Munich Personal RePEc Archive

Does Competition Prevent Industrial Pollution? Evidence from a Panel Threshold Model

Polemis, Michael and Stengos, Thanasis

University of Piraeus, Department of Economics Piraeus, Greece, University of Guelph, Department of Economics, University of Guelph, Guelph, Ontario, Canada.

24 February 2017

Online at https://mpra.ub.uni-muenchen.de/85177/
MPRA Paper No. 85177, posted 14 Mar 2018 12:33 UTC

Does Competition Prevent Industrial Pollution? Evidence from a Panel Threshold Model

Michael L. Polemis[a,b] and Thanasis Stengos[c]

[a] University of Piraeus, Department of Economics Piraeus, Greece, E-mail address: mpolemis@unipi.gr
[b] Hellenic Competition Commission, Athens, Greece
[c] University of Guelph, Department of Economics, University of Guelph, Guelph, Ontario, N1G 2W1: Tel: (519) 824-4120

Abstract

The scope of this paper is to assess the impact of competition on industrial toxic pollution by using for the first time, a panel threshold model which allows evaluating the main drivers of toxic releases under two different market regimes. The empirical analysis is based on a micro level panel data set over the five year-period 1987-2012. We show that this relationship is statistically significant and robust above and below the threshold, even after accounting for alternative specifications of market concentration. Finally, we unmask an inverted "*V-shaped*" relationship between market concentration and industrial pollution.

Keywords: Competition; Concentration; Threshold; Non-linearities; Inverted "*V-shaped*"
JEL codes: L10; L60; Q52; C23; C24

1. Introduction

Manufacturing activities such as metal mining, electric power generation, oil refining, recycling, use chemicals to produce the products we consume (i.e pharmaceuticals, computers, paints, clothing, automobiles, etc). While the majority of toxic chemicals are managed by industrial facilities via strict regulations in order to minimize chemicals into the air, water and land, toxic releases do still occur as part of their everyday business operations (Levinson, 2015). To give but an example, of the scale of their use, it is worth mentioning that in 2015 and only for the US almost 3.36 billion pounds of total chemical disposals (including on-site and off-site releases) were released many of them can be regarded as hazardous waste. Hopefully, nearly 26 billion pounds covering approximately the 92% of total chemical waste (excluding metal mines), was not released into the atmosphere due to the use of preferred waste management practices such as recycling, energy recovery, and treatment (EPA, 2017).

Based on the above considerations, it is common knowledge that industrial pollution affects the entire spectrum of the optimal use of natural and environmental resources to economic activity (Hsueh, 2015; Shapiro and Walker, 2015; Harrington *et al*, 2014; Bi and Khanna, 2012; Levinson, 2009). Over the last ten years researchers have tried to disentangle this relationship. Specifically, one strand of literature tries to explore possible linkages between the level of environmental pollution and serious health problems such as asthma, infant health and mortality, lung cancer, cardiovascular diseases (see for example Rzhetsky *et al*, 2014; Agarwal *et al*, 2010; Currie *et al* 2009; Currie and Schmieder, 2008). The second strand of literature, tries to investigate the possible spillover effects between environmental degradation and market structure (e.g. Simon and Prince, 2016; Branco and Villas-Boas, 2015; Fowlie, 2009). 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). This strand is rapidly growing. Our approach is one of the very few attempts at modeling and estimating the decision of US firms on their participation using facility level data. For this reason, we formulate a number of research questions including *inter alia* the following: How does market concentration affect chemical releases? How does competition generate industry output and emissions? In what way the level of pollution is determined under different market regimes? Why does the level of industrial pollution is lower (higher) under the presence of (non)-competitive conditions in the market? Will more concentrated industries pollute more because of inefficiency driven by lack of technological innovation and limited competition or are they polluting more because of the correlation between market concentration and fixed costs. Lastly, what policy implications could be drawn in order to boost an efficient abatement mechanism?

The main novelty of our study is that we use for the first time in the relevant literature a panel sample splitting methodology (threshold model) accounting for the decomposition of Significant Market Power (SMP) in an industry and linking the possible interactions with the level of industrial pollution.[1] In this way, we argue that an industry needs to cross a certain level of market concentration (competition) in order to restrict environmental degradation. Explanations offered to account for this argument broadly fall into two categories. According to the first, it is the nature of data and differences in empirical methodology (i.e misspecification and measurement

[1] In oligopolistic markets, SMP is evident in an industry/sector when prices exceed marginal cost (MC) and long run average cost (LAVC), so the firm makes positive economic profits.

error, existence of outliers, lack of data quality, etc) justifying that the effect of competition on industrial pollution might be non-linear. But there is also an economic motivation emphasizing for this justification. More specifically, in oligopolistic sectors (i.e energy, steel industry, oil refining, cement industry, etc) which are characterised by high market concentration and absence of effective competition as a result of the existence of SMP by the incumbent, there are usually strict environmental regulations (i.e taxes, tariffs, fees, etc) in order to limit environmental degradation (Halkos and Papageorgiou, 2016). On the other hand, pollution (i.e toxic chemical releases) rises (falls) with an increase (decrease) of market concentration suggesting that environmental damage is more likely to evolve in oligopolistic sectors (Simon and Prince, 2016; Fowlie, 2009). Based on these justifications, we argue that competition might be different in the two regimes (concentrated and non-concentrated industries). In other words, instead of assuming a linear effect in which we attribute the full impact to one variable (i.e market concentration) we allow this effect to vary at different values of market structure.

Our approach strongly accounts for the presence of cross section dependence while it utilizes "*second-generation*" panel unit root tests in order to uncover possible cointegrated relationships an issue that has been overlooked by the existing empirical literature. The reason for using this kind of unit root testing can be justified by the fact that traditional stationarity tests (known as "*first-generation*" tests) suffer from size distortions and the ignorance of cross section dependence (Pesaran 2015).

The contribution of this paper is three-fold. First, it goes beyond the existing literature in that it uses a unique micro level dataset originated from thousands industrial facilities (polluters) dispersed among the 50 US states. This will help us to empirically explore the net effect of competition on facility-level emissions. Second, it utilizes a panel threshold approach with certain innovations such as the inclusion of structural market characteristics and the treatment of a second threshold in the sample. This technique has been widely used by the literature to identify threshold effects when the variable of interest is observable, but the position of the threshold is not known. Third, and foremost, the paper unveils a stable non-linear inverted "*V-shaped*" relationship between market concentration and industrial pollution already hidden by the existing literature. Taken together, this set of findings is important in that it provides useful policy implications towards the abatement of toxic chemical releases in order to achieve sustainability.

Using a panel threshold framework in the spirit of Hansen (2000), we show that the reason for the mixed evidence of the impact of competition on environmental degradation, (proxied by toxic chemical releases) in an industry lies with its level of market concentration. This implies that market structure cannot assert its role in the process of environmental pollution until an industry crosses a certain threshold level of concentration. Our findings remain robust across alternative market concentration measures (CR4 and HHI). However, the driving force that pushes competition to alter its behaviour toward the level of environmental pollution based on a specific threshold point (generating a "*kinked*" curve) provides an interesting opportunity for future theoretical and empirical research.

The rest of the paper is as follows. Section 2 reviews the empirical literature. Section 3 describes the data and the relevant empirical testing for cross-section dependence and unit roots. Section 4 portrays the econometric methodology used in the empirical analysis. Section 5 discusses the empirical findings of the study, while Section 6 performs some necessary robustness checks. Lastly, Section 7 concludes the paper and provides some policy implications.

2. Literature review

There is a widespread belief that competition is regarded as a reliable mechanism for stimulating both allocative and technical efficiency (Leibenstein 1966). As suggested by many researchers (Zhang *et al*, 2005; Zhang *et al*, 2008; Akkemik and Oguz, 2011), in a competitive market, prices and profits provide the firm with incentives to improve efficiency minimising costs. Further, competition in network industries such as electricity would deliver production and allocative efficiency, hence lower prices, or lower mark-up over costs (Fiorio and Florio, 2013; Chiara Del Bo and Florio, 2012). This will lead to higher industrial output, while lower per-unit costs resulting from increased technical efficiency may be passed through in lower prices, thus increasing the quantity demanded and subsequently the level of environmental pollution (Polemis and Stengos, 2016). Although the positive impacts of competition on total welfare are widely acknowledged by the economists the effect that competition has on environmental pollution is under scrutiny.

