroskedasticity and/or no normality in the disturbance term (Buchinsky 1998). Moreover, quantile regressions also provide a richer characterization of the data, allowing us to consider the impact of a covariate on the entire distribution of the dependent variable, not merely its conditional mean. It is worth mentioning that QR is invariant to monotonic transformations such as natural logarithms.

If we assume a random variable Y with a probability distribution function (Halkos, 2011)

$$F(y) = \Pr(Y \leq y) \tag{2}$$

such as $0 < \lambda < 1$ when the i quantile is defined as the lowest y assuring the condition $F(y) \geq \lambda$

$$Q(\lambda) = \inf\{y : F(y) \geq i\} \tag{3}$$

The quantile regression expands the simple model including the independent variables X that assume linear specification for the conditional quantile of the dependent variable Y given the values of the matrix P of the explanatory variables X. Particularly,

$$Q(\lambda \mid X_i \beta(\lambda)) = X'_j \beta(\lambda) \tag{4}$$

where
$$\beta(\lambda) = \arg mm_{\beta(\lambda)} \{\sum_j \rho_\lambda (Y_j - X'_j \beta(\lambda)) \tag{5}$$

For large samples, the coefficients of quantile regressions are normally distributed (Koenker, 2005).

The quantile density function $S(\lambda)$ may be calculated with the use of the Kernel density estimator (Powell 1986, Buchinsky 1995, Jones 1992). Specifically,

$$\hat{S}(\lambda) = \frac{1}{\left[(1/i)\sum_{j=1}^{i} c_i^{-1} L\left(\hat{\varepsilon}_j(\lambda) / c_i \right)\right]} \tag{6}$$

where $\hat{\varepsilon}_j$ the residuals of the quantile regression.

4. Results and discussion

4.1 Efficiency and econometric findings

Table 1 presents the results of the regional environmental efficiency estimates derived from our DEA model before (pre-crisis) and after (post-crisis) the financial crisis (2002 and 2012 respectively). Specifically, before the financial crisis, the efficiency results reveal that, 3 out of 50 states (Alaska, Hawaii and New Mexico) are reported to be environmentally efficient in terms of the anthropogenic emissions since their scores are close to unity. On the other hand, 3 out of 50 states report the lowest efficiency values ranging from 0.500 to 0.594. These are Nevada, Rhode Island and Arizona. The descriptive statistics reveal low disparities of regional environmental efficiencies among US states since the standard deviation and the coefficient of variation (CV) appear to be relatively low ranging from 0.080 to 0.118. Moreover, on average terms USA states have an efficiency level equal to 0.646. This means that US regions on average terms are able to reduce their toxic chemical releases (undesirable output) generated by the manufacturing sectors by 54.6% to reach the efficiency frontier.[1]

On the contrary when we stress our attention to the post-crisis model (see column 2), there are significant differences in the magnitude of the efficiency scores. More specifically, nearly all the US states portray a significant reduction in the efficiency scores with Arizona, New Hampshire and Wisconsin depicting the lowest values. On the contrary Alaska, Hawaii and New Mexico remain the "*leaders*" in the efficiency ranking but with a smaller magnitude. It is worth mentioning that the average efficiency score has been reduced during the last ten years of our analysis by 16,6% reaching the level of 0.564 (or 56,4%). The latter is close to the inefficiency score reported in the pre-crisis model, revealing that the financial crisis has also impacted a large global economy such as the USA.

In the next step we estimate separately the following (parametric) models by employing OLS and QR techniques for the two years of our study (2002 and 2012). Similarly to other empirical studies (Halkos, 2003; Millimet et al., 2003; Zarzoso and Morancho, 2004; Halkos and Tzeremes, 2009; Halkos and Polemis, 2017), we estimate the following simple (parametric) model:

$$EFF_{ij} = \alpha + b_1 LHHI_{ij} + b_2 LHHI_{ij}^2 + u_{ij} \tag{7}$$

i = 2002, 2012 Years and j = 1,2,…2,391 DMUs

where EFF denotes the efficiency scores per DMU at time i =2002, 2012 and $LHHI_{ij}$ is the logged Hirschman-Herfindahl index of the 50 largest firms in each of the six-digit industry sector included in the analysis. Finally u_{ij} is the i.i.d disturbance term.

[1] Since the mean environmental efficiency score equals to 0.649 or 64.9%%, the rest amount 0.546 (54.6%) denotes the inefficiency score.

Table 1: Results of the DEA efficiency scores per state

State	2002	2012
Alaska	1.000 (1)	1.000 (1)
Alabama	0.647 (68)	0.490 (68)
Arkansas	0.627 (47)	0.537 (47)
Arizona	0.594 (10)	0.418 (11)
California	0.698 (103)	0.604 (102)
Colorado	0.625 (11)	0.576 (12)
Connecticut	0.648 (40)	0.549 (40)
Delaware	0.685 (10)	0.511 (10)
Florida	0.602 (39)	0.480 (34)
Georgia	0.621 (79)	0.463 (78)
Hawaii	0.849 (3)	0.868 (3)
Iowa	0.707 (53)	0.565 (53)
Idaho	0.648 (10)	0.463 (9)
Illinois	0.679 (148)	0.660 (149)
Indiana	0.667 (84)	0.557 (87)
Kansas	0.763 (35)	0.592 (35)
Kentucky	0.653 (51)	0.517 (52)
Louisiana	0.754 (68)	0.637 (67)
Massachusetts	0.616 (42)	0.636 (42)
Maryland	0.644 (15)	0.636 (15)
Maine	0.643 (17)	0.396 (18)
Michigan	0.643 (93)	0.559 (96)
Minnesota	0.628 (53)	0.583 (53)
Missouri	0.662 (62)	0.565 (62)
Mississippi	0.629 (33)	0.494 (32)
Montana	0.727 (9)	0.623 (9)
North Carolina	0.700 (76)	0.516 (74)
North Dakota	0.718 (8)	0.629 (7)
Nebraska	0.736 (19)	0.613 (19)
New Hampshire	0.637 (10)	0.436 (10)
New Jersey	0.711 (45)	0.588 (45)
New Mexico	0.880 (4)	0.806 (4)
Nevada	0.500 (1)	0.281 (1)
New York	0.663 (60)	0.508 (61)
Ohio	0.647 (197)	0.537 (198)
Oklahoma	0.686 (27)	0.629 (28)
Oregon	0.616 (35)	0.494 (36)
Pennsylvania	0.623 (135)	0.538 (135)
Rhode Island	0.534 (11)	0.542 (11)
South Carolina	0.676 (60)	0.448 (60)
South Dakota	0.698 (4)	0.581 (4)
Tennessee	0.647 (60)	0.492 (59)
Texas	0.735 (196)	0.614 (197)
Utah	0.699 (14)	0.596 (14)
Virginia	0.647 (49)	0.553 (48)
Vermont	0.625 (7)	0.590 (6)
Washington	0.689 (34)	0.541 (34)
Wisconsin	0.621 (116)	0.470 (114)
West Virginia	0.723 (22)	0.533 (22)
Wyoming	0.720 (7)	0.716 (6)
Descriptives		
Mean	0.676	0.564
Standard deviation	0.080	0.114
Median	0.658	0.555
Max	1.000	1.000
Min	0.500	0.281
Coefficient of variation (CV)	0.118	0.203

Notes: The table reports the mean efficiency scores of all the sample DMUs by region (state) for the years 2002 and 2012. The efficiency scores were estimated with the DEA methodology using annual frontiers and constant returns to scale (CRS). The benchmark best practice frontier for DEA is efficiency equal to 1. The number in parentheses are the power plants (DMUs) in each of the USA state.

The degree of the polynomial has been determined by the maximum number of statistically significant powers. In our case, third and higher degree polynomial specifications have the extra powers of CR8 to be not statistically significant. Besides, we choose a quadratic specification since alternative specifications such as log-linear and log-log could admit only monotonic (and nonlinear) relationships (Dai et al, 2014). As it is evident the key variables of interest are b_1 and b_2. Therefore equation (27), implies that when the relationship between market structure and efficiency is nonmonotonic, the sign of b_1 should be different from b_2.

