THE UNIVERSITY OF ADELAIDE AUSTRALIA

SUB CRUCE LUMEN

The University of Adelaide
School of Economics

Research Paper No. 2010-06
May 2010

Pollution Abatement and Environmental Equity: A Dynamic Study

Nadezhda V. Baryshnikova

Pollution Abatement and Environmental Equity: A Dynamic Study

Nadezhda V. Baryshnikova *
School of Economics, University of Adelaide, Australia

We study pollution abatement and environmental equity in a dynamic panel model using data for 236 plants in the US pulp and paper industry observed over the period 1985–1997. We suggest a theoretical model for the plant manager who incorporates regulatory pressures into his calculations of optimal amount of pollution. Assuming actual pollution abatement exhibits a sluggish adjustment process, the theoretical model leads to an empirical AR(1) panel model. We estimate our model using GMM with both "temporally lagged" and "spatially lagged" instruments. We find that children, people below the poverty line and the smallest minority races are exposed to higher levels of pollution. Our findings show no evidence of environmental inequity against African-Americans or Hispanics, and find that the neighborhoods with a higher percentage of elderly population face significantly lower levels of pollution from the plants.

Keywords: pollution abatement, environmental equity, dynamic panel, instrumental variable, fixed effects.

*Address for correspondance: Nadezhda Baryshinkova, School of Economics, University of Adelaide, Adelaide, SA 5005; Email: nadezhda.baryshnikova@adelaide.edu.au; Phone: +61 8 8303 4821; Fax: +61 8 8223 1460

1. Introduction

The question of whether disadvantaged population groups, such as racial and socioeconomic minorities, are disproportionately exposed to pollution and whether demographic composition influences the amount of pollutants generated has been studied for roughly two decades. One of the goals of the US Environmental Protection Agency (EPA) is to ensure that "everyone enjoys the same degree of protection from environmental and health hazards and equal access to the decision-making process to have a healthy environment in which to live, learn, and work."[2] However, while environmental and political groups continue to lobby for "environmental justice," especially with respect to racial and ethnic minorities, the results of economic studies have been ambiguous. There has been no agreement on whether disadvantaged population groups are exposed to more pollution and if so, which of the racial, age, or socioeconomic minorities are more at risk.

This paper studies whether the racial, education, age, and income based population groups are disproportionately exposed to emissions of air pollutants. Specifically, we are interested in whether plants are allowed to emit more air pollutants if their neighborhood has a more disadvantaged population group. We seek to answer this question by examining air pollution emissions for a sample of 236 pulp and paper plants from 1985–1997 and by correlating them with changes in the characteristics of communities surrounding the plants. Our results show that neighborhoods with a higher percentage of children or a higher percentage of the population below the poverty line are exposed to higher emissions of air pollutants. Inequities were also found for "other" minority races (non-white, non-black, non-hispanic), and communities with low voter turnout or congressional members with low "environment" ratings based on the League of Conservation Voters scorecard.

The initial study in the literature was produced by the United Church of Christ's Commission for Racial Justice (1987). This descriptive study reported that the zip codes which had more pollution as measured by the presence of a treatment, storage, and disposal facility (TSDF), had a higher percentage of minorities (twice that of the areas without TSDFs). They also noted that the relationship between socioeconomic status variables and pollution were not as significant. Since this seminal study, the quantity and quality of the environmental equity studies has improved remarkably.

One area of improvement is the measurement of the dependent variable. Earlier papers use proximity to noxious facilities as a proxy for environmental risk, (e.g. Anderton et al. (1994), Been (1994), Boer et al. (1997), Oakes et al. (1996), Pollock and Vittas (1995)) whereas later studies use actual pollution emissions levels (see Brooks and Sethi (1997), Daniels and Friedman (1999), Gray and Shadbegian (2004), Morello-Frosch et al. (2004), Ringquist (1997)). This paper studies actual pollution levels cited at the plant, specifically emissions of particulate matters less than 10μm [PM10] and emissions of sulfur dioxide [SO_2]. These are common pollutants that are monitored and regulated by the EPA. We incorporate the pulp and paper industry dataset which is extended from Gray and Shadbegian (2004).

A second area of improvement has been the attempt to control for alternative explana-

[2] Environmental Protection Agency (EPA): http://www.epa.gov/environmentaljustice/

tions and addressing the inherent temporal dimension of environmental equity: assessing the "chicken and the egg" question with respect to risk exposure and community demographics. Been (1994) points out this endogeneity problem and resolves it by using pre-siting demographics. Ringquist (1997) uses a control variable approach by controlling for housing prices. Gray and Shadbegian (2004) use instrumental variables and control for alternative explanations.

