Do People Vote With Their Feet?: An Empirical Test of Environmental Gentrification*

H. Spencer Banzhaf
Resources for the Future
1616 P Street NW
Washington, DC, 20036
banzhaf@rff.org

Randall P. Walsh
Department of Economics and Institute for Behavioral Science
University of Colorado
walshr@colorado.edu.

Paper for the NBER Summer Meetings
July 2005

---

* Support for this research was provided by the National Science Foundation NSF SES-03-21566. We thank Lisa Crooks for excellent research assistance. We also appreciate the helpful comments of Pat Bayer, Nicholas Flores, Michael Greenstone, Michael Margolis, Chris Timmins, Ann Wolverton, and participants of the 2004 AERE Workshop, the 2004 CU Environmental and Resource Economics Workshop, and a 2005 AERE session at the ASSA meetings.

I. Introduction

Tiebout's (1956) suggestion that people "vote with their feet" to find the community that provides their optimal tax/public goods pair has played a central role in the theory of local public finance over the past 50 years, motivating such diverse literatures as capitalization and "hedonics," fiscal federalism, and the formation of endogenous public goods. More recently, a new and growing empirical literature has leveraged the equilibrium properties of Tiebout model to identify general equilibrium models of household sorting (e.g. Bayer, McMillan, and Reuben 2005, Ferreyra 2005, Sieg et al. 2004, and Timmins 2003). Given the central importance of Tiebout's insights, there have been surprisingly few direct tests of his premise. In this paper, we use changes across time and space in emission levels of toxic air pollutants to test for the associated changes in neighborhood population and composition that the Tiebout model predicts. Our results provide strong empirical support for the notion that households "vote with their feet" in response to changes in environmental quality. This result has two broader implications. First, it provides direct empirical support for the assumptions underlying the Tiebout model. Second, the presence of what we call "environmental gentrification"—environmentally motivated migration patterns that differ systematically by income or minority status—has important implications for both the analysis of environmental equity and for the design of environmental policies that are aimed at benefiting the less advantaged elements of society.

"Tests" of the Tiebout model generally fall into two categories.[1] Indirect or implicit tests, the most common, have focused on deductive implications of the model. For example, Wallace Oates' (1969) seminal article on the link between local tax and service packages and property values introduced a hedonic model as an implicit test of a Tiebout equilibrium. Brueckner (1982) tested implications of the model related to the efficient provision of public goods. In a recent paper reflecting on the impact of Tiebout's paper today, Oates (2005) highlights the fact that many tests have focused on issues of stratification in demand for public goods and the link between diversity across communities in income and public good provision (e.g. Gramlich and Rubinfeld 1982, Epple and Sieg 1999). Direct tests of migratory responses to public good provision, the second category, have been less common. Graves and

[1] For reviews and discussion, see Rubinfeld (1987) and Oates (2005).

Waldman (1991) found that the elderly retire in counties where public goods are capitalized more into wages than into land prices. Kahn (2000) found migration into California counties with improving air quality. As suggested by these two papers, existing direct tests have utilized identified county differences on a national or regional scale. To the best of our knowledge, our paper is the first direct test of Tiebout-type migration at the local level.

In addition to the broader connection to Tiebout, understanding the way that changes in environmental quality can lead to changes in community composition is of critical importance for the literature on "environmental justice." The environmental justice movement asserts that minority households face greater exposure to environmental pollution than white households. Supporting this assertion are numerous studies documenting a positive correlation between the location of hazardous waste facilities and minority or poor populations (*e.g*. United Church of Christ 1987, Goldman and Fitton 1994, Mohai and Bryant 1992). This correlation is often interpreted as direct evidence of environmental racism. Although others have challenged these findings (e.g. Anderton et al. 1994, Daniels and Friedman 1999), concerns about environmental racism have prompted the US Environmental Protection Agency to create a national Environmental Justice Office and several states to craft legislative responses to the perception of environmental racism.

Understanding the sensitivity of community composition to changes in environmental quality is of key importance for evaluating these claims of discrimination. If "Tiebout sorting" is the driving factor behind the observed correlation between pollution and poor or minority populations, then it may be the case that polluters do not locate in such neighborhoods out of racist motives or even to exploit weaker political opposition, but rather that poor or minority households move to these neighborhoods (or move out differentially less) because of their lower marginal rates of substitution between the environment and housing prices. The potential importance of this sorting for clarifying the issue of environmental equity has led to a small number of studies of the link between changes in environmental quality and changes in community composition. Because our study is in part aimed at overcoming the limitations of this existing literature, we evaluate the current state of this literature in the following section.

II. Environmental Justice and Tiebout

Been (1994) provides the first attempt to evaluate the neighborhood composition effects from changes in pollution exposure. She evaluates the change in community composition following the citing of four hazardous waste landfills, three incinerators and seven landfills—using as controls the changes in the host Census tract relative to the host county. As would be expected, given its small sample size and lack of controls, the study provides limited results. Been (1997) reconsiders the issue evaluating demographic changes for 544 "communities" that hosted active commercial hazardous waste treatment storage and disposal facilities. She finds little direct evidence that the citing of facilities led to changes in demographics.[2] This study again defines host Census tracts as the effected "community." As controls, the study compares changes in host tracts to changes in all non-host tracts. Alternatively, in regression models the study uses only a very sparse set of controls, omitting exiting facilities, lagged effects of older facilities, and other demographic variables other than baseline racial composition.