In Figures 1A and 2A of the Appendix, we plot the dependent variable (EFF) by quantiles for the years 2002 and 2012 respectively. From the careful inspection of Figure 1A, it is obvious that even at the 50th quantile (median) the states depict relatively moderate efficiency scores (around 0.55 or less). It is worth mentioning that the CDF of the dependent variable remains stable until the 30th quantile and equal to 0.5. On the contrary, there is an increasing upward trend when we cross the 60th quantile in which it is evident that efficiency scores exceed the level of 0.65 on average reaching the highest level of 1.00 from the 80th quantile and henceforth. The situation is different when we examining the CDF associated with the dependent variable (EFF) for the year after the financial crisis (2012).

Specifically, from Figure 2A, it is evident that the CDF does not portray significant variation since it follows an increasing upward trend from the low quantiles (especially from the 10th) until the high quantiles (80th quantile). The above findings reveal that neither of the two CDFs of the dependent variable follow a symmetric pattern.

Table 2: Estimation results for the year 2002

Variable	Model 1	Model 2													
	OLS	Quantiles τ													
		Q (0.01)	Q (0.05)	Q(0.10)	Q(0.15)	Q(0.20)	Q(0.25)	Q(0.30)	Q(0.40)	Q(0.50)	Q(0.60)	Q(0.70)	Q(0.80)	Q(0.90)$^{+}$	Q(0.95)$^{+}$
LHHI	-0.0722*** (-3.63)	-0.00031 (-1.68)	-0.00100 (-0.83)	-0.00808 (-1.56)	-0.0420 (-1.53)	-0.156*** (-4.36)	-0.165*** (-33.82)	-0.172*** (-14.83)	-0.191*** (-11.12)	-0.256*** (-7.08)	-0.252** (-2.72)	-0.211*** (-6.17)	0.144*** (3.63)	0.2026*** (11.08)	0.0929*** (33.57)
LHHI^2	0.00796*** (4.14)	0.00002 (1.64)	0.0000774 (0.80)	0.000639 (1.54)	0.00335 (1.52)	0.0125*** (4.37)	0.0132*** (26.68)	0.0141*** (11.98)	0.0161*** (9.91)	0.0235*** (6.07)	0.0260** (2.83)	0.0237*** (7.13)	-0.0164** (-2.58)	-0.0155*** (-9.45)	-0.0071*** (-28.84)
Constant	0.805*** (16.72)	0.50101*** (851.62)	0.503*** (135.81)	0.525*** (32.80)	0.631*** (7.43)	0.988*** (8.81)	1.012*** (93.25)	1.027*** (37.70)	1.082*** (25.59)	1.232*** (15.77)	1.194*** (5.70)	1.126*** (14.93)	0.0886 (0.59)	0.3460*** (6.90)	.7062144*** (93.10)
Observations	2,360	2,360	2,360	2,360	2,360	2,360	2,360	2,360	2,360	2,360	2,360	2,360	2,360	2,360	2,360
Turning point (logged values)	4.535	-	-	-	-	6.240	6.250	6.099	5.932	5.447	4.846	4.451	7.439	6.534	6.486
Turning point (measurement units)	3,429	-	-	-	-	1,738	1,778	1,257	854	280	702	283	678	246	1,738
Shape of curve	U-shape	-	-	-	-	U-shape	U-shape	U-shape	U-shape	U-shape	U-shape	U-shape	Inverted U-shape	Inverted U-shape	Inverted U-shape

Notes: The dependent variable in both models is the efficiency score (EFF) obtained by the DEA method. LHHI is the logged Hirschman-Herfindahl index of the fifty largest companies in the sector. LHHI^2 is the squared logged Hirschman-Herfindahl index of the fifty largest companies in the sector. Model 1 was estimated using OLS and allowing for robust standard errors. Model 2 was estimated using the Bootstrapped quantile regressions methodology at different quantiles τ (0.01, 0.05, 0.10 0.15. 0.20. 0.25. 0.30, 0.40, 0.50, 0.60, 0.70, 0.80) allowing for 100 repetitions. "-" means that the turning point cannot be estimated since the polynomial coefficients of the market concentration variable (HHI) are not statistically significant. $^{+}$ Denotes that the relevant estimates were drawn by using a quantile regression model. The numbers in parentheses denote the t-statistics, while the numbers in square brackets are the p-values. Significant at ***1%, **5% and *10% respectively.

Table 2 reports the parametric results of our analysis. Specifically, column 1 displays the regression findings using OLS and allowing for robust standard errors (Model 1), while columns 2–15 record the results at various quantiles obtained from the Bootstrapped instrumental quantile analysis allowing for 100 repetitions (Model 2). In the case of Model 1 it is evident that all the coefficients are highly statistically significant alternating their signs starting from negative to positive. This implies that efficiency decreases up to a certain "*turning*" point (3,429) and then increases gradually, suggesting a non-monotonic '*U*' shaped relationship between efficiency scores and market concentration.[2]

In contrast, the results of the quantile approach for Model II highlight a different pattern (reported in columns 2–15 of Table 2). In particular, although the coefficients of the market concentration indicator (HHI) alternate their signs across all quantiles of the efficiency scores distribution, they are shown to be statistically significant from the 20^{th} quantile, turning out to be insignificant at low quantiles (τ = 0.01, 0.05, 0.10 and 0.15) of the conditional distribution (see columns 2-5). In addition, from the 20^{th} until the 70^{th} quantile the findings confirm the baseline results of Model 1 since all the coefficients are statistically significant alternating their signs from negative to positive (see columns 6-12). This clearly suggests the existence of a quadratic (non-linear) polynomial HHI form consisting of one "*turning*" point ('*U*-shaped curve). In other words, regional market structure level decreases up to a certain point (estimated low) and then increases.

From the inspection of the relevant table, it is obvious that the turning points display significant variation around the median (50^{th} quantile) with the estimated low to be appeared on the 70^{th} quantile (283) and the estimated high on the first quartile (25^{th} quantile) reflecting a value of 1,778. However, the previous findings are fully