Despite the profound interest by policy makers and government officials on the possible spillovers between market competition and environmental degradation the existing literature is still in its infancy, with controversial results. These can be justified by the fact that many researchers acknowledge that competition may have positive as well as negative effects on environmental pollution (Simon and Prince, 2016; Branco and Villas-Boas, 2015; Fowlie, 2009; Mansur 2007; Shleifer, 2004).

In a seminal theoretical paper, Farber and Martin (1986) argue that increased competition lowers industrial output, and thus, at least lowers average production per plant. Therefore, an increase in the environmental pollution along with the production expansion will create a positive effect on the environmental degradation resulting in less pollution per facility. Moreover, they posit that increased competition leads to less abatement efforts by firms, since firms in more concentrated industries spend more on combating air and water pollution than firms in less concentrated industries.

Subsequent work by Shleifer (2004) indicates that effective competition tends to increase the incentives of a firm to undercut costs in order to reduce prices. This can be broadly implemented to combating pollution in a sense that companies may pursue cost reducing strategies, by reducing pollution control activities.

In another study, Fowlie (2009) develops a theoretical model for analyzing the rate of environmental emissions pass-through in tandem with certain welfare implications when effective competition is absent (unregulated industry). This paper highlights the role of market structure on determining the extent of emissions leakage by acclaiming that the more competitive the industry, the greater the effect of incomplete participation on industry emissions. Moreover, the study links the net welfare effects of pollution with the regulation-induced reallocation of production among heterogeneous producers. However more recent work (Branco and Villas-Boas, 2015) tries to shed some light on the theoretical controversy between competition, abatement and pollution. Their main argument, documents that lower production reduces pollution while presumed lower abatement increases facility-level pollution. In other words, the net effect of an increase in competition on facility-level pollution is ambiguous. Cole at al (2013), use the toxic release inventory database for the period 1990 - 2005 to examine the relationship between ethnic divisions and US toxic releases neglecting totally the role of market structure. They argue that measures of ethnic divisions have a positive relationship with toxic releases. Similarly, Zwickl et al (2014), examine the spatial variation between racial and ethnic disparities in industrial air toxic releases in U.S. cities. They claim that the latter are strongest among regions with median income below 25,000 USD dollars. On the other hand, their empirical findings support the argument that income-based disparities are stronger among regions with median incomes above that threshold.

Lastly, in a similar to ours empirical study by Simon and Prince (2016), it is examined the role of market competition on controlling the level of industrial pollution in a form of toxic chemicals. They employ simple estimation techniques (OLS with fixed effects) using a micro panel dataset consisting of a small time period (five years from 1987 to 2007) and a large cross-section element (thousands of industrial facilities in the US). Their empirical findings indicate a robust linear scheme where competition is a stimulating mechanism achieving a reduced level of industrial pollution. In addition, they claim that competition increases abatement since the relevant estimate comes with a negative sign indicating that each percentage-point increase in the level of market concentration (proxied by Hirschman-Herfindahl Index) is associated with a 2.9% reduction in the abatement ratio.

Based on the above, the existing studies do not properly incorporate the spillovers generated by the inclusion of competition on the pollution-abatement nexus since they totally neglect the notion of a sample splitting variable acting as a separator between two different market regimes (more competition *vs* less competition). Our model estimates an unknown threshold parameter in a data driven approach that "*endogenously*" sorts the data into the two different market regimes, whereby each regime would differ according to the prevailing attitudes of its members towards competition (Polemis and Stengos, 2017). The threshold variable that we use to sort observations is the level of concentration measured by well documented in the literature structural indices (CR_n, and Hirschman-Herfindahl index or HHI). Subsequently, the sample facilities will be sorted according to the level of market concentration placing them into competitive (i.e taking low values of the index) and non-competitive (i.e. taking high values of the index). The purpose of this study is to fill these research gaps by combining certain structural industry characteristics (i.e level of employment, capital intensity, market concentration, value added, etc) drawn from a micro economic perspective with the facility-level of toxic chemical releases.

3. Data and empirical testing

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 classical (global and local) pollutants such as CO_2, NO_X or SO_2 is that the latter are only available at a state level. However, we argue that the empirical findings from the use of toxic chemicals can be easily extrapolated to other pollutants in a future research.

The above information is submitted by U.S. facilities in industry sectors such as manufacturing, metal mining, electric utilities, and commercial hazardous waste management. Under the Emergency Planning and Community Right-to-Know Act (EPCRA), facilities must report their toxic chemical releases for the prior calendar year to EPA by 1st of July of each year. Moreover, the Pollution Prevention Act also requires facilities to submit information on pollution prevention and other waste management activities of TRI chemicals (EPA, 2017). Each basic data file contains 108 data fields, which generally represent these categories:

a) Facility name, address, latitude and longitude coordinates, SIC or NAICS codes and Industry Sector Codes

b) Chemical identification and classification information

c) On-site release quantities

d) Publicly Owned Treatment Works (POTW) Transfer Quantities

e) Off-site Transfer Quantities for Release/Disposal and Further Waste Management

f) Summary Pollution Prevention Quantities

The relevant database is rapidly growing since in 1987, it included nearly 275 toxic chemicals, while by 2015 this number had nearly doubled, to 600 chemicals, with some of the original chemicals dropped from the reporting requirements. However, for purposes of consistency, we restrict our sample to the 234 chemicals that appeared in the 1987 dataset and have been reported in every year since (Simon and Prince, 2016). Lastly, it is worth mentioning that the set of chemicals included in the TRI has evolved over time since primarily new chemicals have been added. However for consistency purposes we have excluded them from the sample selection.[2]

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. Because of the change from SIC to NAICS industry definitions in 1997, the database is provided in two versions (one with 459 four-digit 1987 SIC industries and the other with 473 six-digit 1997 NAICS industries). Especially for the year 2012, and due to data restrictions concerning the level of market concentration as measured by certain indicators (i.e CR4, CR8, CR20, CR50 and HHI), we used data directly from the US Census of Manufacturers. The

[2] Facilities are required to report their toxic releases to the EPA if they meet the following three criteria: (1) They have ten or more full-time employees (or the equivalent); (2) They are in a covered industry (all manufacturing industries, mining, electricity generation, hazardous waste facilities, along with some publishing and wholesale trade industries); and (3) They “manufactured" or "processed" more than 25,000 pounds or "otherwise used" more than 10,000 pounds of any listed toxic chemical during a calendar year. Plants that meet the first two criteria must report releases for each toxic chemical that exceeds the threshold in (3).

latter is only conducted every five years limiting our time span to six years (1987, 1992, 1997, 2002, 2007 and 2012). We must mention though that while there are obvious benefits of having a strongly-balanced panel, there are costs as well. Due to the fact that TRI database constitutes an unbalanced panel, observations were dropped to balance the panel thus restricting a 25 year panel to 6 years.

In order to check the robustness of our findings, we take five measures of market concentration: HHI is the Herfindahl-Hirschman index for the 50 largest firms in the industry[3], CR4 is the four-firm concentration ratio, CR8 is the eight-firm concentration ratio, CR20, is the twenty-firm concentration ratio and finally CR50, is the fifty-firm 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 (Cabral, 2017).

Our sample consists of thousands observations, namely, 2,461 panels (facilities) times 6 years, and the panel data set is strongly balanced. We excluded observations for facilities with missing values for toxic chemical releases. Hence our sample includes 14,767 plants, spread across 356 six-digit NAICS industry codes. Especially, for the years 1987 and 1992 we used the SIC classification. Similarly to Simon and Prince (2016) we used the TRI database in order to incorporate the level of toxic chemical releases in our sample. However, there is significant difference in the magnitude of the two samples. Our sample consists of nearly 11,900 facility year observations, while the aforementioned study includes more than 80,000 observations. This discrepancy is the result of merging the two databases (i.e TRI with the NBER dataset). Lastly, similarly to Simon and Prince (2016), we were able to merge

[3] The calculation of the HHI squares each market share (MS) and places a higher importance on those firms that have a larger market share. The formula is as follows: $HHI = MS_1^2 + MS_2^2 + MS_3^2 + \ldots + MSn^2$. The HHI ranges from zero (Perfect competition) to unity (Monopoly).

chemical releases data with national industry concentration ratios drawn from the NBER database and the Census of Manufacturers (only for the year 2012) since each facility must indicate the primary operated industry.