Despite these improvements, there has been no agreement on the existence of demographic inequities. There is disagreement in literature reviews as well with respect to interpretation of results. Ringquist (2005) provides a helpful literature review on this subject and a meta-study to analyze why conclusions about the existence of environmental inequity differ across studies, in an attempt to extract a generalizable conclusion. He finds that the differences in the choice of pollution measurement, levels of aggregation, and control factors do not explain away the existence of inequity with respect to race. He also finds that there is, to a lesser degree, evidence to support inequity with respect to the poor.

This paper is the first to use the dynamic panel model in the field of environmental equity. We address the issues raised in previous studies by use of this new methodology. The first methodological problem it addresses is that the process of pollution abatement is a dynamic process. It is very hard to reduce pollution overnight since it frequently requires changes in equipment and integration with the existing production technology. Hence, pollution abatement may exhibit a sluggish adjustment phenomenon. Because the empirical literature has so far focused on cross-sectional results, the sluggish adjustment phenomenon of pollution abatement is not captured, nor is the delayed response to dynamic changes in the population demographic composition. For example, assume there is environmental inequity towards the poor. Consider a case where the poor population in a neighborhood decreased over the last period. There was high pollution, but because pollution adjusts sluggishly, it did not yet respond to the change in composition and remained high in the current period. Regressing current high pollution levels on current decreased percentage of the poor may result in a misleading conclusion that the poor are not more exposed to pollution when in fact they are. This paper deals with inertia in the environmental performance of plants using an instrumental variables partial adjustment model.

Another methodological issue addressed by this model is that regressing pollution amounts on demographic characteristics introduces an endogeneity problem. Demographic groups that appear to suffer from environmental inequity could instead have chosen to live near the pollution for social or economic reasons. This paper deals with this endogeneity problem through the use of spatially lagged instruments suggested by Gray and Shadbegian (2004) in their cross-sectional study and we extend this to a dynamic model. A third methodological problem comes from the difficulty in accounting for the unobserved heterogeneity of the firms' production technology and other firm-specific effects. Gray and Shadbegian (2004) used variables such as pulp capacity and paper capacity to control for plant-specific effects. Although it is better to try controlling for these effects than to leave them unaccounted for, some effects are unobservable. Instead, we rely on the methodological advantage of our dynamic panel model. That is, we can first difference out these fixed effects in a dynamic panel setting.

The rest of this paper is organized as follows: Section 2 proposes a dynamic model of pollution adjustment that allows for a lag in abatement. Section 3 describes the Census and the pulp and paper industry data along with the merging procedures. Section 4 describes the methodology and specifies the choice of instruments for estimating the model using Generalized Method of Moments (GMM). Section 5 discusses results. Section 6 concludes that environmental equity is not met with respect to certain demographic groups.

2. Theory: model of dynamic regulation

To model pollution regulation and compliance within the confines of economic theory, we propose a new approach. The approach taken by Gray and Shadbegian (2004)is to model the behavior of environmental regulators (such as the EPA and state level regulators) that maximize the total social benefits subject to the total social costs. The optimal level of pollution is reached when the marginal social benefit of pollution abatement equals the marginal social cost. Our approach is to construct a simple and intuitive dynamic model for the typical plant whose pollution is under environmental regulation. The plant manager is assumed to be a profit maximizing agent whose main disincentive for pollution comes from society through its pollution regulating agencies.

The pollution regulating agencies interact with plants through a variety of regulatory pressures. The manager of plant i, under the regulatory environment, incorporates regulatory pressure into his profit-maximizing calculations. When he does that, he arrives at an optimal level of pollution at time t, P^*_{it}. There is another important factor in determining P^*_{it}, the plant's existing production technology, out of which the pollution comes as a by-product. We assume P^*_{it} depends on the plant's production technology, T, and the regulatory pressures, R.

We further assume that plant i's production technology stays relatively fixed through time, denoted T_i. It is evident that a plant's pollution emissions depend on its technology, since pollution can be viewed as a by-product of the plant's production process. The assumption that plant technology stays constant through time seems justifiable: the typical production process is largely dependent on the capital equipment whose life-cycle tends to be long, factory buildings last a long time once built, and floor plans stay put once the production line is in place. It is particularly justifiable in our case since the pulp and paper industry we study is very capital-equipment intensive.