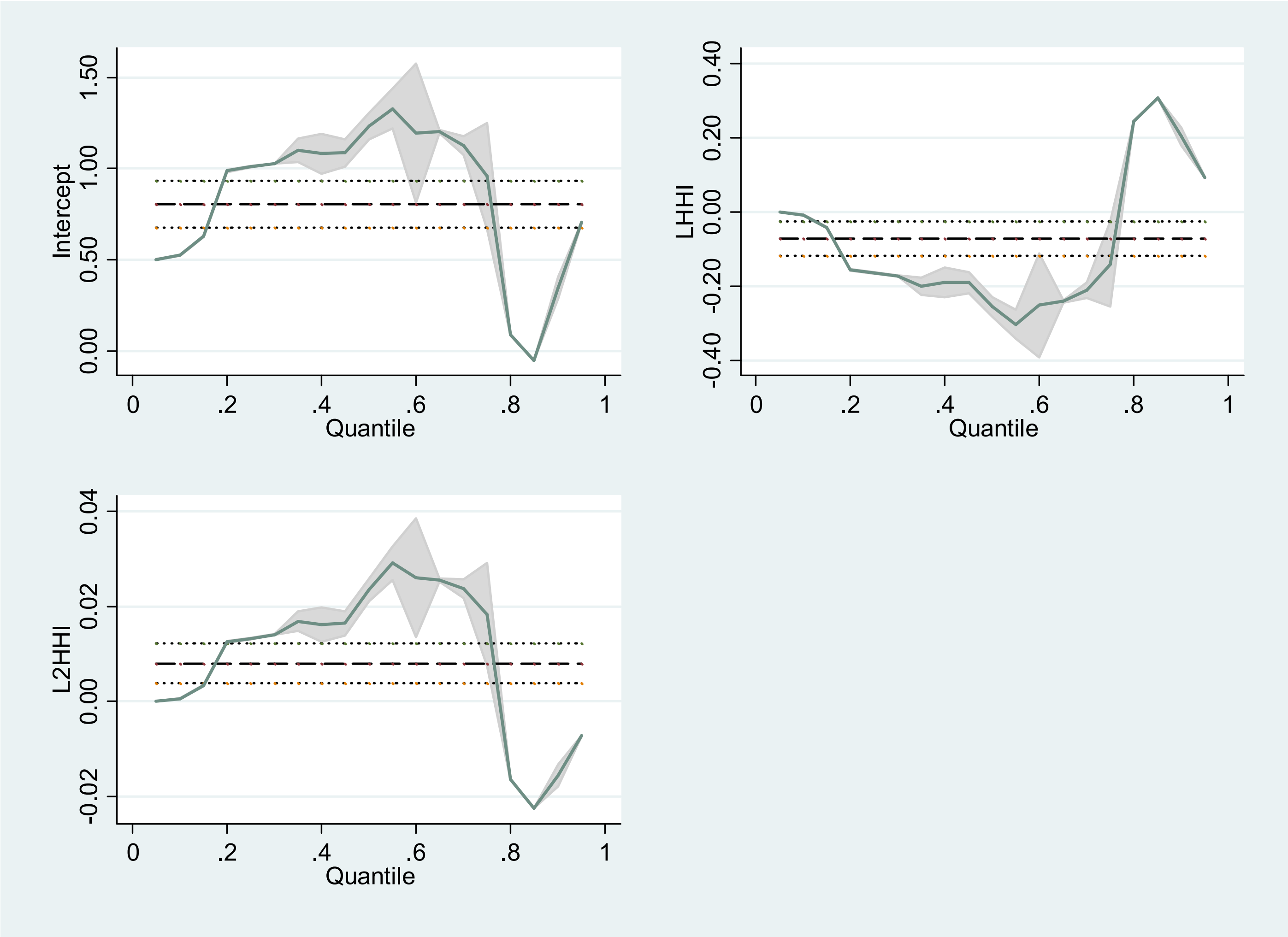
[2] Alternatively we trace an inverted "*U*' shape relationship between competition and efficiency.

reversed when we estimate the QR for higher quantiles (see columns 13-15). In other words, the relationship between efficiency and market structure remains non-monotonic but with a different curvature (inverted U-shape). These findings indicate that as we are moving across a certain point (i.e threshold level) in the cumulative distribution of the dependent variable (efficiency), which in this case this is evident at the 70th quantile, market structure seems to generate a reversal effect on the level of industrial efficiency. In other words, in higher quantiles (see for example 95th quantile), the competition effect is so strong that shifts the U-shaped curve to the right due to the larger estimated turning point of 1,738 units compared to the previous quantiles of the efficiency distribution (678 and 246).[3]

Having discussed the empirical findings of the two models, it is interesting to examine how each covariate's effects vary across quantiles, and contrast them with the (fixed) OLS estimates (Model 1). The following diagram illustrates how the market structure effect vary over quantiles. It also shows how the magnitude of the effects at various quantiles differ considerably from the OLS coefficient, even in terms of the confidence intervals around each coefficient. Regarding the first polynomial term (LHHI) it is evident that its estimated coefficient takes negative values nearly across all the quantiles (see right side of the diagram).

[3] Since competition is the inverse of market concentration (HHI), the relationship between efficiency and polynomial powers of competition is of the U-shaped form.

Figure 1: Variation of LHHI estimates across quantiles for the year 2002



Notes: The dependent variable in both models (OLS and QR) is the efficiency scores obtained by the DEA method (EFF). LHHI is the logged Hirschman-Herfindahl index of the fifty largest companies in the sector. L2HHI is the squared logged Hirschman-Herfindahl index of the fifty largest companies in the sector. The bold dotted horizontal line denotes the fixed OLS estimate. The dotted lines and the grey area denote the confidence bands for the OLS and QR estimates respectively.

However, from the 80th quantile, the estimates change their sign into positive. In contrast the squared polynomial term (L2HHI) follows the opposite direction taking firstly positive values and then negative ones (see lower left of the diagram). The above findings fully confirm the previous results indicating a different (inverted U-shaped) behavior of market structure across the higher quantiles of the efficiency distribution. Compared to the (fixed) OLS estimates, it is evident that for the market structure variable (HHI), the QR effects are much stronger at higher quantiles, with the OLS effect quite far from the median estimate (50th quantile). We also perform the test for the equivalence of the quantile estimates across quantiles, which allows us to estimate the model for each of several quantiles in a single model, accounting for cross-equation hypothesis tests. The p-values (0.000 and 0.000) clearly reject equality of the estimated coefficients for the three quartiles (0.25, 0.50 and 0.75) in each of the two polynomial terms (LHHI and $LHHI^2$).

In the next stage we proceed with the OLS and QR estimates drawn from the post-crisis model (see Table 3). In this case, some interesting results emerge. First, the OLS estimates are not statistically significant despite the fact that the polynomial terms alternate their signs from negative to positive (see column 1). As a consequence the turning point and subsequently the shape of the relationship between efficiency and market structure cannot be estimated.

Second, compared to the previous estimates of the pre-crisis model (see Table & Figure 1) the relationship between efficiency and market structure (competition) although non-monotonic depicts a completely different pattern. Specifically, in lower quantiles of the conditional distribution of the dependent variable (τ = 0.01, 0.05, 0.10, 0.15, 0.20, 0.25 and 0.30), the polynomial terms are statistically significant alternating their signs from positive to negative (see columns 2-8). This is consistent with the existence of an inverted U-shaped curve denoting that regional market structure level increases up to a certain point (estimated high) and then decreases. The magnitude of the turning points differ significantly across the lower quantiles ranging from 912 (1^{st} quantile) to 3,415 (30^{th} quantile).