The starting date for the study was dictated by data availability, while 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. Table 1 depicts the descriptive statistics for our sample variables. For the sample of facilities, the level of toxic releases averages 461,833 pounds (or 10.12 pounds in logged values). Similarly, the level of market concentration (measured by the four concentration ratios) ranges from 38 to 80. This masks a wide disparity across facilities and across time. It is also worth mentioning that the starting year of our sample (1987) where the TRI was reported, the median facility released 50,498 pounds of toxic chemicals. By 2012, median facility releases had fallen to below 16,858 pounds, a roughly 67% reduction.

<Insert Table 1 about here>

3.1 Preliminary Testing for Cross-Section Dependence and Unit Roots

One of the additional complications that arise when dealing with panel data compared to the pure time-series case, is the possibility that the variables or the random disturbances are correlated across the panel dimension. The early literature on unit root and cointegration tests adopted the assumption of no cross-sectional dependence (Pesaran, 2015). However, it is common for macro-level data to violate this assumption which will result in low power and size distortions of tests that assume cross-section independence. For example, cross-section dependence in our data may arise due to common unobserved effects due to changes in federal legislation. Therefore, before proceeding to unit root and cointegration tests we test for cross-section dependence (Halkos and Polemis, 2017). We use the cross-section dependence tests proposed by Breusch and Pagan (1980) and Pesaran (2004).[4]

The tests are based on the estimation of the linear panel model of the form:

$$y_{it} = \alpha_i + \beta_i' x_{it} + u_{it}, \quad i = 1,..N; T = 1,..T \tag{1}$$

where T and N are the time and panel dimensions respectively, α_i the provincial-specific intercept, and x_{it} a $kx1$ vector of regressors, and u_{it} the random disturbance term. The null hypothesis in both tests assumes the existence of cross-section correlation: $Cov(u_{it}, u_{jt}) = 0$ for all t and for all $i \neq j$. This is tested against the alternative hypothesis that $Cov(u_{it}, u_{jt}) \neq 0$ for at least one pair of i and j. The Breusch and Pagan (1980) and Pesaran (2004) tests are a type of Lagrange-Multiplier test that is based on the errors obtained from estimating Equation 1 by the OLS method. Both tests strongly reject the null hypothesis of cross-section independence (P-value = 0.000) for all the models, providing evidence of cross-sectional dependence in the data given the statistical significance of the CD statistic (see Table 2). In light of this evidence we proceed to test for unit roots using tests that are robust to cross-section dependence.

<Insert Table 2 about here>

To examine the stationarity properties of the variables in our models we use the "*second generation*" unit root tests for panel-data proposed by Breitung and Das (2005) and Pesaran (2007) that allow for cross-section dependence. Both tests are based on OLS regressions; however the Breitung and Das approach breaks down if it is assumed that cross-correlation is due to common factors while the Pesaran (2007) test, denoted as *CIPS*, remains valid (Halkos and Polemis, 2017). The test results

[4] Since the cross-section dependence tests demonstrate that the error terms are correlated across facilities, it is worth mentioning that this issue could be alternatively addressed with the estimation of clustered standard errors, on the condition that the regimes have already been determined.

suggest that all the sample variables are stationary I(0).[5] Hence we proceed with testing for the presence of thresholds.

<Insert Table 3 about here>

4. Econometric framework

We proceed with the estimation of the threshold regression model, where the concentration ratio (CR4) is used as the sorting (threshold) variable that classifies the facilities in a competitive and a non-competitive industry regime. Hansen (1999; 2000) provides an estimation method based on a Concentrated Least Squares (CLS) procedure and he obtains the properties of the threshold and slope parameter estimators. In other words, the approach that we employ here does not rely on a known threshold parameter, but one that needs to be estimated along-side the other unknown parameters of the model. However, the method is based on first testing for the presence of a threshold effect. Once we reject the null of no threshold(s) we proceed in the estimation of the model that includes the estimation of the threshold(s) and allows for the sample split. The technique is based on a CLS method that splits the model into the two regimes, whereby there is a full interaction of all the variables with the (estimated) threshold.

We proceed to test for the presence of a (significant) threshold that allows for the comparison between the TR model and the simple linear benchmark without a threshold. It is worth noting that the threshold parameter is not identified under the null hypothesis of no threshold and usual test statistics have non-standard distributions. For that reason, Hansen (1999, 2000) suggests a bootstrap methodology

[5] In accordance with the LM-statistic of Breusch and Pagan (1980), we can reject cross-sectional independence among the sample variables (with p-values virtually equal to zero).

based on the utilization of a heteroskedasticity consistent Lagrange Multiplier (LM) bootstrap procedure to test H_0 of a linear formulation against a threshold formulation.

We argue though that an alternative, but less sophisticated approach to investigating non-linear effects would be to simply use higher order polynomial regressors of market concentration (squared, cubed terms, etc) instead of a panel threshold model. One could also resort alternatively to a semiparametric specification using local smoothers or splines/series to capture possible turning points. However such methods involve bandwidth choices and they do not lend themselves to estimating sharp turning points/thresholds as it is the case in the threshold model that we adopt in a fully interactive way (Kourtelos et al, 2016). Moreover, one important advantage of this methodology is that it avoids the ad hoc, subjective pre-selection of threshold values which has been a major critique of previous studies (Christie, 2014). In contrast to a simple case where the sample is split according to a known pre-assigned threshold value, the method that we use first tests for the presence of such a threshold and then estimates it (see for example Hansen, 2000; Caner and Hansen, 2004 and Kourtellos *et al,* 2016). In principle, one can test for additional sample splits, something that we did and we were able to detect. Based on the above, our threshold model takes the following form:

$$Y_{ijt} = X_{it}b + X_{it}I(z < \gamma)d + n_i + v_t + \varepsilon_{it} \tag{2}$$

where subscripts j = 1,...v, denote the facility (plant) that generates the chemical releases, i = 1, ..., N represent the six-digit code industry and t = 1, … , T indexes the time. n_i is the firm-specific fixed effect that control for differences across facilities such as technological innovations and chemicals used in the production process, capturing individual heterogeneity. We also include the relevant year (time) fixed effect (v_t) which captures the co movement of the series due to external shocks

(Asimakopoulos and Karavias, 2016). Y_{ijt} denotes the dependent variable (LREL). z is the vector of threshold variables namely $CR4_{it}$ and HHI_{it}.[6] In addition, X_{it} is the vector of exogenous control variables (LSHIP, LVADD, LEMP, LINVEST, LCAP) where slope coefficients are assumed to be regime independent. $I(\cdot)$ is the indicator function taking the value 1 when the condition in the parenthesis is satisfied and 0 otherwise which represents the regime defined by each threshold variable (CR4 and HHI respectively), and the threshold value γ that needs to be estimated within the model. Finally ε_{it} denotes the idiosyncratic i.i.d error term.

We complement the threshold model with a benchmark linear analysis in order to draw sharp differences between these results and the traditional benchmark linear specifications. We provide below the general exposition of the two linear models accounting for the presence of CR4 (Model I) and HHI (Model II) respectively[7]

$$\ln REL_{ijt} = a_0 + a_1 \ln SHIP_{it} + a_2 \ln CR4_{it} + a_3 \ln VADD_{it} + a_4 \ln EMP_{it} + a_5 \ln INVEST_{it} + a_6 \ln CAP_{it} + n_i + u_t + \varepsilon_{it} \quad (3)$$

$$\ln REL_{ijt} = a_0 + a_1 \ln SHIP_{it} + a_2 \ln HHI_{it} + a_3 \ln VADD_{it} + a_4 \ln EMP_{it} + a_5 \ln INVEST_{it} + a_6 \ln CAP_{it} + n_i + u_t + \varepsilon_{it} \quad (4)$$

The interpretation of the variables comes as follows. $\ln REL_{ijt}$, denotes the logged total (on-site and off-site) chemical releases emitted by facility j in industry i across the year t. $\ln SHIP_{it}$ is the logged value of shipments as a proxy for market size for industry i during year t. $CR4_{it}$, is one of our four concentration ratios of market structure allowing for certain cyclical behaviour (nonlinearities) in the impact of the covariates on the dependent variable.[8] $\ln VADD_{it}$, is the total value added for industry i during year t as a proxy for industry output expressed in natural logarithm. $\ln EMP_{it}$, is the logged value of total employment for industry i during year t as a proxy for

[6] We have also estimated our threshold model with the inclusion of the other three market concentration measures (i.e CR8, CR20 and CR50) as a robustness check and the results do not change significantly. The results are available upon request.