While this work represents a significant advance in the environmental justice literature, three specific concerns raise doubts about the inference that can be drawn from Been's analysis. First, as Been points out, it is not clear that identifying host Census tracts as the effected community is appropriate. The findings of environmental justice studies have been shown to be quite sensitive to spatial definitions. While relatively homogeneous in population, Census tracts can vary significantly in size and are often quite large, making it possible for their use as a "community" definition to mask demographic shifts that occur within each tract. Additionally, because facilities are often located along Census tract boundaries, attaching facility effects to their host tract introduces significant noise to the empirical model. Second, the use of all non-host census tracts as controls is potentially problematic. If sub-regions of the country that host facilities are experiencing demographic shifts that vary systematically from sub-regions that do not host facilities, the use of all other tracts as controls could lead these systematically varying demographic shifts to mask

[2] However, while she finds contemporaneous correlation between pollution, she does not find them *at the time of the siting*, suggesting some re-sorting must have taken place. Wolverton (2002) replicates this finding using new data and discrete choice models.

environmentally driven composition changes that are occurring within the host sub-regions.[3] Third, using the presence of a hazardous waste treatment storage and disposal facility in a given tract to measure environmental quality abstracts from the fact that there is a great degree of risk heterogeneity associated with different facilities. This heterogeneity arises both from differences in the quantity and toxicity of emissions among facilities and differences across tract-facility pairs in the average distance that tract households are located from the facility.[4]

Several of these issues are addressed in more recent work by Cameron and Graham (2004). They evaluate how demographic composition, measured at the Census block level, varies across time and distance from six superfund sites – finding heterogeneous demographic shifts across the six different facilities. This approach yields much smaller neighborhood definitions, uses neighboring census blocks located further from the facility as controls, and directly accounts for variation in risk associated with varying distances from the facility. Their analysis provides an excellent case study for these six sites. However, due to the limited number of facilities studied, it is difficult to draw broader inferences from their work. In this paper we attempt to move the research on environmentally driven migration forward on two fronts. First, we provide a simple theoretical model and derive the comparative statics for changes in community populations that are associated with changes in environmental quality. We find that over-all migration consistent with the Tiebout hypothesis provides clear predictions for changes in total populations. However, even with differential migration, the model does not provide clear-cut predictions for changes in community composition *relative to other communities*. We clarify the special cases where it does and the information that would be required for better tests of the hypothesis.

Second, we provide an empirical analysis of environmentally-induced migration that attempts to overcome the limitations of previous studies. Toward this end, we evaluate the impact of entry and exit of TRI facilities, as well as changes in toxicity-weighted emission levels, on changes in local community population and racial composition—controlling for demographics and other location specific effects. As our unit of analysis, we use a set of

[3] This concern will hold whether the appropriate "sub-region" is a metropolitan area, urban cores, certain states, or any region whose spatial definition is larger than that of a census tract.

"communities" defined by equally spaced half mile circles and attach TRI emissions to these communities based on their location relative to the facility. Because migratory responses may be highly non-linear functions of demographics, we use both linear regression and a non-parametric matching estimator in the analysis. We find clear evidence of migration correlated with TRI facility emissions and their arrival or exit from a community. Furthermore, we find evidence that TRI facilities cause the composition of a community to become less white over time.

III. Model

To motivate the empirical work that follows, we begin in the spirit of Tiebout (1956) and explore the impacts of changing environmental quality on community composition within a locational equilibrium model. In particular, we use a model of vertically-differentiated communities introduced by Epple, Filimon, and Romer (1984), a more general version of which was recently applied to environmental improvements by Sieg et al. (2004).

Assume a continuum of households that are characterized by their income y and demographic group t. The joint distribution of types and income is given by $f(y,t)$. The marginal distribution of income is given by $f_y(y)$ and the distribution of income conditional on type t is given by $f_y^t(y)$. Household preferences are defined over housing with price P, a numeraire whose price is normalized to 1, and environmental quality G. Household i's preferences are represented by the indirect utility function

$$V_i = V(y_i, P, G). \qquad (1)$$

Each household chooses to live in a community $j \in J$ and, conditional on community choice, chooses a quantity of housing D_i. Each community is characterized by its supply of housing S_j and level of environmental quality G_j, both of which are exogenously determined. To facilitate a characterization of the equilibrium sorting of households across communities, we further assume that household preferences satisfy the "single crossing" property. This

[4] The second factor is driven both by heterogeneity in where facilities are located within tracts (i.e. center vs. border) and in differences in tract size.

condition requires that the slope of an indirect indifference curve in the (G,P) plane be increasing in y.[5]

Given the assumption of single crossing, equilibrium can be characterized by an ordering of communities that is increasing in both P and G. That is, there is a clear ordering of communities from low price low quality communities to high price high quality communities. Further, for each pair of "neighboring" communities (as sorted by this ranking), there will exist a set of boundary households (incomes) that are indifferent between the two communities. Households whose income is below the boundary income will prefer the lower ordered community and those whose income is above the boundary income will prefer the higher ordered community. This leads to perfect income stratification of households across communities.[6] Equilibrium prices P_j and boundary incomes $\overline{Y}_{j,j+1}$ are implicitly defined by the equilibrium conditions of equation (2):

$$V(\overline{Y}_{j,j+1}, P_j, G_j) = V(\overline{Y}_{j,j+1}, P_{j+1}, G_{j+1}) \quad \forall j \in \{1,..., J-1\} \tag{2}$$

$$M \int_{y \in C_j} D(y, P_j, G_J) f_y(y)\, dy = S_j \quad \forall j \in \{1,..., J\},$$

where M is the total mass of households, D(.) is housing demand, and C_j is the set of incomes locating in community j. These equations formalize the J-1 boundary indifference conditions and the requirement that the land markets clear in each of the J communities, yielding 2J-1 equations to identify the 2J-1 endogenous variables.