Table 3: Estimation results for the year 2012

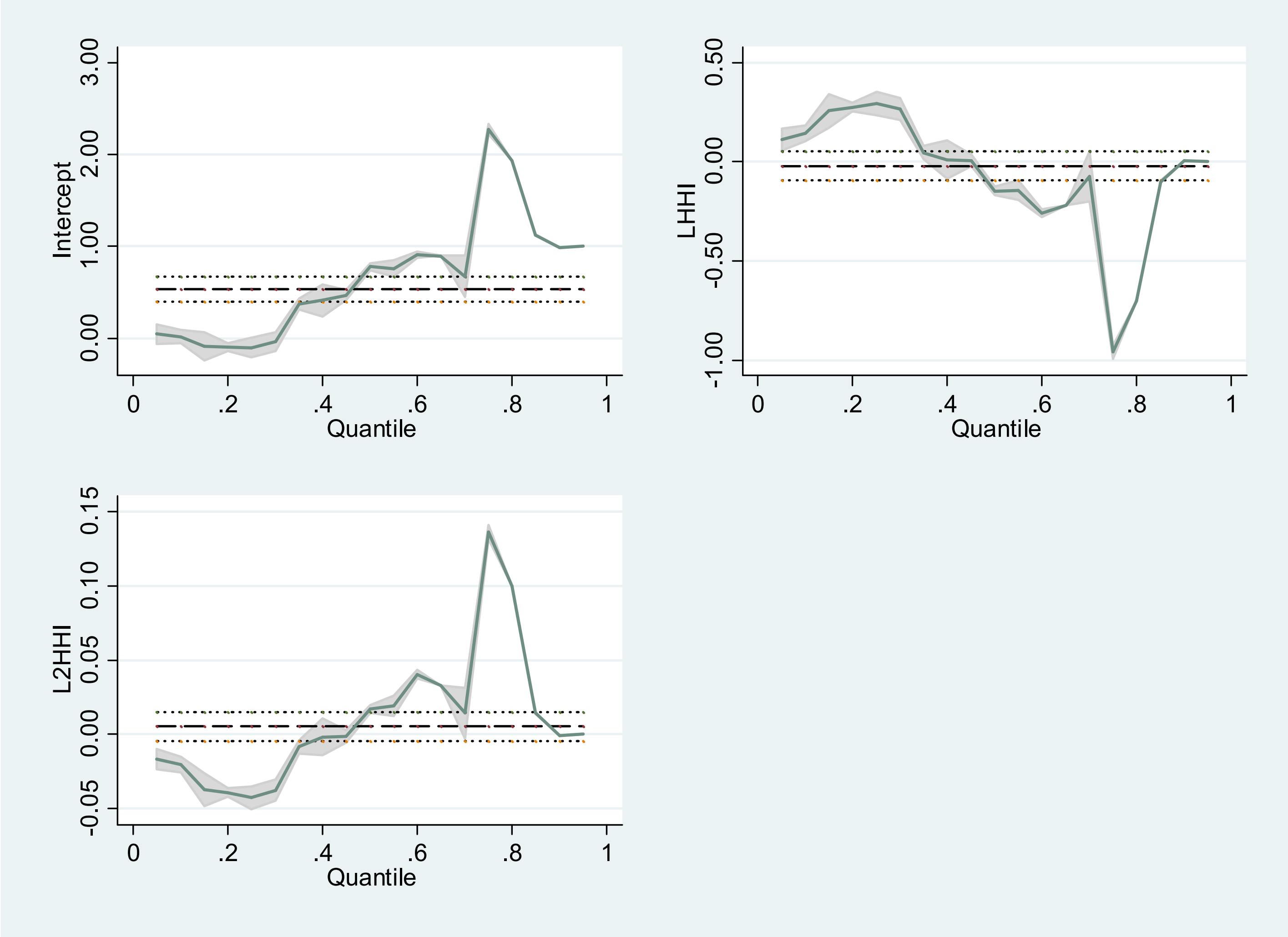
Variable	Model 1	Model 2													
		Quantiles τ													
	OLS	Q (0.01)	Q (0.05)	Q(0.10)	Q(0.15)	Q(0.20)	Q(0.25)	Q(0.30)	Q(0.40)	Q(0.50)	Q(0.60)	Q(0.70)	Q(0.80)	Q(0.90)$^{+}$	Q(0.92)$^{+}$
LHHI	-0.0206 (-0.50)	0.1044*** (3.86)	0.110*** (12.06)	0.143*** (10.78)	0.257*** (7.14)	0.275*** (8.96)	0.292*** (3.85)	0.265** (2.96)	0.0109 (0.17)	-0.147 (-0.82)	-0.259 (-1.82)	-0.0730 (-0.24)	-0.701*** (-3.43)	-0.006 (0.93)	-0.0064*** (12.90)
LHHI^2	0.00532 (0.99)	-0.0176*** (-3.81)	-0.0166*** (-10.83)	-0.0203*** (-9.64)	-0.0373*** (-6.34)	-0.0391*** (-8.47)	-0.0426*** (-4.36)	-0.0375*** (-3.56)	-0.00169 (-0.19)	0.0170 (0.71)	0.0405* (2.19)	0.0143 (0.37)	0.0996*** (3.91)	0.0006 (-0.70)	0.0006*** (-10.21)
Constant	0.536*** (6.95)	0.0457* (1.83)	0.0453*** (4.83)	0.0172 (1.16)	-0.0865** (-2.64)	-0.0985* (-2.15)	-0.101 (-0.70)	-0.0400 (-0.21)	0.414** (3.27)	0.777* (2.52)	0.911*** (3.52)	0.672 (1.15)	1.930*** (5.13)	0.983*** (74.32)	0.984*** (1074.19)
Observations	2,388	2,388	2,388	2,388	2,388	2,388	2,388	2,388	2,388	2,388	2,388	2,388	2,388	2,388	2,388
Turning point (logged values)	-	2.960	3.313	3.522	3.445	3.517	3.427	3.533	-	-	-	-	3.519	-	4.765
Turning point (measurement units)	-	912	2,057	3,328	2,786	3,286	2,674	3,415	-	-	-	-	3,304	-	582
Shape of curve	-	Inverted U-shape	Inverted U-shape	Inverted U-shape	Inverted U-shape	Inverted U-shape	Inverted U-shape	Inverted U-shape	-	-	-	-	U-shape	-	U-shape

Notes: The dependent variable in both models is the efficiency score (EFF) obtained by the DEA method. LHHI is the logged Hirschman-Herfindahl index of the fifty largest companies in the sector. LHHI^2 is the squared logged Hirschman-Herfindahl index of the fifty largest companies in the sector. Model 1 was estimated using OLS and allowing for robust standard errors. Model 2 was estimated using the Bootstrapped quantile regressions methodology at different quantiles τ (0.01, 0.05, 0.10 0.15. 0.20. 0.25. 0.30, 0.40, 0.50, 0.60, 0.70, 0.80) allowing for 100 repetitions. "-" means that the turning point cannot be estimated since the polynomial coefficients of the market concentration variable (HHI) are not statistically significant. $^{+}$ Denotes that the relevant estimates were drawn by using a quantile regression model. The numbers in parentheses denote the t-statistics, while the numbers in square brackets are the p-values. Significant at ***1%, **5% and *10% respectively.

Lastly, the estimates of the polynomial equation (Model 2) change their signs as we are moving across the conditional distribution of the dependent variable (50^{th} quantile) starting from negative (LHHI) to positive ($LHHI^2$). However, these estimates are not statistically significant and therefore the exact pattern of the relationship between market structure and efficiency cannot be determined. The situation clearly changes from the 80^{th} quantile where there is a stable U-shaped pattern. Alternatively, competition and efficiency are linked with an inverted U-shaped curve. Taken these together we argue that, in higher quantiles, the competition effect is not so strong induces the inverted U-shaped curve to shift to the left due to the smaller estimated turning point (582 vs 3,304).

Similarly, Figure 2 portrays the variation of the polynomial estimates at various quantiles and in comparison with the OLS estimates. Regarding the first polynomial term (LHHI) it is evident that its estimated coefficient takes positive values nearly until the median, reversing fully its pattern henceforth. In contrast the squared polynomial term (L2HHI) follows the opposite direction taking firstly negative values and then positive ones. The above findings are consistent with a changing non-linear pattern between efficiency and market structure (competition). Moreover, we see that the effect of market structure differs considerably, having a strong effect on efficiency at lower quantiles and the median estimate is close to the OLS point estimate although not statistically significant. Lastly, we perform the test for the equivalence of the quantile estimates across quantiles. Similarly to the pre-crisis model, the tests clearly reject equality of the estimated coefficients for the three quartiles (0.25, 0.50 and 0.75) in each of the two polynomial terms with p-values equal to 0.000 in both cases.

Figure 2: Variation of LHHI estimates across quantiles for the year 2012



Notes: The dependent variable in both models (OLS and QR) is the efficiency scores obtained by the DEA method (EFF). LHHI is the logged Hirschman-Herfindahl index of the fifty largest companies in the sector. L2HHI is the squared logged Hirschman-Herfindahl index of the fifty largest companies in the sector. The bold dotted horizontal line denotes the fixed OLS estimate. The dotted lines and the grey area denote the confidence bands for the OLS and QR estimates respectively.

4.2 Robustness checks

In order to check for the robustness of our findings, we re-estimate our model which is accordingly adjusted for the presence of an alternative concentration variable namely CR8. This structural indicator captures the impact of the eight largest firms (measured on a 1-100 scale) in the industry. The empirical results when a different concentration ratio is taken into consideration do not reveal significant differences regarding the behaviour and the relevant estimates of the polynomial form.

Regarding the pre-crisis model (2002), we observe that market structure affects efficiency in a non-linear way expressed by a U-shaped curve until the 70^{th} quantile (see columns 2-12 of Table 4). It is worth mentioning that the OLS estimates are statistically significant and in alignment with the previous finding (see column 1). This trend is interrupted as we are moving across the higher quantiles of the distribution (see columns 13-15). Specifically, the relationship between efficiency and market structure remains non-monotonic but with a different curvature (inverted U-shape). In other words, from the 80^{th} quantile, market structure seems to generate a reversal effect on the level of industrial efficiency. The above findings are also reported in Figure 3. From the inspection of the relevant figure, we argue that the effect of market structure differs considerably, having a strong effect on efficiency at higher quantiles and the median estimate is far from the OLS point estimate although statistically significant.