[7] The results do not drastically change if we estimate a semi logged model similar to Simon and Prince (2016).

[8] For the use of concentration ratios see also Polemis and Stengos (2015).

labour. $\ln INVEST_{it}$, denotes the total capital expenditure for industry i during year t as a proxy for capital, and subsequently $\ln CAP_{it}$, is the total real capital stock for industry i during year t as a proxy for intermediate inputs. Moreover, n_i is the unit-specific residual that differs between sectors but remains constant for any particular sector (unobserved sector level effect); while u_t captures the time effect and therefore differs across years but is constant for all sectors in a particular year. Finally ε_{it} denotes the error term.

5. Results and discussion

In this section, we present the results of the threshold fixed effects model along with the benchmark linear specification for each of the two alternative measures of market concentration (threshold variables). In addition, we offer a comparative discussion between the threshold effects and the static panel fixed effects linear specification benchmark models.

5.1 Testing for thresholds and estimating the linear model

We carry out the first part of the empirical analysis by determining the number of thresholds. For this reason, Equation (2) is estimated by OLS, allowing for (sequentially) zero, one and two thresholds respectively. The test statistics LM_1 and LM_2, along with their bootstrap p-values, are shown in Table 4. Specifically, we find that the test for a single threshold LM_1 is highly significant in both models with a bootstrap p-value of 0.00. On the other hand, the test for a second threshold LM_2 is also highly statistically significant, with a bootstrap p-value for each of the two models (Model I and II) equal to 0.00 and 0.00 respectively. As a consequence, we infer that there are two thresholds in all of the regression relationships.

<Insert Table 4 about here>

The point estimates of the thresholds for the two models are also reported in the relevant table. The estimates range from 3.70 to 4.20 (Model I) and 6.00 to 6.50 (Model II). The estimated values for the second threshold split the sample into two regimes. The first regime captures the medium and high levels of concentration since it includes the facilities where the sum of the four largest companies are below the value of 67 (or 4.20 in logged levels), while in the second regime the firms are characterised by significantly high levels of market concentration since only four firms possess more than 67% of the market. Similar interpretation applies to the HHI.[9] As a consequence, the two classes of facilities indicated by the point estimates are those with medium-high and very high level of market concentration respectively. In other words, the existence of a second threshold classifies the industrial facilities into competitive and non-competitive conditions respectively.

After having estimated the appropriate number of thresholds, we proceed with the exposition of results generated from the benchmark linear specification that will be contrasted with the threshold model. From the following table, it is evident that nearly all of the variables are statistically significant in either of the two specifications (with or without the time effects). However, the relevant signs of most of the regressors entering the two models (Model I and Model II) differ drastically.

Specifically, there is evidence supporting the argument that the market concentration proxied by the lnCR4 is positively correlated with a higher level of pollution, in both models. Regarding Model II, our estimates are higher than the ones reported by Simon and Prince (2016), ranging from 0.323 to 0.346 compared with the estimated value of 0.107. We argue that each percentage-point reduction in the lnHHI results in more than three percent reduction in a facility's toxic releases. The range of

[9] The first (second) regime includes the facilities where the sum of the squared market shares of the fifty largest companies is below (above) the value of 665 (or 6.50 in logarithmic scale).

this discrepancy could be attributed to the different samples as well as the model specifications followed in two studies.

Similarly, market size (lnSHIP) increases the level of industrial pollution when lnCR4 and lnHHI are taken into account respectively. In addition, the estimates for the level of industry output (lnVADD) reveal a negative correlation with the level of toxic chemical releases (see Model I and II respectively). The adverse result is evident when intermediate inputs (lnCAP) are taken into account. More specifically, it seems that real capital stock is positively correlated with the level of industrial pollution emitted by the facilities. On the other hand, there is strong evidence that labour intensive facilities do not stimulate toxic chemical releases since the relevant estimates (lnEMP) although positive in most specifications are not statistically significant. The opposite result is evident when the level of capital expenditures (lnINVEST) interacts with industrial pollution. From the magnitude of the relevant elasticities, we argue that a 10 percentage point increase (decrease) in the level of capital expenditures will lead to a 7.5 percentage point increase (decrease) in the level of toxic chemical releases.

< Insert Table 5 about here >

However, as it will be shown below, the results of the benchmark static model compared with the threshold effects model that we use in the present study reveal significant differences in the interpretation of the key variable of interest (market concentration). This means that the benchmark model does not capture the nonlinear effects stemmed from the existence of a double threshold according to the bootstrapped P-values of the relevant LM tests (see Table 4). Therefore, the threshold model is better suited to assess these effects on chemical releases under two different regimes (competitive and non-competitive conditions).

5.2 *The threshold model*

The results for the empirical relation between the (logged) toxic chemical releases and its main drivers under the tow regimes (competitive and non-competitive conditions) are depicted in Table 6. When the level of the four largest industries in the sector (CR4) is taken into account as the threshold variable (Model I) it is evident that nearly all of the control variables are statistically significant and plausibly signed. (Model I)

< Insert Table 6 about here >

The main variable of interest is the level of market concentration measured by the four largest firms in terms of their market shares. Recall, that when entered linearly, the coefficient is positive and statistically significant at the 1% level indicating that a one-percentage point decrease in market concentration reduces the level of industrial pollution by nearly 0.7 of a percentage point (see Table 5). On the other hand, the results for the non-linear model with a (double) threshold on market concentration at 67 percent, do suggest a strong non-linear relationship between competition and pollution. The point estimates suggest that the level of concentration (competition) is positively (negatively) related to the level of toxic chemical releases when time dummies are taken into account (see columns 1 and 2). However, it is evident that the CR4 index is more important in the sample below the threshold (competitive regime) since the relevant coefficient (2.336) is highly statistically significant. This means that a 10% decrease in the level of market concentration leads to a 23% decrease in the total chemical releases. This finding concurs that for already competitive sectors the level of market concentration does affect industrial output and subsequently the level of toxic releases emitted in the atmosphere. These results are in alignment with existing studies (Farber and Martin, 1986; Simon and Prince, 2016) where it is supported that competition effect would tend to lower pollution per facility.

Notably, the other control variables have the expected signs and are all statistically significant in both models (with and without time effects). More specifically, the level of industrial output proxied by the logged value added (lnVADD) has a negative impact on industrial pollution, indicating the presence of a strong technological effect since industrial facilities operating to high value added sectors are more prone to undertake actions limiting chemical releases (i.e energy conservation, waste management, etc). The magnitude of the estimates range from -2.037 (column 6) to -2.715 (column 1) indicating that a one percentage percent increase (decrease) in the level of industrial output (value added) results in approximately 2.5 percent decrease (increase) of the toxic chemical releases.

It is noteworthy that the aforementioned estimates do not vary substantially in their magnitude under the two different regimes. On the contrary the market size as expressed by the (logged) value of shipments (lnSHIP) varies considerably when the threshold value is taken into account. Specifically, for observations falling in the high (non-competitive) regime, further increases in the level of market size increase toxic chemical releases by 2.5 percentage points. This is contrasted against increases in the level of value added below the threshold value which displays a less direct effect on pollution. While the coefficients in both models (with and without time dummies) remain positive, are small in their magnitude and statistically significant not exceeding the value of 0.85. The relevant magnitude of the estimates although smaller than their counterparts in the model above the threshold (see columns 2 and 6) they are significantly higher than the ones reported by the existing literature (see Simon and Prince, 2016). The discrepancy could be justified by the fact that we uncovered a non-linear relationship between competition and industrial pollution.

Also as expected, increases in employment (lnEMP) reduce toxic pollution, but the estimated coefficients are small and only significant at the model above the threshold (-0.411 and -0.425 respectively). The magnitude of the relevant estimates denotes that for the non-competitive regime a 10 percent increase (decrease) of the level of employment in the industry would tend to lower (increase) pollution per facility by nearly 4 percent. This could be attributed to the fact that an increase in the level of employment would lower the capital to labour ratio (K/L) and hence the level of industrial emissions.

The coefficient on the level of total capital expenditure (lnINVEST) is positive and highly statistically significant in all of the specifications. For the model below the threshold (see columns 1 and 5), it is evident that a one-percentage point increase in capital investment can stimulate toxic releases by nearly 0.44 of a percentage point on average in each of the two specifications (with and without time effects). The rate of change is larger in its magnitude when we account for the model above the threshold (non-competitive) in which elasticities are equal to 0.527 and 0.492 respectively (see columns 2 and 6). These findings compared with the previous ones indicate that industrial pollution is evident (hidden) in capital (labour) intensive sectors. Lastly, the positive effect of intermediate inputs proxied by the total real capital stock (lnCAP) on pollution is evident in all of the specifications (see columns 1,2,5 and 6).