We model the regulatory pressure R in two parts: a relatively time-independent regulatory pressure R_i and a relatively time-dependent regulatory pressure R_{it}. R_{it} itself consists of two parts: an observable part O_{it} and an unobservable part U_{it}. Thus, the model can be represented as the following equation:

$$P^*_{it} = \beta_1 T_i + \beta_2 R_i + \beta_3 O_{it} + \beta_4 U_{it} \tag{1}$$

The variable R_i captures the possibility that the heterogeneity among plants may induce regulators to impose plant-specific regulatory pressures. A plant may be considered more politically important by the regulators for various reasons, such as its political visibility, its unionization, its voting district's pro-environmental voting records, etc. In general, both T_i and R_i encompass numerous aspects that are easy to list but hard to account for due to their qualitative and unobservable nature. Gray and Shadbegian (2004) attempted to control for at least some of these aspects, such as the pulp and paper capacity, return on assets, and Occupational Health and Safety Administration (OSHA) violations. Our dynamic panel setting allows us to difference these fixed effects away and focus on the main issue of environmental justice. The decomposition of R_{it} into an observable part O_{it} and an unobservable part U_{it} is done to facilitate our empirical study, which uses only observable data, such as population demographic characteristics. The unobservable data U_{it} will eventually be incorporated into the error term of our regression.

At time t, the plant manager will compare the actual level of pollution P_{it} with the optimal level of pollution P^*_{it}, then their difference will be the target amount of pollution to be abated for the period $t+1$, denoted as $Abate^*_{it+1}$ in the following equation:

$$Abate^*_{it+1} = P_{it} - P^*_{it} \tag{2}$$

We also make the sluggish adjustment assumption. That is, we assume that the actual pollution abatement for the $t+1$ period, $Abate_{it+1}$, is only a fraction γ of the theoretical pollution abatement $Abate^*_{it+1}$:

$$Abate_{it+1} = \gamma(P_{it} - P^*_{it}) \tag{3}$$

It is reasonable to assume that the adjustment process is sluggish over time. One possible scenario is that many pollution abatement projects require investments in pollution abatement capital, which takes time to plan and install. Another possible explanation is that it may take time and some learning by doing to fully incorporate pollution abatement processes into the existing production technology. In either of these scenarios, the pollution will be gradually reduced over time as the abatement projects become fully operational.

The actual pollution at time $t+1$ will be expressed as the difference between the actual pollution in the last period P_{it} and the actual pollution abatement for the $t+1$ period:

$$P_{it+1} = P_{it} - Abate_{it+1} \tag{4}$$

Combining equations (1), (3), and (4), we reach the following equation:

$$P_{it+1} = \gamma(\beta_1 T_i + \beta_2 R_i) + (1-\gamma)P_{it} + \gamma\beta_3 O_{it} + \gamma\beta_4 U_{it} \tag{5}$$

Notice the first term in equation (5) can be regarded as the unobservable fixed effect. Hence, we can take the standard step to difference out the fixed effect and arrive at our main equation:

$$\Delta P_{it+1} = (1-\gamma)\Delta P_{it} + \gamma\beta_3(\Delta O_{it}) + \gamma\beta_4(\Delta U_{it}) \tag{6}$$

A special case to equation (6) is when $\beta_3 = 0$. When none of the observable regulating variables such as population characteristics matter in deciding the optimal pollution, and assuming that the U_{it} term becomes part of the error term, equation (5) simplifies into the following:

$$P_{it+1} = F_i + (1-\gamma)P_{it} + \varepsilon_{it} \tag{7}$$

where $F_i = \gamma(\beta_1 T_i + \beta_2 R_i)$ is the fixed effect and ε_{it} is the error term. And similarly, equation (6) simplifies into the following AR(1) model:

$$\Delta P_{it+1} = (1-\gamma)\Delta P_{it} + \Delta\varepsilon_{it} \tag{8}$$

In our empirical analysis section, we use the above model to estimate the effect of regulatory pressure on two air pollutants: small particulates (PM10) and sulfur dioxide (SO_2) for the pulp and paper industry in the United States, observed between 1985 and 1997.

3. Data

In our empirical study, we use two data sets: the yearly plant-level pollution data that comes from the Gray and Shadbegian (2004) study of US pulp and paper mills, and the decade-level population demographic characteristics data (1970–2000) from the US Census. The first data set contains 306 plants from the pulp, paper, and cardboard industries (SICs 2611, 2621, and 2631) observed during the period 1985–1997. Due to the dynamic panel nature of our study, we require the plants to have at least three consecutive observations. We are left with 277 plants after this requirement. Merging with Census data reduces the usable number of plants to 236.[3] Due to taking the first difference to eliminate the fixed effects, we lose the first two years of observations. Thus, we are left with the period 1987–1997. In 1992, the Census of Manufacturing reported a total of 529 plants. It should be noted that the plants in our data set tend to be larger than the average plant in the industry as the result of EPA coverage.