Lastly, we conclude our analysis with the description of the empirical findings of the post-crisis model (2012) drawn from the two estimation techniques (OLS and QR). The relevant estimates are reported in Table 5. As a general statement, we may argue that the OLS and QR estimates are statistically significant in the majority of the quantiles.

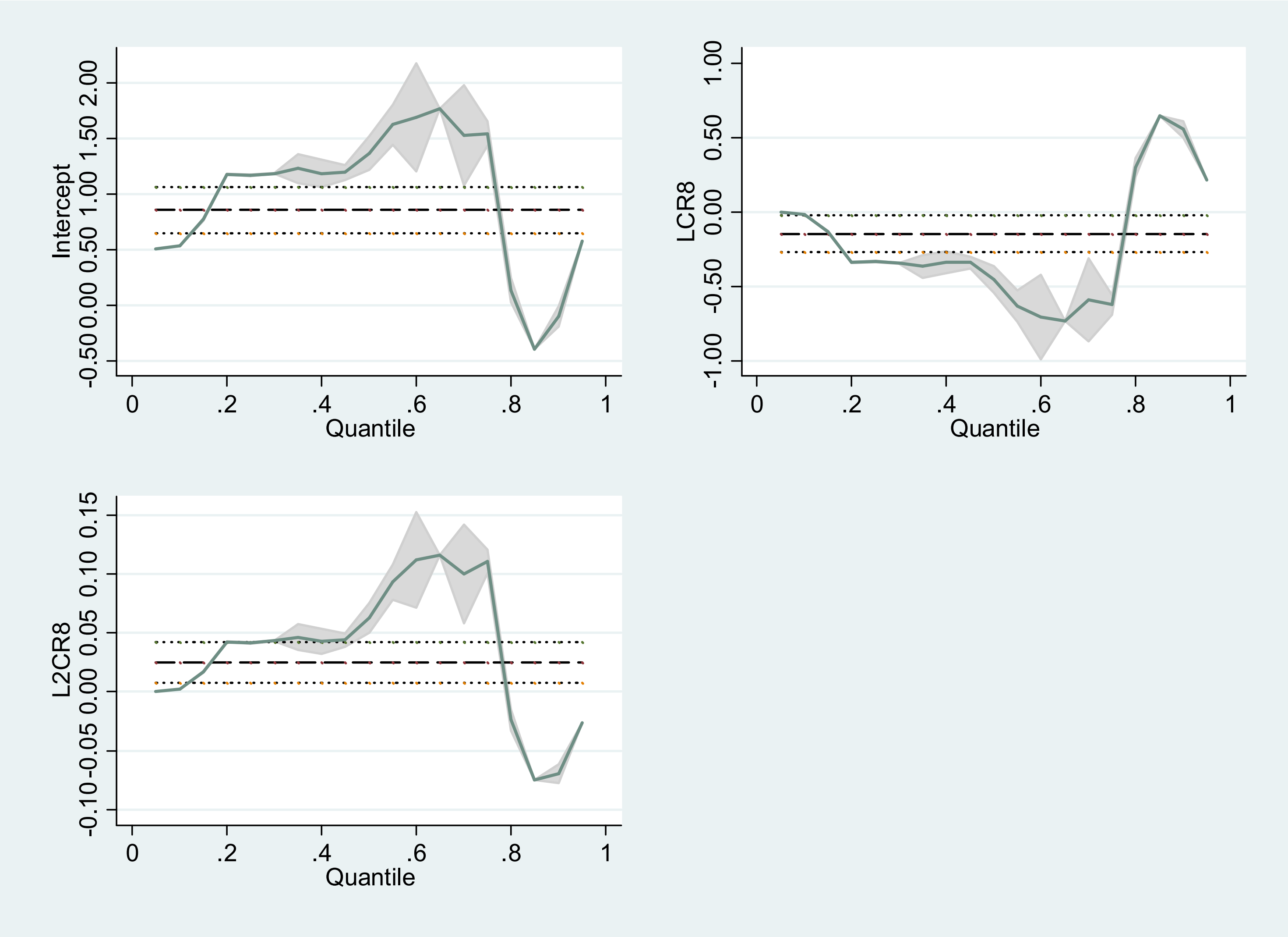
This means that in higher quantiles the competition effect shifts the U-shaped curve to the right due to the larger estimated turning point (85) compared to the previous quantiles of the efficiency distribution (46 and 11). In other words, after the global financial crisis that also hit the USA, market concentration (competition) in the industrial sectors is related with a decrease (increase) of the efficiency after crossing a relatively high threshold (85). In this case, the impact of concentration on industrial efficiency alternates its sign depending on the different competitive regime (low vs high concentration). The latter is also evident in the following diagram.

Table 4: Estimation results for the year 2002

Variable	**Model 3**	**Model 4**													
		Quantiles τ													
	OLS	**Q(0.01)**	**Q (0.05)**	**Q(0.10)**	**Q(0.15)**	**Q(0.20)**	**Q(0.25)**	**Q(0.30)**	**Q(0.40)**	**Q(0.50)**	**Q(0.60)**	**Q(0.70)**	**Q(0.80)**	**Q(0.90)$^{+}$**	**Q(0.95)$^{+}$**
ln(CR8)	-0.147** (-2.71)	-.00123*** (-3.63)	-0.00245 (-0.72)	-0.0179 (-0.82)	-0.134* (-2.18)	-0.338*** (-5.28)	-0.333*** (-31.73)	-0.344*** (-17.71)	-0.337*** (-10.49)	-0.456*** (-5.04)	-0.707* (-2.48)	-0.591** (-3.12)	0.299 (1.63)	0.555*** (10.01)	0.213*** (47.08)
ln(CR8)^2	0.0249** (3.04)	.00016*** (3.61)	0.000296 (0.70)	0.00221 (0.80)	0.0168* (2.14)	0.0423*** (5.27)	0.0417*** (25.27)	0.0432*** (14.54)	0.0428*** (9.04)	0.0628*** (4.35)	0.112* (2.55)	0.0998*** (3.35)	-0.0239 (-0.89)	-0.069*** (-8.76)	-0.026*** (-40.48)
Constant	0.856*** (10.09)	.5024553*** (748.31)	0.505*** (75.10)	0.536*** (12.51)	0.771*** (6.31)	1.180*** (9.18)	1.172*** (73.48)	1.186*** (38.64)	1.186*** (22.96)	1.369*** (10.22)	1.693*** (3.92)	1.529*** (5.48)	0.134 (0.48)	-0.100 (-1.06)	0.573*** (73.89)
Observations	2,391	2,391	2,391	2,391	2,391	2,391	2,391	2,391	2,391	2,391	2,391	2,391	2,391	2,391	2,391
Turning point (logged values)	2.952	4.060	-	-	4.018	4.017	4.026	3.981	3.937	3.631	3.156	2.961	-	3.991	4.058
Turning point (measurement units)	9	11	-	-	97	99	98	96	86	43	14	9	-	10	11
Shape of curve	U-shape	U-shape	-	-	U-shape	U-shape	U-shape	U-shape	U-shape	U-shape	U-shape	U-shape	-	Inverted U-shape	Inverted U-shape

Notes: The dependent variable in both models is the efficiency scores obtained by the DEA method (EFF). Ln(CR8) is the sum of the market shares of the eight larger companies in the sector. Ln(CR8)^2 is the squared sum of the market shares of the eight larger companies in the sector. Model 3 was estimated using OLS and allowing for robust standard errors. Model 4 was estimated using the Bootstrapped quantile regressions methodology at different quantiles τ (0.01, 0.05, 0.10 0.15. 0.20. 0.25. 0.30, 0.40, 0.50, 0.60, 0.70, 0.80) allowing for 100 repetitions. "-" means that the turning point cannot be estimated since the polynomial coefficients of the market concentration variable (CR8) are not statistically significant. $^{+}$ Denotes that the relevant estimates were drawn by using a quantile regression model. The numbers in parentheses denote the t-statistics, while the numbers in square brackets are the p-values. Significant at ***1%, **5% and *10% respectively.