The discussion now turns to the alterative measure of market concentration namely the logged value of the Hirschman-Herfindahl index (lnHHI). Although difficult in its computation, the HHI provides a better measure of market concentration since it takes into account all the market shares of the firms in an industry (here the first fifty firms) compared to the concentration ratio of the four largest firms (Cabral, 2017).

Table 6 depicts the results for the empirical relation between market concentration (lnHHI) and the other covariates with the level of toxic releases under a competitive and non-competitive regime. According to the relevant table, nearly all of the main covariates are statistically significant and plausibly signed (Model II). Our key variable of interest is the level of market concentration (lnHHI). In this case, the impact of concentration on industrial pollution alternates its sign depending on the different competitive regime.

More specifically, the relationship between competition and toxic releases is negative (positive) when the threshold is high (low). This means that for observations falling into low regime (competitive) market concentration induces firms to increase output and hence the total level of pollution highlighting a positive net effect of competition, while the opposite holds for the high regime (non-competitive). This finding traces out the existence of an inverted "*V-shaped*" relationship between market concentration and industrial pollution at facility level.[10] More specifically, we are the first to uncover a non-linear statistically significant relationship between competition and industrial toxic releases for both above and below the threshold (665 units or 6.5 in logarithmic scale).

In particular, when the market concentration of the average facility is below the threshold, a one percent increase in the level of competition will reduce toxic emissions by 0.81 and 1.15 percent respectively (see columns 3 and 7). In this case we are on the upward slopping part of the curve. However, if the average facility is above the threshold then a one percent decrease in the level of competition will result in an increase of toxic releases by 0.52 and 0.65 percent respectively in both specifications (with and without time effects). This means that we are on the downward sloping part

[10] Since competition is the inverse of market concentration, we can also argue that there is a "*V*" shape relationship between competition and industrial pollution at plant level (facility).

of the concentration-pollution curve. As a consequence, the impact of competition on industrial pollution is larger quantitatively when it is below the estimated threshold.

Lastly, regarding the remaining variables we find that the estimated value of shipments is positive and statistically significant in both regimes ranging from 0.550 to 0.954. Similarly to Model I, the magnitude of this variable is larger when the observations fall above than below the threshold (0.929 compared to 0.595 when time dummies exist). The level of value added and the employment are negatively correlated with the level of industrial pollution, while capital expenditures reveal a strong positive effect on toxic releases. Finally, the coefficient of capital stock is around to unity on average when the observations are classified above and below the threshold.

6. Robustness checks

In order to check for the robustness of our findings, we re-estimate our basic linear model which is accordingly adjusted for the presence of three distinct concentration variables namely CR8, CR20 and CR50 respectively. These structural indicators capture the impact of the eight, twenty and fifty largest firms (measured on a 1-100 scale) in the industry respectively.

The empirical results when different aspects of market power are taken into account do not reveal significant differences regarding the competition variables and the set of the other covariates including the interaction terms. It is worth mentioning that, these interaction terms completely change the meaning of the coefficient on concentration. In other words the latter indicates the marginal effect of concentration when all the other RHS variables (that are interacted with concentration) are equal to zero (i.e when the logged values of these variables are equal to one). Despite the presence of so many interaction terms the empirical findings do not reveal significant discrepancies between the linear specifications and the TR model as already examined implying that the results are rather robust.

Nearly in all of the specifications, the control variables are statistically significant with the appropriate signs (see Table I in the Appendix). More specifically, when entered linearly, the estimate of the market concentration in all of the specifications (with and without time effects) is positive and statistically significant at the 1% level. It is noteworthy that the relevant magnitude ranges from 0.996 to 3.137 In other words a ten-percentage point increase in market concentration increases toxic chemical releases by 9.6% and 31.3% respectively. As a consequence, the negative relationship between the level of competition and the pollution emitted from industrial facilities seems to be in alignment with the possibly positive effect on abatement (Simon and Prince, 2016). In contrast to other studies, the possibility of a non-linear effect is well captured by the inclusion of the cross terms. More specifically, nearly all of the interaction terms in the three models are statistically significant denoting the existence of a possible non-linear relationship between the level of market power (and hence competition) and environmental damage.

Lastly, the inclusion of market concentration as an indicator of SMP might raise a possible endogeneity issue. Knowledge of the actual causality direction between market concentration and industrial pollution has important implications for modeling suitable environmental policies. Specifically, if the causality runs from market concentration to pollution, then environmental policies for combating toxic emissions may not affect the level of competition in the industry. On the other hand, if the causality is reversed, then environmental policies aimed at restricting industrial output and thus emissions may negatively affect the level of market structure by distorting effective competition. To tackle the presence of a possible endogeneity in the concentration variables, we have also used the lagged CR4 and HHI as regressors and our results remained fairly robust to whether we used current or lagged values of market concentration. All in all, we feel that the issue of endogeneity is not as severe in our case.[11]

7. Conclusions and policy implications

Toxic chemical prevention is emphasized in the U.S. environmental agenda as one of the primary means of industrial pollution abatement. The latter however, has been thoroughly investigated at an industry level neglecting the role of market structure and effective competition. In this study, we use a unique data set at the plant level comprising by thousands of industrial facilities dispersed among the US states over the period 1987-2012, in order to investigate the effects of industrial pollution prevention activities on toxic chemical releases under the presence of two market regimes (competitive and non competitive conditions).

For this reason, we utilised for the first time a static panel threshold model which allows for the presence of non-linear effects. The methodology applied supports new empirical findings that are of interest to policy makers and government officials regarding the non-linear nature of pollution. Moreover, our empirics cast doubt on the existence of a unilateral positive or negative effect of competition on pollution since we claim that a rather mixed (non-linear) effect prevails.

We uncover that the non-linear relationship is statistically significant above and below the estimated (double) threshold value, even after allowing for alternative specifications of market concentration. Our empirical findings do indicate that on average, each percentage-point reduction (increase) in the HHI results in a nearly one percent reduction (increase) in a facility's toxic releases when we are below the threshold level. On the other side of the curve, a ten percent reduction (increase) in the

[11] To preserve space, the results are available from the authors upon request.

concentration index induces an approximately six percent increase (reduction) in the level of pollution.

Taken together, these sets of findings fully justify the existence of an inverted "V-*shaped*" curve linking concentration and industrial pollution. This relationship provides new insights into the environmental policy design toward releases abatement since the policy makers 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 (threshold) of 665 units approximately indicates a negative effect on facilities' emissions levels whereas a decreasing line indicates a positive effect. Moreover, our models concur that the results remained robust under different specifications not driven by endogeneity.

The empirical findings indicate that when concentration level increases up to that point industries' toxic releases levels are also increasing. However after that estimated peak ("*turning point''*) it is evident that the regression line slightly decreases henceforth, revealing a negative effect of competition on environmental degradation. In other words, within this interval, the logged level of concentration has a positive impact on environmental pollution (decreasing part of the curve) creating an inverted "V-*shaped*" curve. Lieb (2003) asserts that the upturn of an inverted "U-*shaped*" Environmental Kuznets Curve (EKC) may be justified by the achievement of the internalization of the pollution externality on top of that the control chances are exhausted. On the contrary, the declining part of the curve may be because of a shock.

These set of empirical findings could be important for policy makers, academic researchers and practitioners. More specifically, they call for the need to strengthen the effectiveness of ecological-friendly policies by taking into consideration the market structure and the subsequent level of competition in an industry in order to drastically abate chemical pollution. Specifically, policy makers and government officials have to stimulate investments in value added sectors (i.e energy sector) and more likely to promote the use of renewable energy sources. This can be accompanied by more financial resources for research and development and more cost effective mitigation methods.

Acknowledgments

We are grateful to Timothy Antisdel from the Environmental Protection Agency (EPA) for his valuable assistance in collecting the data. We would also like to thank Thanos Pantos from the University of Piraeus for his help in processing and organizing the dataset. The usual disclaimer applies.

References

Agarwal, N., Banternghansa, C., and Bui, L. 2010. Toxic exposure in America: Estimating fetal and infant health outcomes from 14 years of TRI reporting. *Journal of Health Economics*, 29: 557-574.