The second data set used in our study is compiled from the 1970, 1980, 1990, and 2000 US Census of Population data for Census block groups.[4] It contains the demographic characteristics of the population within a 50-mile radius of each plant. To determine which Census block groups fall within the 50-mile radius of the plant, the distances are calculated between the plant and the centroid of each block group. Then, the values for these block groups are aggregated to determine the demographic characteristics for each plant's 50-mile radius neighborhood. For the purpose of constructing spatially-lagged instruments, the population demographic characteristics for the area between 50 and

[3]This paper does not include the analysis of predicted probability of plant closing since the number of plants that closed in the current dataset is very small (8 plants out of total of 306). The closed plants were not omitted from the study if they had at least 3 consecutive observations in the period of 1985-1997.

[4]This data has been complied in the Census-CD data sets prepared by Geolytics, Inc. and merged using the GIS.

100 miles from the plant (the "doughnut") are also constructed in the same fashion and included in this data set.

In our yearly plant-level pollution data set, we focus on the air pollution measures of small particulates (PM10) and sulfur dioxide (SO_2). These measures originally come from the Aerometric Information Retrieval System database for 1985–1990 and from the National Emissions Inventory for 1990–1997. Since there is a large difference in magnitude in the reported emissions among different plants, we take the natural logarithm of the pollution levels as our pollution measures.

To study the environmental justice issues, we merge our plant-level yearly pollution data with the Census decade-level data set. The demographic characteristics we consider represent different schemes, such as socioeconomic status, racial composition, and sensitivity of population towards pollution. To measure socioeconomic status, we use the variable (PPOOR), the percentage of the population below the poverty line. We also use the variable (PHSDROP), the percentage of the population who are high school dropouts. For racial background characteristics, we use the percentage of population that is African-American (PBLACK), Hispanic (PHISP) and non-white other than African-American and Hispanic (POTHER) using the Census method for racial identification. Children and the elderly are considered to be groups that are especially sensitive to air pollution. We compute the percentage of population under six years old (PKIDS) and the percentage of total population over the age of 65 (PELDERS). Population density of the area surrounding the plant is measured by (DENSITY). To control for the propensity to vote, we use the percentage of the population over 18 voting in the county in the previous presidential election (TURNOUT). However, political activity on its own may not be as important if the voters do not support environmental regulation. Hence, we interact the voter turnout with a measure of the area's support for environmental regulation, the state membership in 3 conservation groups in the late 1980's, per 1000 population (TURNOUT*CONVMEMB). We use state pro-environmental Congressional voting (ENVIRVOTE) to control for the area's propensity for environmental regulation and pressure to reduce pollution. The data comes from The League of Conservation Voters scorecard for environmental voting records of each member of congress. We use the yearly average score for all representatives of the state where the plant is located.

Ideally our demographic data should contain yearly characteristics instead of decade-level data. Unfortunately this yearly data is not available from the Census. However, given that population characteristics tend to change gradually and smoothly over the years, we can obtain reasonable estimates for the yearly population demographics using natural cubic spline interpolation. To perform the interpolation, we require data to be available for all four Census years from 1970 to 2000. We are left with 236 plants after this requirement.

The natural cubic spline interpolation uses a cubic polynomial for approximation, as shown in the following formula:

$$Y_i(t) = a_i + b_i t + c_i t^2 + d_i t^3 \tag{9}$$

where $Y_i(0) = y_i$ and $Y_i(1) = y_{i+1}$ for $i = 0, ..., n-1$. On top of the end condition, we require $Y_i'(1) = Y_{i+1}'(0)$ for $i = 0, ..., n-1$. The second derivatives are also set to

$Y_i''(1) = Y_{i+1}''(0)$ for $i = 0, ..., n-1$. In addition to that, $Y_0''(0) = Y_n''(1) = 0$. These conditions will solve uniquely for a_i, b_i, c_i, d_i.

Including the demographic characteristics can create an endogeneity problem; that is, some groups may come to live in polluted areas due to the cheaper housing rate or other social factors. To correct this problem, we use the spatially lagged instruments suggested by Gray and Shadbegian (2004). For each decade, we calculate each of the demographic variables for the plant neighborhood within a 50–100 mile radius of the plant. As the distance from the plant increases, the level of pollution decreases while the demographic characteristics in the general area remain similar.

Tables 1, 2 and 3 provide the descriptive statistics for each year for the log of each of the air pollutants and the demographic variables. The means of both pollutants have been declining over the years, but there is considerable variation in the emission variables across the plants.