Figure 3: Variation of LCR8 estimates across quantiles for the year 2002



Notes: The dependent variable in both models (OLS and QR) is the efficiency scores obtained by the DEA method (EFF). LCR8 is the sum of the market shares of the eight larger companies in the sector. L2CR8 is the squared sum of the market shares of the eight larger companies in the sector. The bold dotted horizontal line denotes the fixed OLS estimate. The dotted lines and the grey area denote the confidence bands for the OLS and QR estimates respectively.

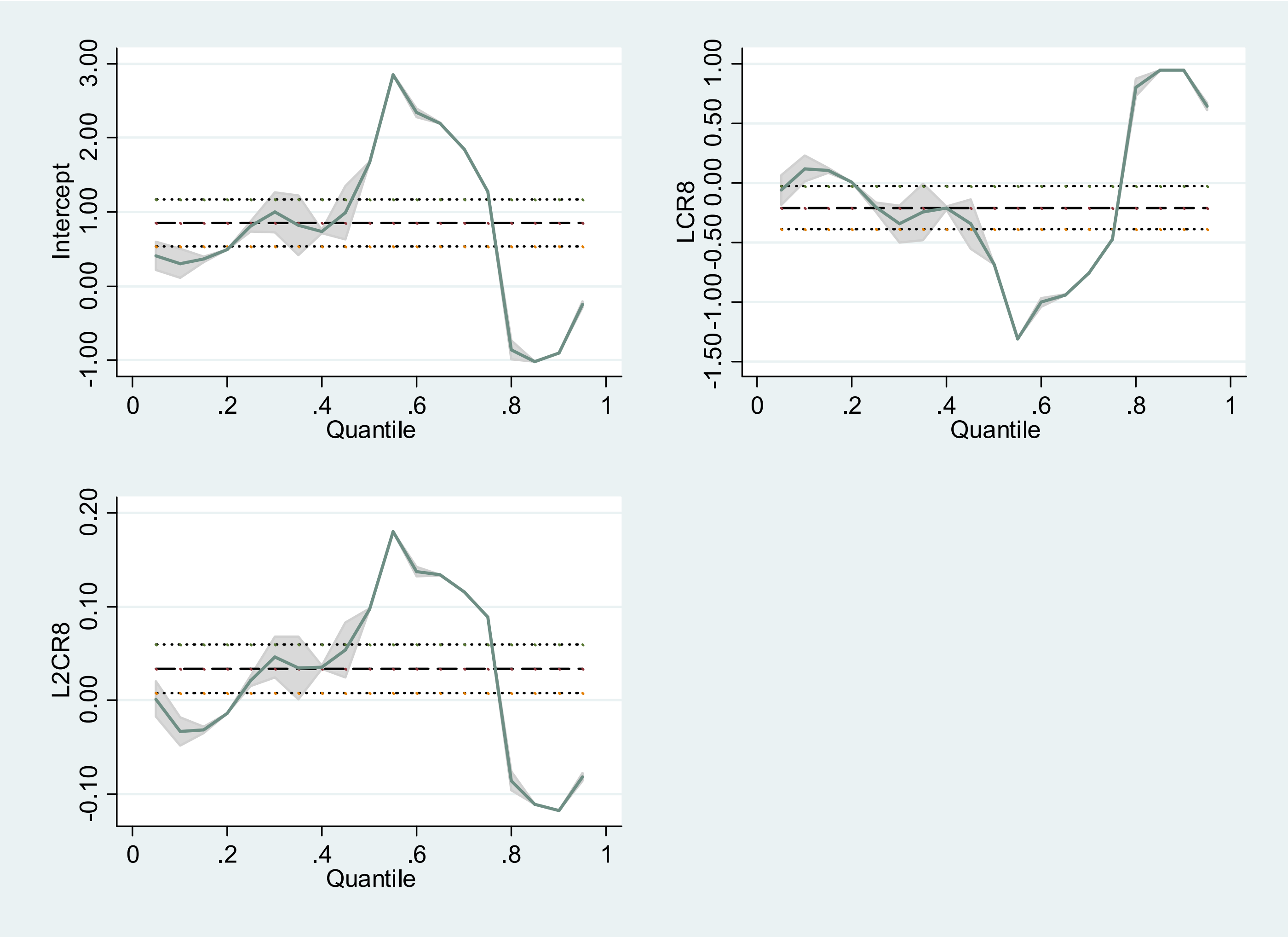
A closer look though at the estimation results reveals significant differences. More specifically, the OLS estimates uncover a non-monotonic U-shape form between efficiency and market structure, which is evident also in most of the quantiles from the 25th (first quartile) to the 70th quantile (see columns, 7-12). This is in contrast with the previous baseline findings where market structure was proxied by the HHI (see Table 3). Another interesting result is that the conditional distribution of efficiency is fully reversed in higher quantiles (from 80th quantile) revealing an inverted U-shaped form.

Table 5: Estimation results for the year 2012

Variable	Model 3	Model 4													
		Quantiles τ													
	OLS	**Q(0.01)**	**Q (0.05)**	**Q(0.10)**	**Q(0.15)**	**Q(0.20)**	**Q(0.25)**	**Q(0.30)**	**Q(0.40)**	**Q(0.50)**	**Q(0.60)**	**Q(0.70)**	**Q(0.80)**	**Q(0.90)**	**Q(0.95)**
ln(CR8)	-0.211* (-2.46)	-0.0355 (-0.72)	-0.0632 (-0.43)	0.118 (0.78)	0.103* (1.53)	0.00630 (0.09)	-0.208** (-2.58)	-0.347*** (-4.08)	-0.212* (-1.88)	-0.687* (-1.85)	-1.004*** (-4.56)	-0.760* (-1.90)	0.801** (2.64)	0.9476*** (5.09)	0.6398*** (7.70)
ln(CR8)^2	0.0338** (2.71)	-0.0054 (-0.64)	0.00121 (0.06)	-0.0334 (-1.51)	- 0.0313** (-3.13)	-0.0137 (-1.26)	0.0211 (1.62)	0.0465*** (3.41)	0.0351* (2.09)	0.0974* (2.09)	0.137*** (4.22)	0.116* (1.91)	-0.0859* (-1.99)	-0.1172*** (-5.02)	- 0.0814*** -7.64
Constant	0.854*** (5.97)	0.38665*** (5.27)	0.409 (1.49)	0.306 (1.25)	0.361*** (3.33)	0.493*** (4.97)	0.806*** (6.83)	0.994*** (7.90)	0.737*** (3.84)	1.669* (2.26)	2.337*** (6.60)	1.845** (2.90)	-0.863* (-1.72)	-0.9074** (-2.46)	-0.2461* (-1.52)
Observations	2,388	2,388	2,388	2,388	2,388	2,388	2,388	2,388	2,388	2,388	2,388	2,388	2,388	2,388	2,388
Turning point (logged values)	3.121	-	-	-	1.645	-	4.929	3.731	3.020	3.527	3.664	3.276	4.662	4.043	3.930
Turning point (measurement units)	13	-	-	-	44	-	85	54	10	34	46	19	46	11	85
Shape of curve	U-shape	-	-	-	Inverted U-shape	-	U-shape	U-shape	U-shape	U-shape	U-shape	U-shape	Inverted U-shape	Inverted U-shape	Inverted U-shape

Notes: The dependent variable in both models is the efficiency scores obtained by the DEA method (EFF). Ln(CR8) is the sum of the market shares of the eight larger companies in the sector. Ln(CR8)^2 is the squared sum of the market shares of the eight larger companies in the sector. Model 3 was estimated using OLS and allowing for robust standard errors. Model 4 was estimated using the Bootstrapped quantile regressions methodology at different quantiles τ (0.01, 0.05, 0.10 0.15. 0.20. 0.25. 0.30, 0.40, 0.50, 0.60, 0.70, 0.80) allowing for 100 repetitions. "-" means that the turning point cannot be estimated since the polynomial coefficients of the market concentration variable (CR8) are not statistically significant. The numbers in parentheses denote the t-statistics, while the numbers in square brackets are the p-values. Significant at ***1%, **5% and *10% respectively.