Akkemik, K. Ali., and Oğuz Fuat. (2011). "Regulation, Efficiency and Equilibrium: A General Equilibrium Analysis of Liberalization in the Turkish Electricity Market," *Energy* 36(5): 3282-3292.

Antweiler, W., Copeland, B., and Taylor, S. (2001). Is Free Trade Good for the Environment? *American Economic Review*, 91: 877-908.

Asimakopoulos, S., and Karavias, Y., (2016). The impact of government size on economic growth: A threshold analysis. *Economics Letters*, 139: 65-68.

Bennett, M., Pierce, L., Snyder, J., and Toffel, M. (2013). Customer-Driven Misconduct: How Competition Corrupts Business Practices. *Management Science*, 59: 1725-1742.

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

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

Breitung., J., Das, S (2005). Panel unit root tests under cross-sectional dependence, *Statistica Neerlandica*, Netherlands Society for Statistics and Operations Research, 59(4): 414-433.

Breusch T.S and Pagan A.R., (1980). The Lagrange Multiplier Test and its Applications to Model Specification in Econometrics. *The Review of Economic Studies* 47, (1), 239-253.

Cabral, L (2017) Introduction to Industrial Organization, Second Edition. The MIT Press, London.

Caner, M., and Hansen B. (2004). Instrumental Variable Estimation of a Threshold Model. Econometric Theory 20: 813–843.

Chiara D.B, and Florio M. (2012). Electricity Investment: An Evaluation of the New British Energy Policy and its Implications for the European Union. Department of Economics, Management and Quantitative Methods, University of Milan, Working Paper no 2012-15, April.

Christie, T. (2014). The effect of government spending on economic growth: Testing the non-linear hypothesis. *Bulletin of Economic Research*, 66(2), 183-204.

Cole, M.A., Elliott, R.J.R and Khemmarat K. (2013). Local exposure to toxic releases: Examining the role of ethnic fractionalization and polarization. *Ecological Economics* 93: 249-259.

Currie, J., and Schmieder, J. (2008). Fetal Exposure to Toxic Releases and Infant Health. NBER Working Paper 14352.

Currie, J., Neidell, M., and Schmieder, J. (2009). Air Pollution and Infant Health: Lessons from New Jersey. Journal of Health Economics, 28: 688-703.

Environmental Protection Agency (2017). TRI National Analysis 2015. Available at http://www.epa.gov/trinationalanalysis.

Farber, S., and Martin, R. (1986). Market Structure and Pollution Control under Imperfect Surveillance. *Journal of Industrial Economics*, 35: 147-160.

Fiorio, C., and Florio. M. (2013). Electricity Prices and Public Ownership: Evidence From the EU15 Over Thirty Years. *Energy Economics* 39: 222–232.

Fowlie, M (2009). Incomplete Environmental Regulation, Imperfect Competition, and Emissions Leakage. *American Economic Journal: Economic Policy,* 1:2, 72–112.

Halkos, G and Polemis M. (2017). Does financial development affect environmental degradation? Evidence from the OECD countries? *Business Strategy and the Environment* DOI: 10.1002/bse.1976

Halkos, G., and Papageorgiou, G. (2016). Optimal abatement and taxation for internalizing externalities: A dynamic game with feedback strategies, MPRA Paper 74467, University Library of Munich, Germany.

Hansen, B.E. (1999). Threshold Effects in Non-Dynamic Panels: Estimation, Testing, and Inference. *Journal of Econometrics* 93: 345–368.

Hansen, B.E. (2000). Sample Splitting and Threshold Estimation. *Econometrica*, 68 (3) 575-603.

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.

Kourtellos, A, Stengos, T., Chih T.M. (2016). Structural Threshold Regression, *Econometric Theory*, 32(04): 827-860.

Leibenstein, Harvey. (1966). Allocative Efficiency versus X-efficiency. *American Economic Review*, 56(3): 392–415.

Levinson, A. (2015). A Direct Estimate of the Technique Effect: Changes in the Pollution Intensity of US Manufacturing, 1990-2008. *Journal of the Association of Environmental and Resource Economists*, 2(1): 43-56.

Levinson, A. (2009). Technology, International Trade, and Pollution from US Manufacturing, *American Economic Review*. 99(5): 2177-2192.

Lieb C.M. (2003). The Environmental Kuznets Curve: A survey of the empirical evidence and the possible causes. University of Heidelberg Depart of Economics, Discussion Paper Series No 391.

Mansur, E.T (2007). Prices vs. Quantities: Environmental Regulation and Imperfect Competition, NBER Working Papers 13510.

Pesaran, M.H., (2004). General Diagnostic Tests for Cross Section Dependence in Panels, Cambridge Working Papers in Economics 0435.

Pesaran, M.H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics*, 22(2): 265-312.

Pesaran, M.H. (2015). Time Series and Panel Data Econometrics, OUP Catalogue, Oxford University Press.

Polemis, M., Stengos T, (2015). Does market structure affect labour productivity and wages? Evidence from a smooth coefficient semiparametric panel model, *Economics Letters*, 137(C): 182-186.

Polemis, M., and Stengos T, (2017). Electricity Sector Performance: A Panel Threshold Analysis, *The Energy Journal*, 0 (Number 3).

Rzhetsky, A., Bagely, S. Wang, K., Lyttle, C., Cook, E., Altman, R., and Gibbons, R. 2014. Environmental and State-Level Regulatory Factors Affect the Incidence of Autism and Intellectual Disability. PLOS Computational Biology, 10: 1-11.

Shapiro, J.S., Walker, R. (2015). Why is Pollution from U.S. Manufacturing Declining? The Roles of Trade, Regulation, Productivity, and Preferences. NBER Working Papers 20879.

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*, http://dx.doi.org/10.1016/j.jeem.2016.03.001.

Zhang, Yin-Fang., David Parker., and Colin Kirkpatrick. (2005). Competition, Regulation and Privatization of Electricity Generation in Developing Countries: Does the Sequencing of the Reforms Matter? *Quarterly Review of Economics and Finance* 54(2–3): 358–379.

Zhang, Yin-Fang., David Parker., and Colin Kirkpatrick. (2008). Electricity Sector Reform in Developing Countries: An Econometric Assessment of the Effects of Privatization, Competition and Regulation, *Journal of Regulatory Economics* 33 (2): 159–178.

Zwickl, K., Ash M., Boyce J.K. (2014). Regional variation in environmental inequality: Industrial air toxics exposure in U.S. cities. Ecological Economics. 107: 494-509

List of Tables

Table 1: Descriptive statistics

Variable	Observations	Mean	Standard Deviation	Min	Max
lnREL	14,772	10.12	2.77	0.69	18.78
lnSHIP	14,292	16.18	1.49	4.38	20.18
CR4	14,299	38.10	20.27	2.00	100.00
CR8	14,297	51.57	22.59	2.00	100.00
CR20	14,266	67.51	22.30	5.00	100.00
CR50	14,230	80.07	20.86	8.00	100.00
HHI	11,878	617.8	591.00	15.00	2.99
lnCR4	14,271	3.47	0.62	0.69	4.60
lnCR8	14,297	3.82	0.55	0.69	4.60
lnCR20	14,225	4.14	0.44	1.60	4.60
lnCR50	14,117	4.34	0.34	2.08	4.60
lnHHI	11,729	5.97	1.09	2.71	8.00
lnEMP	11,906	3.67	0.91	0.10	6.25
lnVADD	11,906	8.58	1.11	4.69	11.61
lnINVEST	11,906	5.98	1.36	0.99	9.72
lnCAP	11,906	8.65	1.28	3.92	11.51

Note*:* All variables except for the concentration measures (i.e CR4, CR8, CR20, CR50 and HHI) are expressed in natural logarithms. lnREL denotes the logged total (on-site and off-site) chemical releases. lnSHIP is the logged value of shipments. lnCR4 is the logged concentration ratio of the four largest companies in the sector. lnCR8 is the logged concentration ratio of the eight largest companies in the sector. lnCR20 is the logged concentration ratio of the twenty largest companies in the sector lnCR50 is the logged concentration ratio of the fifty largest companies in the sector. lnHHI is the logged Hirschman-Herfindahl index of the fifty largest companies in the sector. lnVADD, is the natural logarithm of the total value added. LEMP denotes the logged value of total employment. lnINVEST stands for the total capital expenditure and lnCAP is the total real capital stock.