4. Methodology

In this section we provide an overview of the econometric methodology used in our empirical analysis to study the possibility of demographic discrimination and test for it in various groups. The model we are using is a panel AR(1) model with fixed effects:[5]

$$P_{it} = \alpha_i + \rho P_{it-1} + \eta D_{it-1} + \varepsilon_{it} \tag{10}$$

where α_i represents the plant i's unobservable fixed effect, which is invariant across time, $\varepsilon_{it} \sim iid(0, \sigma_\varepsilon^2)$ is the unobserved error term, which could include the unobserved regulatory pressure as we have discussed in our theory section, $\rho = (1-\gamma)$, with γ being the abatement coefficient from our theory, and P_{it} is the measured pollution variable (PM10 or SO_2) of plant i at time t. D_{it} is a vector of demographic characteristics for plant i at time t.

This model is very good at capturing the heterogeneity among the plants via the fixed effect α_i. α_i represents the plant's specific characteristics, such as the plant-specific technology for each plant. These characteristics cannot be easily observed or quantified, hence, they cannot be controlled for in a cross-sectional model. The lagged AR(1) variable is used to represent the sluggish adjustment of the dependent variable, pollution.

To deal with the incidental parameter problem, we eliminate the unobserved fixed effect by taking the first difference and arrive at the equation[6]

$$\Delta P_{it} = \rho \Delta P_{it-1} + \eta \Delta D_{it-1} + \Delta\varepsilon_{it} \tag{11}$$

where $\Delta P_{it} = P_{it} - P_{it-1}$ represents the sluggish adjustment of the firm's pollution, $\rho = (1-\gamma)$ and $\eta = \beta_2\gamma$ according to our dynamic theoretical model. ΔD_{it-1} represents the yearly change in the demographic composition around the plant. A nice feature of equation (11) is that it has differenced out the fixed effects, and has eliminated the need to find the control variables and estimate them.

[5] This equation corresponds to equation (5) in our theoretical model of pollution abatement.

[6] This corresponds to equation (6) in the theory section.

To estimate equation (11), we use the Generalized Method of Moments (GMM) framework analogous to Arellano and Bond (1991). Our modification is that we use both temporally and spatially lagged variables, which must both be included in the instrumental variable matrix.

By using differencing to remove the fixed effects, we made the lagged regressor ΔP_{it} correlated with the error term in equation (11). To resolve this issue we instrument for the lagged variable ΔP_{it-1} with the optimal temporally-lagged instrumental variable matrix which is suggested by Arellano and Bond (1991). The instruments are all values of the levels P_{it} lagged two periods or more. It is worth noting that the strength of these instruments depends on the unknown parameter of interest, ρ. In particular, when the coefficient ρ is close to 1 these instruments are weak.[7] This would occur if γ is close to zero, implying that plant managers have little control in reducing pollution to the optimal levels they set.

Another set of instruments is needed to resolve the endogeneity problem in ΔD_{it-1}. Although temporally lagged instruments are convenient and readily available, they will be weak if the time series is too persistent. Another choice of instruments would be pre-siting demographics as suggested by Been (1994). Unfortunately, since many pulp and paper mills were built before 1960, detailed demographic data from before the mills were built is mostly unavailable. Instead, we use demographic variables from the 50-100 miles around the plant, $\Delta\widetilde{D}_{it-1}$, as instruments for ΔD_{it-1}, which were used by Gray and Shadbegian (2004). Hence, the instrumental variable matrix consists of both a temporally lagged part for pollution and a spatially lagged part for demographic variables:

$$Z_i = \begin{bmatrix} P_{i,1} & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & \vert & \Delta\widetilde{D}_{i3} \\ 0 & P_{i,1} & P_{i,2} & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & \vert & \Delta\widetilde{D}_{i4} \\ 0 & 0 & 0 & P_{i,1} & P_{i,2} & P_{i,3} & & 0 & & 0 & \vert & \Delta\widetilde{D}_{i5} \\ \vdots & & & & & & \ddots & & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & \ldots & P_{i,1} & \cdots & P_{i,11} & \vert & \Delta\widetilde{D}_{i11} \end{bmatrix}$$

5. Results

Table 4 reports the results of the first stage regression of demographic variables on all the instruments and exogenous variables. All regressions include time dummies. The spatially lagged demographic variables are highly significant with the p-values of 0 for all the regressions.

Table 5 reports the two-step GMM estimates with the standard errors in parentheses. The Sargan test statistic rejects the null hypothesis of serial autocorrelation in the errors. Both Wald test and Kleibergen test, which is robust to weak instruments, indicate the joint significance of the coefficients.

For the pollution variable PM10, the point estimate for the AR(1) coefficient is $\hat{\rho} = 0.552$. For SO_2 the point estimate for the AR(1) coefficient is $\hat{\rho} = 0.474$. Both AR(1) coefficients are significantly positive. From our theory, the AR(1) coefficient has an additional economic meaning: $\rho = 1 - \gamma$, where γ is the pollution adjustment coefficient.