We use the model to consider two issues important for the analysis of migration and environmental gentrification. First, we consider the implied distribution of households across communities for demographic groups with different income distributions, $f_y^t(y)$. Second, we evaluate how the predicted demographic composition of communities change in response to changes in environmental quality. Consider two demographic groups, Type 1 and Type 2.

[5] For a discussion of the single crossing property in this context see Epple & Sieg (1999).
[6] It is straightforward to relax this assumption by introducing heterogeneity in tastes, so that there is heterogeneity of income within each community, but perfect stratification by tastes for each income (see Epple and Sieg 1999 and Sieg et al. 2004). Accordingly, this assumption is not critical for the following implications of the model and is relaxed in the following simulation exercise.

Assume that their conditional income distributions $f_y^{\,1}(y)$ and $f_y^{\,2}(y)$ are such that the mean income for Type 1 individuals is less than the mean income for Type 2 individuals. Figure 1 provides a graphic representation of the distribution of these demographic types in a system of two communities, with Community 1 having the lower (*P*, *G*) pair and Community 2 the higher. All households to the left of $\overline{Y}_{1,2}$ sort into Community 1, all to the right sort into Community 2. The figure shows that, due to differences in the income distributions, in equilibrium Community 2 will have a much higher concentration of Type 2 individuals and Community 1 a much higher concentration of Type 1 individuals.

The above result is completely expected given the model. However, the comparative statics associated with a change in environmental quality in one of the communities is more subtle. Consider the impact of an increase in the environmental quality in the lowest G community in a system of two communities. Evaluating the resulting demographic responses requires identifying the shift in the income boundary, $d\overline{Y}_{1,2} \big/ dG_1$. To evaluate this shift, we assume that housing demand is separable from *G* and apply the implicit function theorem to the boundary indifference condition and two market clearing conditions from equation (2). This yields the following comparative static relationship:

$$\frac{d\overline{Y}_{1,2}}{dG_1} = \frac{-V_{G_1}^{\;1}}{(V_y^{\;1} - V_y^{\;2}) - f_y(\overline{Y}_{1,2})\left[\dfrac{D(P_1,\overline{Y}_{1,2})V_{P_1}^{\;1}}{\int_0^{\overline{Y}_{1,2}} D_{P_1}(P_1,y)\,f(y)dy} + \dfrac{D(P_2,\overline{Y}_{1,2})V_{P_2}^{\;2}}{\int_{\overline{Y}_{1,2}}^{\infty} D_{P_2}(P_2,y)\,f(y)dy}\right]}, \tag{3}$$

where

$$V^1 = V(\overline{Y}_{1,2}, P_1, G_1)\,,$$

$$V^2 = V(\overline{Y}_{1,2}, P_2, G_2)\,,$$

D(.) is the household demand function, and subscripts denote partial derivatives.

The key to signing the derivative in equation (3) is to recognize that the single crossing property implies that $(V_y^{\ 1} - V_y^{\ 2}) < 0$ implying that $d\overline{Y}_{1,2} \,/ dG_1$ is positive.[7] Figure 2 illustrates the impact of an increase in G_1 on the equilibrium sorting In response to the change, the indifference boundary $\overline{Y}_{1,2}$ moves to the right and the set of household in the shaded region A relocate from Community 2 to Community 1. If G_1 were to fall instead, or if G_2 were to increase instead of G_1, an opposite shift would occur. At the aggregate level this change leads to an increase in population for Community 1 and a decrease in population for Community 2. (And to an increase in prices in Community 1 and decrease in Community 2.)

What is the change in the community 1's composition relative to community 2? Surprisingly, the model does not offer clear predictions. As the bordering households move from Community 2 to Community 1, Community 1 gets richer, but meanwhile Community 2 loses its lowest income residents and therefore also experiences an increase in average income. Similar statements can be made for Type 1 and Type 2 residents. Because the percentage of Type 2 residents in the set A is higher than that of the original residents of Community 1 the share of Type 2 residents in Community 1 increases. As with income the percent of Type 2 individuals also increases in Community 2. Thus, we have the counter-intuitive result that increasing the level of G_1 leads to an increase in average income and an increase in the percentage of Type 2 households *for both communities!*

We thus have three propositions.

Proposition 1 (Scale Effect). For any two communities, a marginal increase in G in one community relative to the other will cause population to rise in the community experiencing the improvement and to fall in the other community. This scale effect can be stated in relative terms, as well as absolute terms, since the population in the first community must also rise relative to the second.

Proposition 2 (Absolute Composition Effect). Ceteris paribus, a marginal increase in G in any community will increase its average income and increase the share of higher average income demographic groups (e.g. racial groups).

---

[7] By the definition of $\overline{Y}$, $V(\overline{Y}_{1,2}, P_1, G_1) = V(\overline{Y}_{1,2}, P_2, G_2)$. Since all those with incomes

Proposition 3 (Relative Composition Effect). The change in the average income, or mean share of demographic groups positively correlated with income, in a community experiencing a marginal increase in G, relative to another community is indeterminate.

These results suggest that for small changes in Community 1's environmental quality, there are no clear predictions for the relative change in community compositions. However, if we consider a larger change—one that raises the level of environmental quality in Community 1 above that of Community 2—clearer predictions arise. This change will cause the populations in Communities 1 and 2 to switch places, resulting in an increase in average income and an increase in the share of Type 2 individuals in Community 1, while average income and Type 2 share drops in Community 2. Thus, for large changes in quality, the model predicts that the population share of groups with higher average incomes will increase in communities experiencing improvements in environmental quality.