Table 2: Cross-section dependence test

Variable	CD test	P-value	Correlation	Absolute (correlation)
lnREL	133.05***	0.000	0.692	0.692
lnSHIP	143.34***	0.000	0.768	0.768
CR4	155.27***	0.000	0.831	0.831
CR8	157.21***	0.000	0.842	0.842
CR20	155.89***	0.000	0.836	0.836
CR50	159.52***	0.000	0.858	0.858
HHI	85.33***	0.000	0.468	0.535
lnCR4	157.56***	0.000	0.845	0.845
lnCR8	161.55***	0.000	0.866	0.866
lnCR20	163.82***	0.000	0.882	0.882
lnCR50	164.14***	0.000	0.886	0.886
lnHHI	103.39***	0.000	0.596	0.608
lnEMP	122.95***	0.000	0.806	0.806
lnVADD	127.28***	0.000	0.835	0.835
lnINVEST	127.54***	0.000	0.837	0.837
lnCAP	134.45***	0.000	0.883	0.883

Note*:* Under the null hypothesis of cross-sectional independence the CD statistic is distributed as a two-tailed standard normal. Results are based on the test of Pesaran (2004). The p-values are for a one-sided test based on the normal distribution. Correlation and Absolute (correlation) are the average (absolute) value of the off-diagonal elements of the cross-sectional correlation matrix of residuals. lnREL denotes the logged total (on-site and off-site) chemical releases. lnSHIP is the logged value of shipments. lnCR4 is the logged concentration ratio of the four largest companies in the sector. lnCR8 is the logged concentration ratio of the eight largest companies in the sector. lnCR20 is the logged concentration ratio of the twenty largest companies in the sector lnCR50 is the logged concentration ratio of the fifty largest companies in the sector. lnHHI is the logged Hirschman-Herfindahl index of the fifty largest companies in the sector. lnVADD, is the natural logarithm of the total value added. LEMP denotes the logged value of total employment. lnINVEST stands for the total capital expenditure and lnCAP is the total real capital stock. Significant at ***1%.

Table 3: Panel unit root tests

	Statistic	
Variable	**Breitung and Das (2005)**	**Pesaran (2006)**
lnREL	-23.0000***	-51.4745***
lnSHIP	-23.5125 ***	-46.2411***
lnCR4	-37.8671 ***	-54.1824 ***
lnCR8	-39.4864***	-54.5119 ***
lnCR20	-41.8846 ***	-55.9387 ***
lnCR50	-39.3244***	-57.3004***
lnHHI	-39.6548***	-53.0001***
lnEMP	-34.5414***	-54.7660***
lnVADD	-24.3578***	-48.9113***
lnINVEST	-21.8293***	-47.5499***
lnCAP	-23.6705***	-47.2386***

Note: The number of lags has been set to two according to BIC. The Augmented Dickey Fuller test is used rather than Phillips-Perron test. The null hypothesis assumes that the variable contains unit root. lnREL denotes the logged total (on-site and off-site) chemical releases. lnSHIP is the logged value of shipments. lnCR4 is the logged concentration ratio of the four largest companies in the sector. lnCR8 is the logged concentration ratio of the eight largest companies in the sector. lnCR20 is the logged concentration ratio of the twenty largest companies in the sector lnCR50 is the logged concentration ratio of the fifty largest companies in the sector. lnHHI is the logged Hirschman-Herfindahl index of the fifty largest companies in the sector. lnVADD, is the natural logarithm of the total value added. LEMP denotes the logged value of total employment. lnINVEST stands for the total capital expenditure and lnCAP is the total real capital stock. Significant at ***1%.

Table 4: Test for the existence of threshold(s)

Test for single threshold	***Model with lnCR4***	***Model with lnHHI***
Threshold estimate γ	3.700	6.00
LM_1	488.46***	402.76***
Bootstrap P-value	0.00	0.00
Test for double threshold	***Model with lnCR4***	***Model with lnHHI***
Threshold estimate γ	4.20	6.50
LM_2	137.28***	111.74***
Bootstrap P-value	0.00	0.00

Note: Test of Null of No Threshold Against Alternative of Threshold Allowing Heteroskedastic Errors (White Corrected). The trimming percentage is set to 0.15 and the Bootstrap replications are set to 1000. Significant at ***1%

Table 5: Linear Estimation Results

	With time effects		*Without time effects*	
Variable	**Model I**	**Model II**	**Model I**	**Model II**
lnSHIP	0.433*** (0.0756)	0.401*** (0.0765)	0.445*** (0.0754)	0.412*** (0.0764)
lnCR4	0.682*** (0.0923)	-	0.692*** (0.0446)	-
lnHHI	-	0.323*** (0.0533)	-	0.346*** (0.0258)
lnVADD	-1.451*** (0.0992)	-1.430*** (0.101)	-1.440*** (0.0991)	-1.411*** (0.101)
lnEMP	0.0922* (0.0552)	0.0710 (0.0563)	0.0772 (0.0547)	0.0430 (0.0557)
lnINVEST	0.491*** (0.0705)	0.519*** (0.0720)	0.485*** (0.0704)	0.506*** (0.0719)
lnCAP	0.750*** (0.0639)	0.745*** (0.0650)	0.745*** (0.0639)	0.747*** (0.0648)
Constant	3.462*** (0.688)	4.014*** (0.699)	3.337*** (0.686)	3.904*** (0.698)
Diagnostics				
Observations	11,878	11,607	11,878	11,607
Facilities	2,375.6	2,321.4	2,375.6	2,321.4
Years	6	6	6	6
R-squared (within)	0.1254	0.1326	0.1248	0.1233
F-statistic	170.11*** [0.000]	164.15*** [0.000]	282.14*** [0.000]	272.33*** [0.000]
HW	32.21*** [0.000]	31.26*** [0.000]	32.06*** [0.000]	31.73*** [0.000]
W-T	405.281*** [0.0000]	205.875*** [0.0001]	387.344*** [0.000]	194.132*** [0.0002]

Note: The market concentration variable is either the concentration ratio of the four largest companies (CR4) in the sector (Model I) or the Hirschman-Herfindahl index of the 50 largest companies (HHI) in the sector (Model II). lnREL denotes the logged total (on-site and off-site) chemical releases. lnSHIP is the logged value of shipments. lnCR4 is the logged concentration ratio of the four largest companies in the sector. lnHHI is the logged Hirschman-Herfindahl index of the fifty largest companies in the sector. lnVADD, is the natural logarithm of the total value added. LEMP denotes the logged value of total employment. lnINVEST stands for the total capital expenditure and lnCAP is the total real capital stock. W-T denotes the Wooldridge test for autocorrelation in panel data. The use of the fixed compared to random effects is justified after a Hausman test for each model. All models include state year fixed effects. Standard errors are in parentheses. To preserve space, we have deleted the results of the time dummies and their interactions with the threshold variables CR4 and HHI respectively. The numbers in square brackets are the p-values. Significant at ***1%, **5% and *10% respectively. HW is the Huber/White test for groupwise heteroscedasticity. W-T is the Wooldridge test for autocorrelation in panel data.