Figure 4: Variation of LCR8 estimates across quantiles, and with the (fixed) OLS estimates for the year 2012



Notes: The dependent variable in both models (OLS and QR) is the efficiency scores obtained by the DEA method (EFF). LCR8 is the sum of the market shares of the eight larger companies in the sector. L2CR8 is the squared sum of the market shares of the eight larger companies in the sector. The bold dotted horizontal line denotes the fixed OLS estimate. The dotted lines and the grey area denote the confidence bands for the OLS and QR estimates respectively.

5. Conclusions

In this study, we use a unique data set at the plant level comprising by hundreds of industrial facilities dispersed among the US states before and after the financial crisis (2002 and 2008), in order to delineate the effects of industrial pollution prevention activities on toxic chemical releases under the presence of two market regimes (competitive and non-competitive conditions).

In order to measure the relevant efficiency scores, we use DEA techniques allowing for the inclusion of desirable and undesirable (toxic chemical releases) outputs in the production function. In the next stage, we utilise the bootstrapped quantile regression methodology to uncover possible non-linear relationships between efficiency and competition at the mean and at various quantiles. In this way, we impose no functional form constraints on parameter values over the conditional distribution of the dependent variable (efficiency). At the same time, we estimate at which part of its conditional distribution function, the efficiency is located and draw substantial conclusions about the range of policy measures obtained.

Our empirical findings justify for the first time in the empirical literature the existence of a non-monotonic relationship between market concentration (competition) and efficiency in the industrial sector. Moreover, we notice that the relationship between efficiency and market concentration did not remain unchanged in the aftermath of the economic crisis. This relationship provides new insights into the environmental policy since the regulators must take into account if they are on the upward or the downward slopping part of the curve in order to pursue the effective environmental policies. Moreover, we argue that the existence of a non-linear relationship when toxic chemical releases form the undesirable output in the production process provides new insights into the policy agenda toward emissions releases abatement. This means that policy makers and practitioners must take into account if they are on the upward or the downward slopping part of the curve. It is worth emphasizing that the increasing nonparametric regression line up to a certain concentration level approximately indicates a negative effect on facilities' emissions levels whereas a decreasing line indicates a positive effect.

References

Apergis, N., and Christou, C, (2015). The behaviour of the bank lending channel when interest rates approach the zero lower bound: Evidence from quantile regressions, *Economic Modelling,* **49(C)**: 296-307.

Apergis N. and Polemis M. (2016). Competition and efficiency in the MENA banking region: a non-structural DEA approach. *Applied Economics,* **48(54)**: 5276-5291.

Branco F. and Villas-Boas J. (2015). Competitive Vices. *Journal of Marketing Research*, **52(6)**: 801-816.

Bi X., and Khanna M. (2012) Reassessment of the Impact of the EPA's Voluntary 33/50 Program on Toxic Releases. *Land Economics*, **88**: 341-361

Buchinsky M. (1995). Estimating the asymptotic covariance matrix for quantile regression models: A Monte Carlo study. *Journal of Econometrics*, **68:** 303–338.

Buchinsky M. (1998). Recent advances in quantile regression models. *Journal of Human Resources*, *33*(1), 88–126.

Chen C., Polemis M., Stengos T. (2018). On the examination of non-linear relationship between market structure and performance in the US manufacturing industry. *Economics Letters*, **164**: 1-4

Cooper W.W., Seiford L.M., Tone K. (2007). *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-solver Software.* 2nd Ed. Springer Science and Business Media, New York, NY.

Dai M., Liu Q. and Serfes K. (2014). Is the effect of competition on price dispersion nonmonotonic? Evidence from the U.S. airline industry. *The Review of Economics and Statistics*, **96(1)**: 161–170.

Distante, R., Petrella, I., and Santoro, E. (2018). Gibrat's law and quantile regressions: An application to firm growth, *Economics Letters,* **164**: 5–9

Färe R., Grosskopf S., Lovel C.A.K. and Pasurka S. (1989). Multilateral productivity comparisons when some outputs are undesirable: a non parametric approach. *Review of Economics and Statistics*, **71(1)**: 90–98.

Färe R. and Primont D. (1995). *Multi-Output Production and Duality: Theory and Applications*. Kluwer Academic Publishers, Boston.

Halkos, G. (2003). Environmental Kuznets Curve for sulfur: evidence using GMM estimation and random coefficient panel data models, *Environment and Development Economics*, **8(04)**: 581-601.

Halkos G.E. (2011). Nonparametric modelling of biodiversity: Determinants of threatened species, *Journal of Policy Modeling*, **33(4):** 618-635.

Halkos, G. E., and Polemis, M. L. (2017). Does Financial Development Affect Environmental Degradation? Evidence from the OECD Countries. *Business Strategy and the Environment*, **26:** 1162–1180

Halkos G. and Polemis M. (2018). The impact of economic growth on environmental efficiency of the electricity sector: A hybrid window DEA methodology for the USA. *Journal of Environmental Management*, **211:** 334-346.

Halkos G. and Tzeremes N.G. (2009). Exploring the existence of Kuznets curve in countries' environmental efficiency using DEA window analysis. *Ecological Economics*, **68**: 2168–2176.

Harrington D.R., Deltas, G., and Khanna M. (2014). Does Pollution Prevention Reduce Toxic Releases? A Dynamic Panel Data Model. *Land Economics*, **90** (2): 199-221.

Hsueh L. (2015). Regulatory effectiveness and the long-run policy horizon: The case of U.S. toxic chemical use. *Environmental Science and Policy*, **52**: 6 – 22.

Jones M.C. (1992). Estimating densities, quantiles, quantiles densities and density quantiles. *Annals of the Institute of Statistical Mathematics*, **44(4)**, 721-727.

Koenker R. (2005). *Quantile regression*. New York: Cambridge University Press.

Machado, J.A.F., Mata, J., (2000). Box–Cox quantile regression and the distribution of firm sizes. *Journal of Applied Econometrics,* **15** (3): 253–274.

Mansur, E (2007) Do oligopolists pollute less? Evidence from a restructured electricity market. *Journal of Industrial Economics*, **55** (4): 661–689

Millimet D., List J. and Stengos T., (2003). The Environmental Kuznets Curve: Real Progress or Misspecified Models?, *The Review of Economics and Statistics*, **85(4)**, 1038-1047.

Polemis., (2017) Revisiting the Environmental Kuznets Curve: A semi-parametric analysis on the role of market structure on environmental pollution, Michael Polemis, *Letters in Spatial and Resource Sciences* https://doi.org/10.1007/s12076-017-0195-9.

Newson R. (2012). Sensible parameters for univariate and multivariate splines. *Stata Journal*, **12-3**: 479-504.

Picazo-Tadeo A., Beltrá-Esteve M. And Gómez-Limón J.A. (2012). Assessing eco-efficiency with directional distance functions. *European Journal of Operational Research*, **220**: 798-809.

Polemis M. and Stengos T. (2017). Does Competition Prevent Industrial Pollution? Evidence from a Panel Threshold Model, Working Paper Series **17-07**, The Rimini Centre for Economic Analysis.

Powell J. (1986). Censored regression quantiles. *Journal of Econometrics*, **32:** 143–155.

Scheel H. (2001). Undesirable outputs in efficiency valuations. *European Journal of Operational Research*, **132**: 400–410.

Seiford LM, Zhu J (2002) Modeling undesirable factors in efficiency evaluation. *European Journal of Operational Research*, **142**:16–20

Shephard RW (1970). *Theory of cost and production function.* Princeton University Press, Princeton.