[7]This is because when the lag coefficient is close to 1 the process becomes a random walk.

Specifically, for PM10, the AR(1) coefficient of 0.552 implies that the pollution adjustment rate is $\gamma = 1 - 0.552 \approx 45\%$. That is, the plants tend to complete 45% of the total targeted PM10 pollution abatement (which is the difference between the actual pollution and the optimal level of pollution) over the following year. Similarly, the plants tend to complete 53% of the targeted SO_2 pollution abatement over the following year. This result shows that for both pollution measures, PM10 and SO_2 in the pulp and paper industry, the pollution has been reducing steadily following the dynamic AR(1) process over the observed years.

The variable PKIDS is found to be significant for the PM10 emissions and insignificant for SO_2 emissions. Quantitatively, 1% increase in the percentage of children is associated with 0.54% higher pollution levels of PM10 and, hence, less reduction in pollution. For elders, the effect is opposite. Plants located in areas with a higher percentage of elderly emit significantly lower levels of SO_2. That is, a neighborhood with 1% higher elderly population tends to have 1.19% lower pollution levels of SO_2. It is interesting to note that although each population group is considered sensitive to pollution, their coefficients are of opposite sign. Population density around plants (DENSITY) is found to be insignificant.

We find that plants surrounded by higher percentage of people below the poverty line emit higher levels of pollution, although not significantly for the amounts of PM10 (a 1% increase in population below poverty is correlated with a 0.24% increase in pollution).

Hispanics and African-American population are found to be insignificant, with negative coefficients for all except African-Americans for the SO_2 levels of pollution. Other non-white population (POTHER) is found to be exposed to higher levels of pollution, although not significant for the SO_2 levels. This suggests that 1% increase in POTHER is associated with 0.056% higher PM10 emissions. HSDROP is found to be insignificant in exposure to PM10 and significantly negative for SO_2 emissions. A 1% increase in the population who dropped out of high school is associated with 0.153% lower emissions of SO_2.

The propensity for collective action (TURNOUT) is significant for PM10 emissions and insignificant for SO_2 emissions. A 1% increase in voter turnout is correlated with 1.51% lower emissions. However, this effect is smaller when plants are located in an area with already high support for environmental regulation, measured by CONVMEMB. Also, plants located in the states that are environmentally strong, measured by the State Congressional pro-environmental voting (ENVIRVOTE), have significantly lower emissions of PM10.

6. Conclusion

In this paper we construct a dynamic model of pollution where the amount of current emissions depends on the emissions in the previous year and the demographic characteristics of the population in the plant's neighborhood. The model accounts for endogeneity in demographics and the slow process of pollution abatement. We take the data on the pulp and paper industry with pollution measurements of particulate matter of less than 10μm, PM10, and sulfur dioxide, SO_2, and merge the dataset with the demographic characteristics in the plant's neighborhood as was reported in the Census. The Census decade-level population data is interpolated using the natural cubic spline model to obtain yearly approximations. We use both "temporally lagged" and "spatially lagged" instruments and estimate a panel AR(1) model using two-step GMM.

Our results indicate that environmental inequity cannot be adequately explained away by the fact that certain demographics may be more inclined to seek out housing around plants. Even after controlling for endogeneity, plant-specific technology, political activism, and education we still observe environmental inequity with respect to the poor, the smallest minority races, and children. Our findings show no evidence of inequity with respect to African-Americans or Hispanics. Also, the neighborhoods with a higher percentage of elderly population face significantly lower levels of pollution from the plants.

We restricted our study to two types of pollutants from one industry in the US, so we must caution against generalizing the results. The pulp and paper industry is one of the largest polluting industries in the US, and the outputs studied here are common and dangerous pollutants. We cannot, however, claim that these results are generalizable to all types of pollution or to other industries. Further research could involve extending this model to other regulated polluting industries and other toxic releases.

Acknowledgements

Special thanks to Wayne Gray for providing the data and ongoing advice and support. Frank Kleibergen, Andrew Foster, Stephanie McWhinnie and John Herbst are gratefully acknowledged for helpful comments and guidance, as well as the seminar participants at Brown University, University of Connecticut, Melbourne University, Australian National University RSSS, University of New South Wales and the anonymous referees.

REFERENCES

Anderton, D., Anderson, A., Oakes, J.M., Fraser, M., 1994. Environmental equity: The demographics of dumping. Demography 31, 229-248.

Arellano, M., Bond, S., 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. The Review of Economics Studies 58, no. 2, 277–297.

Been, V., 1994. Locally Undesirable Land Uses in Minority Neighborhoods: Disproportionate Siting or Market Dynamics? The Yale Law Journal 103, no. 6, 1383-1422.