The above analysis abstracts from issues of tastes and of spillovers across multiple communities. To gain a better sense of the complicated interactions involved in a richer framework, we conclude this section by considering a model of preferences that includes an additional heterogeneity parameter that captures taste for environmental quality, α_i, as in Epple and Sieg (1999) and Sieg et al. (2004). We also increase the number of communities in the model. The complexity of the richer model precludes analytical comparative statics. Instead, we assume a functional form for preferences that is consistent with a constant price and income elasticity of demand and then simulate equilibrium sorting before and after an improvement in the environmental quality in the lowest community.

Given y_i and α_i, conditional on locating in community j, assume household i's utility is given by

$$V(P_j, G_j | y_i, \alpha_i) = \left(\frac{1}{1-\nu} y_i^{1-\nu} - \frac{1}{1+\eta} P_j^{1+\eta} \right) G^{\alpha_i}, \qquad (4)$$

where ν and η, the income and price elasticity of demand, are set to 0.7 and -0.7 respectively. As above we consider a model with two types of households. The distribution of incomes are

---

higher than $\overline{Y}$ prefer Community 2, $V(\overline{Y}_{1,2} + e, P_1, G_1) < V(\overline{Y}_{1,2} + e, P_2, G_2), \forall e > 0$.

assumed to be log-normal with Type 1 and Type 2 households having a mean income of \$40,000 and \$60,000 respectively, both with a standard deviation of \$20,000. For each group, tastes are assumed independent of income and normally distributed. The mean of α is set equal to 0.85 for Type 1 individuals and 1.15 for Type 2 individuals and the standard deviation of α is set equal to 0.25 for both groups. Finally, it is assumed there are five communities of identical size with environmental quality levels set to {0.5, 1, 1.5, 2, 2, 5}.

Figure 3 displays the distribution across income-taste space for the two groups as well as the indifference boundaries in the baseline equilibrium. With heterogeneous tastes, the "boundary" between communities, previously a point in the Y dimension only, is now a locus in Y-α space. It is clear from the figure that while some households of each type will locate in each community, Type 1 households will be concentrated in lower ranked communities because on average they have a lower mean taste and income level. We should also note that while we are interpreting alpha as a taste parameter, it is also consistent with the presence of housing discrimination against Type 1 households in higher *G* communities.[8]

Consider now the impact of an increase in environmental quality in community 1 from 0.5 to 0.95 on the equilibrium sorting of households. Figure 4 shows the how the community boundaries move in response to this change and Table 1 summarizes the change in community compositions from the improvement in G_1. The results are largely analogous to the analytical example. All four of the indifference loci shift up and to the right, with the largest shifts occurring in the communities closest to Community 1. As a result, population increases in Community 1 and decreases in communities 2 through 5. Average income increases in four of the five communities, and the Type 2 share increases in all five.

As with the above analytical example, the magnitude of these effects is largest in the communities closest to Community 1. If instead we consider increasing the environmental quality in Community 1 to $G_1=1$ and decreasing the environmental quality in Community 2 to $G_2 = 0.5$, the new equilibrium leads to a switching of populations between the two communities. The result is that Community 1 (Community 2) sees an increase (decrease) in population and an increase (decrease) in Type 2 share.

Taken together, the analytical and simulation results provide two important insights relative to identifying environmental gentrification using reduced form estimation techniques. First, the analysis consistently predicts that increases in environmental quality should be correlated with increased populations. Second, it demonstrates that the link between environmental changes and community composition can be muddied with small changes in environmental quality failing to yield a clear link between improvements in environmental quality and relative increases in community share for wealthier demographic groups. However, for large-scale improvements that change the relative ranking of communities, the composition effect is always predicted.

Estimation strategy

We test for such demographic scale and composition effects in this paper. Our work is most similar in spirit to Kahn (2000), who presents general evidence for the Tiebout hypothesis, documenting over-all county-level population growth in California that is correlated with ozone improvements. However, his work did not consider composition effects, and the fairly aggregate county-level analysis bears further scrutiny at more local levels.

Our analysis is at a much finer level of aggregation, relating changes in community populations (in level terms and percentage terms) and changes in racial composition to changes in exposure to air pollution. For composition effects, we focus on race, which corresponds to the Type 1 and Type 2 individuals in our model, rather than income because it has been a focus of much of the environmental justice literature, is measured with less noise, and is reported by the census at the block level, which is important for reasons described below. Environmental quality is measured as the toxicity weighted exposure to air pollution released from sites listed in the Toxic Release Inventory (TRI). We examine the effect on migration of changes in exposure to air pollution – evaluating both the effect of changes in emissions from existing sites and the effect of entering and exiting polluters.

One potential problem with this analysis is that, if firms do site their facilities based on race as claimed by the environmental justice literature, the relationship between changes in

[8] Unfortunately, it is difficult (if not impossible) to distinguish between tastes and housing discrimination that is correlated with community characteristics based solely on the observed sorting of households across locations.

pollution and changes in demographics would be endogenous, biasing the estimated coefficients upward. (By similar reasoning, if they avoid more densely populated areas, as found by Been 1997 and Wolverton 2002, the estimated coefficients for the scale effect would be biased toward zero.) To address this problem, we also identify lagged demographic responses to TRI sites that were sited before the demographic changes occur and that can therefore be treated as exogenous.

IV. Data

Constructing the dataset necessary to test for environmentally induced migration requires three related tasks. First, we identify a set of spatially delineated "communities." Second, we construct demographic composition measures for each community for 1990 and 2000. Finally, for each community we construct measures of the toxicity-weighted level of exposure to air pollution in 1990 and 2000 based on data from the U.S. EPA's Toxic Release Inventory (TRI). This section of the paper begins by discussing the spatial definition and data matching for the communities in our data set. We then discuss in more detail the demographic and pollution measures utilized in the empirical analysis.