Table 6: Regression Estimates for the Double Threshold Model

	With time effects				*Without time effects*			
	Model I		**Model II**		**Model I**		**Model II**	
Variables	**(1) Threshold** $\hat{a}_1 \leq 4.20$	**(2) Threshold** $\hat{a}_2 > 4.20$	**(3) Threshold** $\hat{a}_1 \leq 6.5$	**(4) Threshold** $\hat{a}_2 > 6.5$	**(5) Threshold** $\hat{a}_1 \leq 4.20$	**(6) Threshold** $\hat{a}_2 > 4.20$	**(7) Threshold** $\hat{a}_1 \leq 6.5$	**(8) Threshold** $\hat{a}_2 > 6.5$
lnSHIP	0.855*** (0.136)	2.581*** (0.275)	0.595*** (0.143)	0.929*** (0.140)	2.515*** (0.274)	0.713*** (0.129)	0.550*** (0.141)	0.954*** (0.139)
lnVADD	-2.715*** (0.175)	-2.045*** (0.275)	-1.880*** (0.239)	-2.260*** (0.164)	-2.037*** (0.274)	-2.541*** (0.169)	-1.815*** (0.237)	-2.324*** (0.163)
lnEMP	-0.0190 (0.0872)	-0.411*** (0.117)	0.119 (0.156)	-0.317*** (0.0738)	-0.425*** (0.117)	-0.0713 (0.0860)	0.119 (0.156)	-0.289*** (0.0734)
lnINVEST	0.415*** (0.128)	0.527** (0.210)	0.716*** (0.157)	0.396*** (0.123)	0.492** (0.205)	0.462*** (0.126)	0.641*** (0.153)	0.432*** (0.122)
lnCAP	1.564*** (0.122)	0.732*** (0.212)	0.777*** (0.160)	1.367*** (0.116)	-0.633*** (0.209)	1.543*** (0.119)	0.832*** (0.158)	1.364*** (0.116)
lnCR4 $\leq$ 4.20	2.336** (1.029)	-	-	-	1.315*** (0.425)	-	-	-
4.20 < lnCR4	-	-2.556* (1.650)	-	-	-	-0.691*** (0.214)	-	-
lnHHI $\leq$ 6.50	-	-	0.813* (0.451)	-	-	-	1.150*** (0.421)	-
6.50 < lnHHI	-	-	-	-0.516* (0.291)	-	-	-	-0.652*** (0.119)
Constant	7.457*** (2.023)	-14.35*** (3.761)	-2.312 (3.016)	6.249*** (1.521)	-14.09*** (3.759)	7.725*** (2.020)	-1.768 (3.007)	6.358*** (1.515)
Diagnostics								
Observations	2,626	1,220	2,063	3,035	2,626	1,220	2,063	3,035
Facilities	525	244	413	607	525	244	413	607
Years	6	6	6	6	6	6	6	6
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time effects	Yes	Yes	Yes	Yes	No	No	No	No
R-squared (within)	0.212	0.122	0.200	0.181	0.206	0.113	0.198	0.176
F-statistic	70.21*** [0.000]	16.78*** [0.000]	73.09*** [0.000]	66.71*** [0.000]	112.93*** [0.0000]	25.78*** [0.0000]	84.41*** [0.0000]	107.82*** [0.0000]
HW	8.50 [0.1309]	8.98 [0.1098]	26.68*** [0.0001]	4.40 [0.4938]	8.50 [0.1306]	8.86 [0.1146]	26.27*** [0.0001]	4.46 [0.4855]
W-T	43.192*** [0.0028]	15.464** [0.0171]	41.762*** [0.0030]	208.276*** [0.0001]	44.196*** [0.0027]	16.707** [0.0150]	41.874*** [0.0029]	212.557*** [0.0000]

Note: lnREL denotes the logged total (on-site and off-site) chemical releases. lnSHIP is the logged value of shipments. lnCR4 is the logged concentration ratio of the four largest companies in the sector. lnHHI is the logged Hirschman-Herfindahl index of the fifty largest companies in the sector. lnVADD, is the natural logarithm of the total value added. LEMP denotes the logged value of total employment. lnINVEST stands for the total capital expenditure and lnCAP is the total real capital stock. All models include state year fixed effects.Significant at ***1%, **5% and *10% respectively. HW is the Huber/White test for groupwise heteroscedasticity. W-T is the Wooldridge test for autocorrelation in panel data.

Appendix

Table I: Alternative Linear Estimation Results

	With time effects			*Without time effects*		
Variable	**Model III**	**Model IV**	**Model V**	**Model III**	**Model IV**	**Model V**
lnSHIP	3.242^{***} (1.042)	4.670^{***} (1.116)	0.828^{**} (0.403)	-0.130 (0.746)	0.600 (0.793)	0.537 (0.492)
lnCR8	0.996^{***} (0.166)	-	-	0.466^{***} (0.137)	-	-
lnCR20	-	2.022^{***} (0.346)	-	-	1.330^{***} (0.332)	-
lnCR50	-	-	3.137^{***} (0.359)	-	-	1.960^{***} (0.507)
lnVADD	0.723 (2.903)	-2.928 (3.080)	1.245 (1.112)	9.491^{***} (2.127)	7.958^{***} (2.180)	4.284^{***} (1.374)
lnEMP	16.90^{***} (1.212)	17.71^{***} (1.269)	6.499^{***} (0.514)	12.98^{***} (0.901)	12.79^{***} (0.908)	8.734^{***} (0.632)
lnINVEST	-5.332^{***} (1.084)	-4.775^{***} (1.122)	-3.016^{***} (0.531)	-6.905^{***} (0.948)	-6.629^{***} (0.958)	-4.293^{***} (0.696)
lnCAP	-12.17^{***} (1.156)	-11.55^{***} (1.177)	-3.240^{***} (0.549)	-11.52^{***} (1.098)	-10.91^{***} (1.120)	-6.357^{***} (0.746)
lnCR8 × lnSHIP	-0.661^{***} (0.232)	-	-	0.0913 (0.167)	-	-
lnCR8 × lnVADD	-0.438 (0.649)	-	-	-2.408^{***} (0.474)	-	-
lnCR8 × lnEMP	-3.715^{***} (0.268)	-	-	-2.841^{***} (0.198)	-	-
lnCR8 × lnINVEST	1.280^{***} (0.245)	-	-	1.641^{***} (0.214)	-	-
lnCR8 × lnCAP	2.862^{***} (0.262)	-	-	2.719^{***} (0.248)	-	-
lnCR20 × lnSHIP	-	-0.980^{***} (0.249)	-	-	-0.0709 (0.176)	-
lnCR20 × lnVADD	-	0.375 (0.688)	-	-	-2.068^{***} (0.485)	-
lnCR20 × lnEMP	-	-3.890^{***} (0.281)	-	-	-2.795^{***} (0.199)	-

lnCR20×lnINVEST	-	1.161*** (0.253)	-	-	1.583*** (0.216)	-
lnCR20×lnCAP	-	2.716*** (0.267)	-	-	2.578*** (0.253)	-
lnCR50× lnSHIP	-	-	-0.0785 (0.101)	-	-	-0.0394 (0.116)
lnCR50×lnVADD	-	-	-0.696** (0.273)	-	-	-1.317*** (0.317)
lnCR50×lnEMP	-	-	-1.528*** (0.121)	-	-	-1.948*** (0.142)
lnCR50×lnINVEST	-	-	0.877*** (0.132)	-	-	1.108*** (0.162)
lnCR50×lnCAP	-	-	0.959*** (0.140)	-	-	1.617*** (0.177)
Constant	5.915*** (0.922)	2.946* (1.650)	-5.504*** (1.529)	5.761*** (0.920)	1.964 (1.643)	-1.873 (2.323)
Diagnostics						
Observations	11,756	11,718	11,756	11,756	11,718	11,718
Facilities	2,177	2,177	2,177	2,177	2,177	2,177
Years	6	6	6	6	6	6
R-squared (within)	0.164	0.165	0.158	0.162	0.161	0.162
F-statistic	153.74*** [0.000]	153.80*** [0.000]	146.29*** [0.000]	205.88*** [0.000]	204.35*** [0.0001]	205.91*** [0.000]
HW	31.04*** [0.000]	24.91*** [0.0001]	23.77*** [0.0006]	31.33*** [0.000]	26.18*** [0.0001]	23.77*** [0.0002]
W-T	238.951*** [0.0001]	170.125*** [0.0002]	199.510*** [0.0002]	202.977*** [0.0001]	162.978*** [0.0001]	170.438*** [0.0002]

Note: The market concentration variable is either the concentration ratio of the eight largest companies (CR8) in the sector (Model III), the twenty largest companies (CR20) in the sector (Model IV) or the fifty largest companies (CR50) in the sector (Model V). lnREL denotes the logged total (on-site and off-site) chemical releases. lnSHIP is the logged value of shipments. lnCR8 is the logged concentration ratio of the eight largest companies in the sector. lnCR20 is the logged concentration ratio of the twenty largest companies in the sector. lnCR50 is the logged concentration ratio of the fifty largest companies in the sector. lnVADD, is the natural logarithm of the total value added. LEMP denotes the logged value of total employment. lnINVEST stands for the total capital expenditure and lnCAP is the total real capital stock. W-T denotes the Wooldridge test for autocorrelation in panel data. The use of the fixed compared to random effects is justified after a Hausman test for each model. All models include state year fixed effects.Standard errors are in parentheses. To preserve space, we have deleted the results of the time dummies and their interactions with the threshold variables CR8, CR20 and CR50 respectively. The numbers in square brackets are the p-values. Significant at ***1%, **5% and *10% respectively. HW is the Huber/White test for groupwise heteroscedasticity. W-T is the Wooldridge test for autocorrelation in panel data.