Boer, J.T., Pastor, M. Jr., Sadd J., Snyder L, 1997. Is there environmental racism? The demographics of hazardous waste in Los Angeles County. Social Science Quarterly 78, no. 4, 793–810.

Brooks, N., Sethi, R., 1997. The distribution of pollution: Community characteristics and exposure to air toxics. Journal of Environmental Economics and Management 32, 233–250.

Daniels, G., Friedman, S., 1999. Spatial inequality and the distribution of industrial toxic releases: Evidence from the 1990 TRI. Social Science Quarterly 80, no. 2, 244–262.

Gray, W., Shadbegian, R., 2004. Optimal pollution abatement - whose benefits matter, and how much? Journal of Environmental Economics and Management 47, 510–534.

Morello-Frosch, R., Pastor, M, Sadd, J., 2004. Waiting to inhale: The demographics of toxic air release facilities in 21st-century California. Social Science Quarterly 85, no. 2, 240–440.

Oakes, J.M., Anderton, D.L., Anderson, A.B., 1996. A longitudinal analysis of environmental equity in communities with hazardous waste facilities. Social Science Research 25, 125–148.

United Church of Christ (UCC) Commission for Racial Justice, 1987. Toxic wastes and race in the United States: A national report on the racial and socioeconomics characteristics of communities with hazardous waste sites, New York: Public Data Access, Inc.

Pollock, P.H., Vittas, M.E., 1995. Who bears the burdens of environmental pollution? Race, ethnicity, and environmental equity in Florida. Social Science Quarterly 76, no. 2, 294–310.

Ringquist E., 1997. Equity and the distribution of environmental risk: The case of TRI facilities. Social Science Quarterly 78, no. 4, 811–829.

Ringquist E., 2005. Assessing evidence of environmental inequities: A meta-analysis. Journal of Policy Analysis and Management 24, no. 2, 223–247.

Table 1: Summary Statistics for Pollution and Race Variables.

Year	**Obs**	**PM10**	**SO_2**	**PBLACK**	**PHISP**	**POTHER**
1987	118	5.05 (1.96)	6.78 (1.64)	13.04 (12.20)	1.86 (2.37)	0.56 (1.36)
1988	133	4.89 (2.15)	6.70 (1.66)	12.94 (11.96)	1.94 (2.35)	0.67 (1.38)
1989	148	4.80 (2.17)	6.54 (1.91)	12.83 (12.08)	2.16 (2.82)	0.71 (1.49)
1990	166	4.50 (2.28)	6.34 (2.07)	12.11 (12.29)	2.23 (3.13)	0.83 (1.64)
1991	170	4.32 (2.28)	6.23 (2.17)	11.95 (12.28)	2.49 (3.66)	0.90 (1.77)
1992	207	4.18 (2.35)	5.94 (2.44)	12.63 (12.50)	2.73 (3.67)	0.84 (1.80)
1993	226	4.03 (2.34)	5.60 (2.76)	12.34 (12.41)	2.81 (3.65)	0.95 (1.88)
1994	231	3.97 (2.38)	5.45 (2.91)	12.34 (12.37)	2.95 (3.72)	1.00 (1.96)
1995	235	3.88 (2.41)	5.29 (3.00)	12.39 (12.46)	3.10 (3.81)	1.03 (2.01)
1996	233	3.76 (2.41)	5.27 (3.03)	12.58 (12.51)	3.28 (3.93)	1.04 (2.06)
1997	233	3.74 (2.42)	5.25 (3.03)	12.70 (12.53)	3.46 (4.04)	1.02 (2.11)

Notes:

(a) Both pollution measurements are in logs

(b) Mean with standard deviation in parenthesis

(c) Total number of plants = 236

Table 2: Summary Statistics for Pollution Sensitive Group and Socioeconomic Status Variables.

Year	**Obs**	**PKIDS**	**PELDERS**	**PPOOR**	**PHSDROP**
1987	118	8.64 (0.69)	12.51 (1.69)	12.75 (4.34)	16.89 (3.53)
1988	133	8.66 (0.65)	12.64 (1.68)	12.68 (4.42)	16.45 (3.50)
1989	148	8.68 (0.63)	12.75 (1.65)	12.61 (4.40)	16.16 (3.54)
1990	166	8.70 (0.60)	12.86 (1.67)	12.34 (4.34)	15.72 (3.53)
1991	170	8.66 (0.58)	12.93 (1.67)	12.27 (4.29)	15.33 (3.49)
1992	207	8.61 (0.56)	12.99 (1.65)	12.35 (4.63)	15.11 (3.49)
1993	226	8.57 (0.55)	13.05 (1.65)	12.18 (4.45)	14.73 (3.35)
1994	231	8.48 (0.55)	13.09 (1.71)	12.07 (4.32)	14.34 (3.26)
1995	235	8.39 (0.55)	13.10 (1.72)	11.96 (4.25)	13.96 (3.19)
1996	233	8.28 (0.56)	13.09 (1.73)	11.85 (4.13)	13.59 (3.12)
1997	233	8.16 (0.57)	13.08 (1.74)	11.70 (4.01)	13.18 (3.04)

Notes:

(a) Mean with standard deviation in parentheses

(b) Total number of plants = 236

Table 3: Summary Statistics for voting and density variables.