Definition of Communities

Our analysis requires a set of communities whose boundaries remain fixed between 1990 and 2000. One approach would be to use census tracts, block-groups, or blocks as our community definition. This approach is problematic because these definitions change from Decennial Census to Decennial Census. Been (1997) found that one-fifth of tracts nationally had changed boundaries between censuses. Been (1997) and Wolverton (2002) address this problem by aggregating up to the greatest common boundaries, but in the process still eliminate some areas and end up with higher levels of aggregation.

An alternative approach has recently become available with the release of Geolytic's Neighborhood Change Database. This product provides access to 1960 through 2000 Census data consistently aggregated to 2000 Census Tract definitions.[9] While this new dataset solves

For a discussion of these issues, see Bayer, McMillan, and Reuben (2005).
[9] Cameron and Crawford (2004) use this data set to analyze the change in the demographic composition for census tracts surrounding six facilities located in New Jersey and California.

the problem of changing definitions by providing aggregation to the 2000 tract boundaries, the use of Census tracts as "communities" may itself be problematic for three reasons. First, census tracts are locally defined to create relatively homogenous entities. Although some see this as a virtue as it gives more integrity to the concept of community, such gerrymandering may also bias results. For example, if polluting sites have an impact on the demographics of only the most local neighborhoods, and these neighborhoods are conjoined with other neighborhoods with similar characteristics to form census tracts, it would induce correlation between the polluting site and the wider geographic entity (the tract). Second, although roughly equal in populations, census tracts range greatly in size. For example, in California, Tracts range in size from less than a tenth of a square mile to more than one thousand square miles. This creates problems in controlling for this large degree of heterogeneity when estimating migration models. Finally, previous research on the correlation between race and environmental quality has shown that results can be quite sensitive to community definitions (Anderton et al. 1994, Hersh 1995). Census tracts may be too aggregate a unit and in any case preclude sensitivity analysis along these dimensions.

For these reasons we take a different approach to neighborhood definitions. We define neighborhoods as a set of half-mile circles (alternatively one-mile circles) evenly distributed across our study area. Using the GIS software package ARCVIEW we construct weights that are used to attach environmental quality data from the TRI and demographic data from Census blocks to our communities. One limitation of this approach is that we have more confidence in imputed demographics for our communities based on Census blocks (the smallest unit) than for those variables based only on the larger Census block-groups. Partly for this reason, we focus on race rather than income when measuring composition effects.

The specifics of the data construction are as follows. First, to keep the data construction task manageable, we restrict our analysis to the state of California. California is attractive because of its racial heterogeneity and because of its size relative to other potential states. To further restrict the size of the data task and to reduce the heterogeneity between different communities, we limit our analysis to locations that were denoted as urban in the 1990 census. We construct our communities by first placing an equidistant grid across our study area. Grid points are one-half mile apart for the half-mile circles and one mile apart for the one-mile circles. If the study area were small in scale, for instance a single county, it would be possible to treat lines of latitude and longitude as an equally-spaced grid. However, because of the size of California, the distance between lines of longitude varies by approximately 13% as one moves from the southern border to the northern border of the state. To avoid an uneven sampling density between northern and southern portions of the state it is necessary to account for this variation.[10]

Once grids that cover the entire state have been constructed (one each for the half mile and one mile circles) a quarter mile (half mile) buffer is placed around each point in the grid, yielding a set of half mile (one mile) diameter circles that are evenly distributed across the state. The set of circles that fall within Census 1990 Urban Area boundaries are then selected and all circles that lie across water are dropped. This process yields 6,218 "communities" using one-mile circles and 25,166 "communities" based on half-mile circles. Figure 5 shows the distribution of communities across the study area.

Census Data

As noted above, we aggregate demographic data from the 1990 and 2000 censuses into our circle-communities. We collected block-level data on the total populations of each racial group, both as individuals and as households, and economic variables including homeownership rates, rental rates, and self-assessed home values. We also collected block-group-level data including average incomes, educational attainment, and workforce descriptors

Demographic count data on numbers of individuals or households by race and other categories are assigned to our communities. Specifically, for each block, a share of each demographic count is assigned to communities based on the percentage of the block's geographic area lying within each community.[11] Even for our half-mile communities, most blocks are assigned entirely to a single community, and 99% are assigned to five or fewer. Table 2 summarizes the opposite mapping, the number of blocks assigned to each community. Because of the splitting of blocks between the 1990 and 2000 census there are more blocks

[10] The grids are constructed using the following factors: 1 degree of latitude = 69.172 miles and 1 degree of longitude = cos (latitude) * 69.172.
[11] For block-group-level data, the values were distributed to the blocks based on population shares, then distributed to the communities as for the block-level data.

assigned per circle in 2000 than in 1990. The 50th percentile ranges from a low of 10 blocks per circle for 1990 half-mile circles to 38 for 2000 one-mile circles.

Table 3 provides descriptive statistics of the demographic characteristics of the half-mile communities. Note that they are easily comparable, since they are all circles of equal size (about 0.785 square miles). Most communities have small populations, with an average 1990 population of about 772 persons, but with a wide interquartile range of 98 to 1162. Most communities are also largely non-Hispanic white, with the mean share of 67% across communities. However, this share decreased on average by 9% from 1990 to 2000. Moreover, whites are disproportionately located in less dense communities; weighting by population reduced the mean white share to 54%. Hispanics are the largest and fastest growing minority, with an average 19% share in 1990, 27% when weighted by population. Blacks and Asians represent smaller minorities, with average 1990 shares of 5% and 8% respectively. Again, there is substantial variation in these data. Among communities with more than 100 residents, there is at least one where each racial group is completely absent and at least one where each racial group makes up 90% or more of the population.