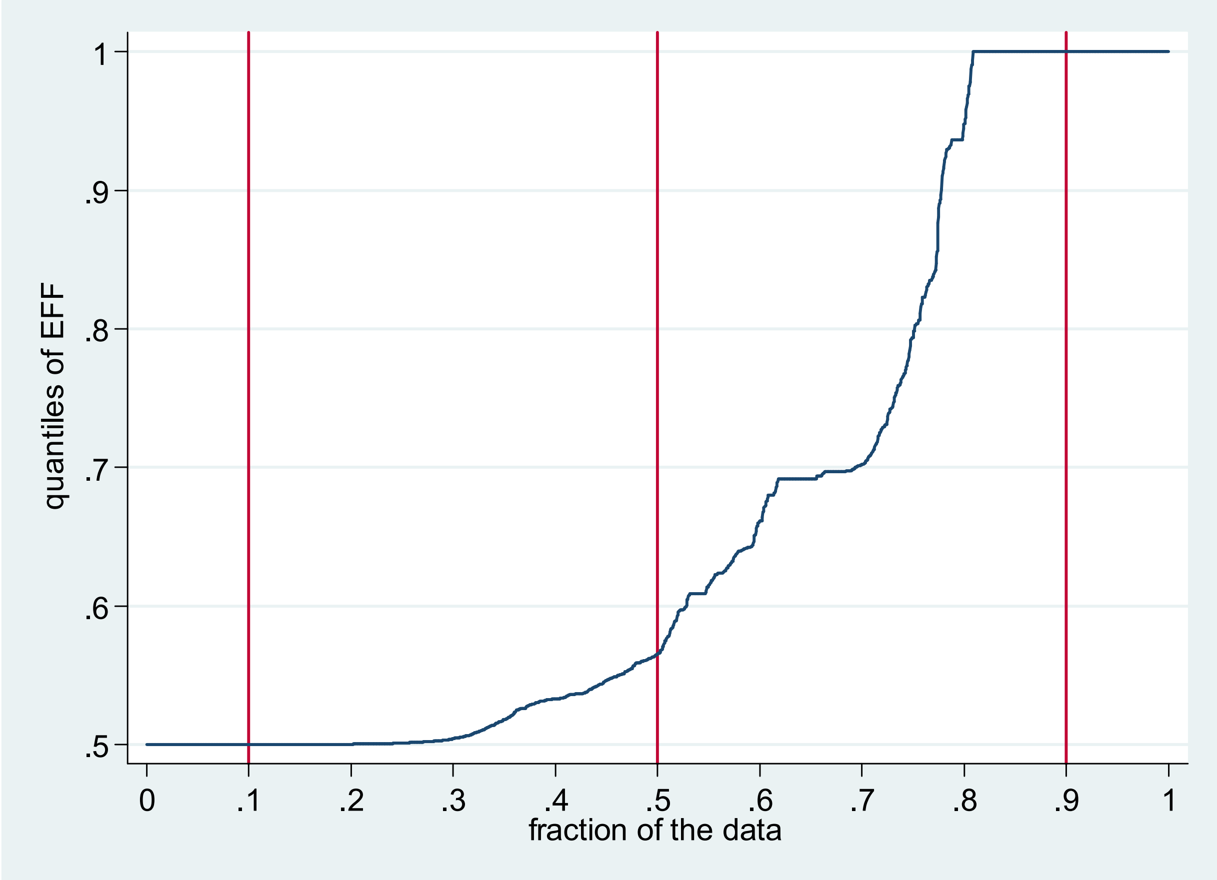
Shleifer A. (2004). Does Competition Destroy Ethical Behavior? *American Economic Review*, **94**: 414-418.

Simon, DH., and Prince J.T. (2016). The Effect of Competition on Toxic Pollution Releases, *Journal of Environmental Economics and Management*, **79**: 40-54.

Zarzoso, I.M and Morancho. A.B., (2004). Pooled mean group estimation of an environmental Kuznets curve for CO_2. *Economics Letters,* **82,** 121–126.

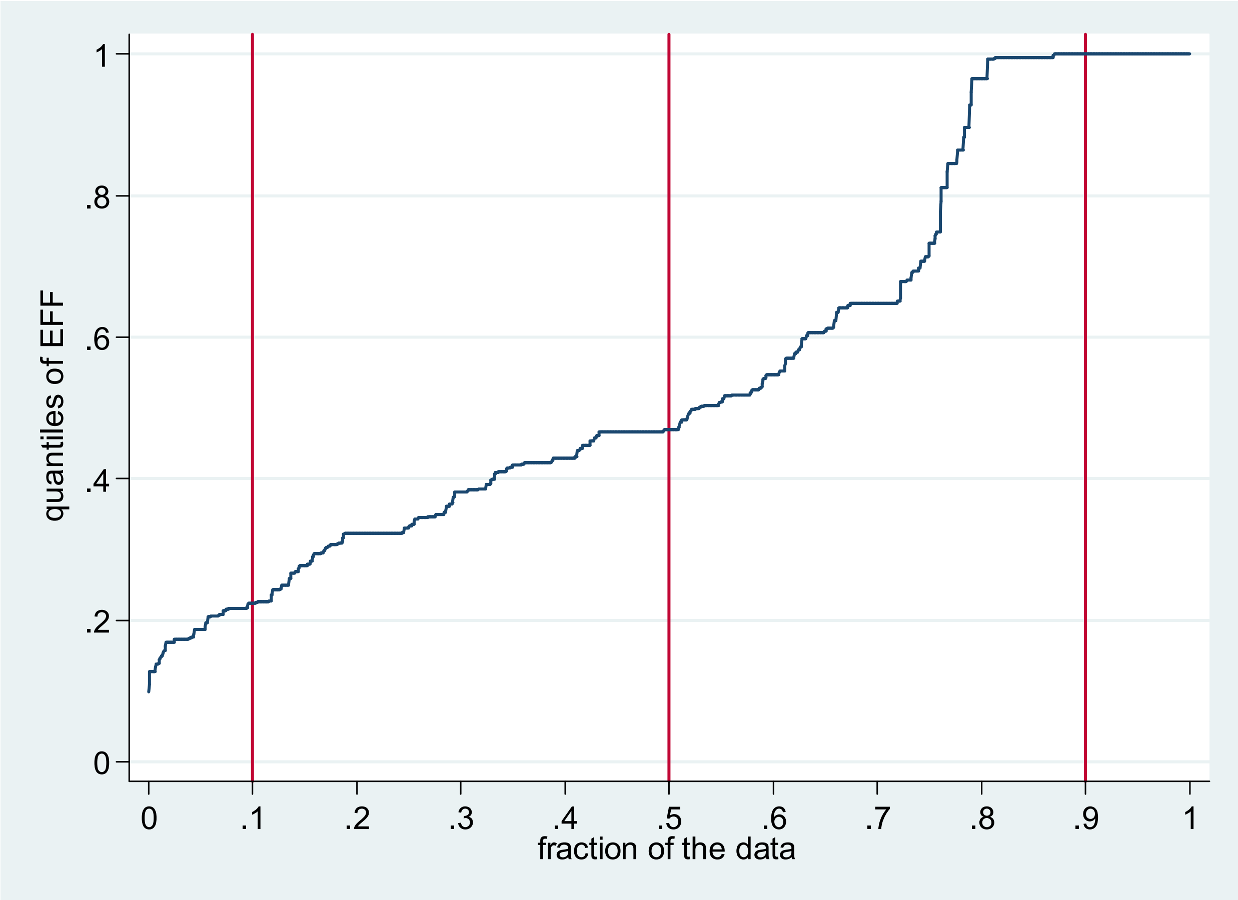
Appendix

Figure 1A: Empirical cumulative distribution function (CDF) of efficiency (2002).



Notes: The continuous (blue) line is the estimated CDF associated with the dependent variable (EFF) across the various quantiles (τ) for the year 2002. The vertical (red) lines represent the 10^{th}, 50^{th} (median) and 90^{th} quantile respectively.

Figure 2A: Empirical cumulative distribution function (CDF) of efficiency (2012).



Notes: The continuous (blue) line is the estimated CDF associated with the dependent variable (EFF) across the various quantiles (τ) for the year 2012. The vertical (red) lines represent the 10^{th}, 50^{th} (median) and 90^{th} quantile respectively.