Year	**Obs**	**ENVIRVOTE**	**TURNOUT**	**CONVMEMB**	**DENSITY**
1987	118	58.93 (16.39)	0.55 (0.08)	8.68 (3.34)	261.58 (401.91)
1988	133	58.79 (16.25)	0.52 (0.09)	8.69 (3.31)	273.94 (393.47)
1989	148	63.20 (17.77)	0.53 (0.09)	8.76 (3.26)	288.93 (430.43)
1990	166	61.22 (16.08)	0.53 (0.09)	8.92 (3.23)	276.56 (413.32)
1991	170	51.69 (16.37)	0.53 (0.09)	9.00 (3.23)	281.78 (416.44)
1992	207	46.83 (20.17)	0.59 (0.09)	9.00 (3.43)	286.85 (402.36)
1993	226	60.02 (13.97)	0.59 (0.09)	9.03 (3.36)	284.80 (394.77)
1994	231	54.18 (15.96)	0.59 (0.09)	9.03 (3.33)	286.93 (395.81)
1995	235	46.42 (18.37)	0.60 (0.09)	9.03 (3.32)	286.62 (396.06)
1996	233	48.76 (17.95)	0.60 (0.09)	8.99 (3.31)	289.74 (400.66)
1997	233	54.29 (21.91)	0.60 (0.09)	8.99 (3.31)	291.79 (403.90)

Notes:

(a) Mean and standard deviation in parenthesis

(b) Total number of plants = 236

Table 4
First stage regression for the demographic variables.

Dependent Variables	Instrumental Variable	F-statistic
PKIDS	0.483^{**} (17.40)	75.44
PELDERS	0.323^{**} (10.59)	33.13
PPOOR	0.897^{**} (57.78)	80.26
PBLACK	0.385^{**} (14.83)	5.55
PHISP	0.839^{**} (33.73)	34.70
POTHER	0.889^{**} (60.09)	58.77
PHSDROP	0.327^{**} (11.40)	20.75
DENSITY	0.298^{**} (6.63)	10.39

Notes:

(a) Time dummies are included in all equations.

(b) T-statistics are reported in parenthesis.

(c) ** indicates significance at 5% level.

(d) The F-stat p-values are 0 for all regressions; (df = 87, 2012)

Table 5
Results for regression on all demographic variables.

Independent Variables	**PM10**	**SO_2**
AR(1)	0.552**(0.016)	0.474** (0.012)
PKIDS	0.538** (0.117)	−0.182 (0.174)
PELDERS	−0.191 (0.117)	−1.187** (0.294)
PPOOR	0.063 (0.039)	0.236** (0.077)
PBLACK	−0.025 (0.065)	0.005 (0.120)
PHISP	−0.120 (0.064)	−0.148 (0.104)
POTHER	0.056** (0.025)	0.061 (0.047)
PHSDROP	0.056 (0.037)	−0.153** (0.069)
DENSITY	−0.002 (0.002)	0.005 (0.004)
ENVIRVOTE	−0.002** (0.0007)	0.0002 (0.001)
TURNOUT	−1.510** (0.695)	−1.071 (1.116)
TURNOUT_CONVMEMB	0.159** (0.051)	−0.003 (0.087)
Sargan test	102.2 (64)	88.8 (64)
Wald test	3999.7 (10)	4174.4 (10)
K test	165.6 (10)	242.3 (10)
Number of observations	2100	2098

Notes:

(a) Time dummies are included in all equations.

(b) Pollution variables are in logs and demographic variables are in percentages.

(c) Sample Period: 1985-2000 (236 plants)

(d) Standard errors are reported in parenthesis for 2-step GMM estimates robust to heteroskedasticity.

(e) ** indicates significance at 5% level.

(f) In the case of all tests, degrees of freedom for χ^2 statistics are reported in parenthesis.

(g) The Sargan test is a two-step version of the test for serially uncorrelated errors.

(h) The Wald statistic is a test of the joint significance of the independent variables.

(i) The K statistic is the Kleibergen test for joint significance. This test is robust to weak instruments.