TRI Data

As a measure of pollution exposure we use the EPA's Toxic Release Inventory (TRI) to find releases of environmental pollution at firms throughout the United States. Firms handling more than 10,000 pounds each year of certain hazardous chemicals have been required to report these emissions since 1987.[12] This censoring at the reporting threshold gives rise to a kind of errors-in variables problem. As we discuss below, as in the usual case, it is likely to have a "conservative" effect on our results, biasing them to zero.

TRI facilities have been a focus of the environmental justice literature.[13] After early work looking only at the presence of a facility or total tons of emissions of all kinds, recent work has begun looking at emissions weighted by *toxicity*, so that more potent pollutants are given more weight than less potent ones. In the first, crude attempt at such weighting, Sadd et al.(1999) estimate an ordered probit model on Los Angeles neighborhoods with no TRI

---

[12] The list of reporting chemicals greatly expanded in 1994. To maintain a consistent comparison of TRI emissions over time, we have limited the data to the common set of chemicals used since 1988.
[13] E.g., Arora and Cason (1996), Brooks and Sehi (1997), Kriesel et al. (1996), Morello-Frosch et al. (2001), Rinquist (1997), Sadd et al. (1999), Wolverton (2003).

releases, with non-carcinogenic releases, and with carcinogenic releases. In follow up work, Morello-Frosch et al. (2001) use a linear index of cancer risks. Both these papers make the arbitrary assumption that carcinogenic releases are more potent (and important) than non-carcinogenic releases. However, such is not necessarily the case, as some pollutants such as particulates have strong correlations with cardio-vascular deaths and some carcinogens may have very weak potency.

As an alternative, we use a toxicity weighted index of all emissions in the TRI. We use toxicity weights developed by the US Environmental Protection Agency and available in its Risk Screen Environmental Indicators model (RSEI).[14] These weights have been used in recent work by Ash and Fetter (2002), who also go further to look at modeled exposure based on atmospheric dispersion.

The latitude and longitude for each facility was taken from a recent careful quality control analysis by the EPA.[15] This geographic information allows a match of facilities to our communities. We construct buffers (quarter mile and half mile) around each TRI site and then assign emissions from a given TRI site to the communities that lie within the given buffer. The sample of TRI sites is the 2,311 California TRI sites located such that a half-mile buffer intersects at least one community. Figure 6 illustrates the approach used to assign emissions for the case of half mile TRI buffers and one-mile circles. In the figure, the shaded circles are half-mile buffers around four TRI sites. The un-shaded circles are communities. Emissions from a given TRI site are assigned to communities based on the percentage of their buffers that lie within a given community. For instance, in Figure 6, 3.1% of the emissions from TRI site A are assigned to community N1, 17.9% to community N2, and 52.4% to community N3. In this way, we can consistently aggregate emission levels from multiple TRI sites into a total exposure in each community.

Table 2 summarizes the assignment of all TRI sites, active at some time during the 1988 to 2000 time period, to communities for each community-size, buffer-size pair. For those communities that are exposed to at least one TRI site, the table summarizes the distribution of the number of sites to which each community is exposed. In all cases, the 50th

[14] Information about this model is available at http://www.epa.gov/opptintr/rsei/.
[15] This quality control analysis provides a predicted accuracy for each Site's location data. Fifteen sites are dropped due to poor quality geo-coding data.

percentile is either 2 or 3, with the maximum exposure going as high as 69 sites in the case of one-mile circles and one-mile TRI buffers. On a community basis, Table 3 indicates that 10 percent of half-mile communities were exposed in the baseline period (1988-1990), with 4 percent losing exposure by 1998-2000 and 1 percent gaining exposure. It also shows mean toxicity weighted exposure among all communities and among those exposed.[16]

Additional Spatial Variables

A number of factors, including other spatially distributed amenities, are likely to drive sorting across communities and should be controlled for as well as possible. As controls, we include in some models coarse controls for location effects, including distance to the coast and degrees latitude. However, our main approach to controlling for unobserved spatial amenities is local fixed effects. We use two sets of fixed effects: school districts and zip codes. Both are very local measures that are consistent with the notion that households are likely to choose a larger area based on other factors and then sort within that area based on the most local amenities. School districts have the advantage of mapping directly into an important local public good whose quality is otherwise notoriously difficult to measure. We find the share of each community that falls within each of the 226 school districts in our urban areas and assign a continuous variable on [0,1] to that community for each school district. Seventy-six percent of half-mile communities lie entirely within one school district, 21 percent within two, and the remainder within three or four. Table 2 shows the number of communities falling within each school district. Zip codes are even more local measures. Here, we simply assign each community to one of the 883 zip codes in our area based on the zip code of the community centroids. Table 2 reports on the distribution of communities across zip codes. The median zip code is assigned 21 half-mile circles and 6 one-mile circles.

V. Estimation and Results

Using these data, we test for differential changes in community population/demographics that are induced by changes and/or baseline differences in TRI emissions. Our primary results

[16] This is in contrast to the approach that has typically been taken in the environmental justice literature of assigning the pollution from a given facility to the census tract in which it resides. This traditional approach is

center around models using half-mile diameter communities and half-mile diameter buffers around TRI facilities. We consider respective 1-mile diameters in sensitivity analyses.

Contemporaneous Patterns in exposure

We begin by confirming, in our new data, the correlation between pollution and minorities as found in the environmental justice literature. To this end, we first test for correlations between 1990 TRI locations and minority populations, estimating probit models for proximity to a TRI site and tobit models that exploit continuous differences in exposure to TRI emissions. Both models take the same basic structure. The Tobit case is shown in equation 5.

$$y_i^* = b_0 + b_R R_i + b_D D_i + b_L L_i + e_i \tag{5}$$
$$y_i = 0 \text{ if } y_i^* \leq 0$$
$$y_i = y_i^* \text{ if } y_i^* > 0$$

where i indexes communities, y_i is the observed toxicity-weighted TRI exposure of a community, y_i^* is the corresponding latent variable, R is a vector of racial characteristics, D is a vector of other demographic variables, and L is a vector of locational variables (including school district fixed effects). Table 4a presents the results of three Tobit models with varying levels of control. We find a positive and statistically significant relationship between a community's TRI exposure and the presence of minority populations, even when controlling for other demographic factors and for locational factors with school-district level fixed effects. Not surprisingly, Probit models (not reported here) find the same basic relationships. Under the Probit specification, a one percentage point increase in a community's Black (Hispanic) population is predicted to increase the probability of TRI exposure by 0.15 (0.30) percent. These findings are consistent with previous results from the environmental justice literature.

The notion that populations are fixed and pollution is endogenous, as expressed in most work on environmental justice, conflicts with the Tiebout's assertion that households "vote with their feet." In the remainder of our analysis, we take adopt a "Tiebouvian"

particularly problematic for handling facilities located on or near the boundary of their census tract. The ability to overcome this problem is one of the advantages of the methodology used here.

perspective and attempt to identify the impact of changes in the level of toxic emissions in a given neighborhood on the population of households choosing to live in said neighborhood.

We begin by re-evaluating the contemporaneous correlation between TRI emissions and neighborhood composition. For each race (non-Hispanic White, Black, Hispanic, and Asian/Pacific Islander) we regress the groups representation in a given community on the presence of a TRI site and TRI emissions, as well as density and school district fixed effects. The basic model is presented in equation 6.

$$R_{ir} = \gamma_{0r} + \gamma_{yr} y_i + \gamma_{Dr} D_i + \gamma_{Lr} L_i + u_{ir}, \tag{6}$$

where R_{ir} here is the percentage of each community's population in racial group r. The results, shown in Table 4b, again indicate that both the presence of a TRI site and TRI emissions are associated with smaller White and Asian populations and greater Hispanic populations.

Of course, as discussed above, it is not possible to say whether these correlations are due to systematic firm behavior or systematic sorting by households. Accordingly, having confirmed this pattern in a single cross-section, or next goal is to develop a model which allows us to test for migratory responses associated with the presence of toxic emissions from TRI sites.

Migration: Scale Effects

The strongest prediction of the Tiebout model is that the introduction of a TRI facility should cause individuals to leave the community (and that the exit of a facility should cause them to enter). To test this hypothesis, we regress both level changes and percentage changes in population from 1990 to 2000 on TRI exposure and other controls.[17] TRI exposure is measured as the three-year lagged average, anchored respectively on 1990 and 2000, of the toxicity-weighted emissions of the 1988-defined chemicals, allocated to each community as described previously. As exposure variables, we include measures of both shocks and baseline exposure. As shocks, we include discrete indicators for when a community changes

[17] To develop an operable definition of a percentage change, we use the average of the 1990 and 2000 levels in the denominator. Within our data, this measure is approximately normally distributed and is bounded above and below by +2 and –2 respectively.

status from exposed to not exposed or from not exposed to exposed,[18] plus continuous measures of the change in emission levels (which picks up the magnitude of those entering or exiting facilities as well as changes at those continuously emitting). Because in reality populations do not adjust instantaneously, we also include an indicator and continuous measure of 1990 exposure to pick up lagged reactions to previous exposure. While there is a potential issue of endogeneity when considering simultaneous 1990-2000 changes in both population and environmental quality, these concerns are diminished when evaluating the relationship between 1990-2000 population changes and 1990 TRI emissions.

The model for this analysis is presented in equation 7.

$$\begin{aligned}\Delta POP_i \ = \ & d_0 + d_{BL} I_i^{BL} + d_{NEW} I_i^{NEW} + d_{EXIT} I_i^{EXIT} \\ & + d_y y_i^{1990} + d_{\Delta y+}(\Delta y_i | \Delta yi > 0) + d_{\Delta y-}(\Delta y_i | \Delta yi < 0) \\ & + d_R R_i + d_D D_i + d_L L_i + w_i \qquad (7)\end{aligned}$$

where ΔPOP is the change (or percentage change) in population from 1990 to 2000; I^{BL}, I^{NEW}, and I^{EXIT} are indicator variables for whether the community had any 1990 baseline exposure, went from no exposure to some exposure, or went from some exposure to no exposure; y_i^{1990} is the level of baseline toxicity-weighted exposure; $\Delta y_i | \Delta yi > 0$ is the change in toxicity-weighted exposure, if positive, and $\Delta y_i | \Delta yi < 0$ is the change in toxicity-weighted exposure, if negative.

We estimate four basic models with different levels of control for confounding factors (the *R*, *D*, and *L* variables). Our first model includes no other controls. While clearly lacking any pretense to identifying causality, this model does give a signal as to the overall pattern of changed exposure resulting from migration. Our second model controls for the baseline demographic variables listed in Table 3, including squares of these terms. As an important spatial amenity, it also includes the FBI crime rate imputed from nearby jurisdictions. Finally, this model also includes spatial effects measured by latitude and distance to the coast in kilometers. Our third and fourth models contain the same demographic controls but replace the spatial variables with school district fixed effects and zip code fixed effects respectively.

[18] Note that the discrete variables indicate proximity of a community to any facilities over time, which is related to but not the same as the entry and exit of firms.

Finally, all models are estimated with and without baseline population weights for the communities.

The inclusion of these spatial fixed effects is key to our identification strategy. One major concern with this type of analysis is that changes in a neighborhood's TRI emissions are likely associated with changes in that neighborhood's economic conditions. Such changes in economic conditions can reasonably be expected to be associated with changes in the neighborhood's population and/or demographic mix—leading to problems of endogeneity and biased estimates. By using school district or zip code fixed effects to control for this potential endogeneity, we are in effect arguing that the relevant scale for considering the role of economic conditions in neighborhood composition is larger than the half-mile circles that comprise our neighborhoods. Put another way, we are assuming that migratory responses associated with economic conditions operate at spatial scales greater than a zip code area (or school district) while responses to toxic emissions can be picked up at the scale of our half-mile circles.

Before turning to the results, note that some of the limitations associated with these data, combined with our local fixed effects, are likely to bias our results toward zero, so that our results can be described as "conservative." First, note that we are essentially comparing "treatment" communities within our TRI buffers to "control" communities outside them, but that we cannot know the *true* area of impact of the TRI facilities. If a TRI facility's actual impact is narrower than our buffers, the treatment communities will be contaminated by control areas, diluting the differential. On the other hand, if its actual impact is wider than our buffers, the TRI facilities will have some impact on control communities, again diluting the identified differential. Moreover, this latter effect is accentuated by our local fixed effects, since the very local controls are the most likely to be affected by the TRI facility. This is one reason we focus on the results of the smallest, half-mile diameter, buffers. Second, as noted above, TRI emissions are censored based on threshold quantities of individual chemicals handled by the facilities. This means that some unobserved, but low-level, emissions are occurring undetected. To the extent this means that some communities diagnosed as controls are in fact exposed, it again dilutes the estimated differential.[19] For these reasons, if we find migration effects with these data, using this design, we have reason to be confident in the existence of Tiebout effects related to pollution.

The results from our scale-effect models are presented in Tables 5a and 5b. Both the weighted and un-weighted models fit reasonably well given the cross-sectional nature of the data, with R^2's of 0.04 to 0.18 for models with controls but no fixed effects and 0.09 to 0.58 for the fixed effect models. Aside from the important impact of the TRI sites, we find that denser communities gain more people from 1990 to 2000, as do communities with lower housing prices but higher rental rates. We also find statistically significant non-linear adjustments to baseline racial composition.

Table 5a presents the estimated scale effects associated with toxic emissions from TRI sites from the un-weighted regressions. The table includes estimates for models with both changes in population level and percentage changes in population as the dependent variable. The table reports three different treatment effects. The “Average Effect of Baseline TRI Exposure” estimates the average effect on a neighborhood’s 1990 to 2000 population change from being exposed to TRI emissions in 1990. The “Average Effect of New TRI Exposure” estimates the average effect on a previously un-exposed neighborhood of becoming exposed to TRI emissions. And the “Average Effect of Exiting TRI Exposure” estimates the average effect on a previously exposed neighborhood of losing all of its TRI exposure. These treatment effects are calculated as combinations of the estimated coefficients on both indicator and continuous variables. Specifically,

$$\text{Avg Baseline Treatment} = d_{BL} + d_y \left(\frac{1}{card(BL)} \sum_{i \in BL} y_i^{1990} \right) \tag{8}$$

$$\text{Avg New Treatment} = d_{NEW} + d_{\Delta y+} \left(\frac{1}{card(NEW)} \sum_{i \in NEW} \Delta y_i \right)$$

[19] If the unobserved facilities merely added additional exposure to already-exposed communities, it would tend to bias our estimates of the importance of pollution as measured as a continuous variable. However, our biggest impacts appear to be in the extensive margin (exposed/unexposed), and the propensity score matching results presented below focus only on this margin.

$$\text{Avg Exit Treatment} = d_{EXIT} + d_{\Delta y-}\left(\frac{1}{card(EXIT)}\sum_{i \in EXIT} \Delta y_i\right)$$

Thus, for example, the estimated effect of average TRI exposure, relative to no exposure, is the estimated indicator variable for exposure, plus the estimated coefficient on the continuous measure of exposure times the average exposure among those communities with exposure. Similar logic holds for the effect of new and exiting exposure. Note that the first two treatments are relative to communities that never experience exposure, while the "exit" treatment is relative to the set of communities that had baseline exposure.

Both the change in level and percent change models provide statistically significant policy relevant evidence of migratory scale effects consistent with the Tiebout hypothesis. Focusing on the percent change model, on average, baseline exposure to TRI Emissions is associated population declines that range from 10% to 16% depending on the model. Likewise, the appearance of new toxic emissions in a previously un-treated neighborhood is associated with population declines between 5% and 9%. Finally, the model predicts consistent responses in the opposite direction for communities that lose exposure. On average these communities are predicted to experience population gains of 5% to 7%.

These un-weighted models take as their unit of analysis communities. They tell a story about what is happening at different places. As such, we view this approach as appropriate for evaluating the effect of Tiebout forces across spatially differentiated neighborhoods. However, from a policy perspective, we might be equally interested in understanding the average effect of these changes on the population. To better understand how populations are behaving, we re-run these regressions weighting by the baseline population. These weighted regressions are reported in Table 5b. The table indicates a similar qualitative pattern of migratory responses, but with level effects somewhat higher and percentage effects much lower than the unweighted regressions. This result is not unexpected as the weighting scheme down-weights less populated areas where larger percentage changes in populations are more likely to occur. In general, the effects continue to be statistically significant—with the exception of the estimated effect for new TRI exposure which remains negative, but loses significance in some models. We interpret these results as strong evidence in support of the scale effects predicted by our simple theory model.

To verify the robustness of these results, we employ a large number of sensitivity analyses. First, we replicated these models with one-mile-diameter communities. For baseline exposure, the effects are of greater magnitude (even in percentage terms) and greater significance for unweighted models and, for weighted models, are likewise greater for the models with no controls and basic controls, but quite similar for the models with fixed effects. The estimated effects for new and exiting TRI exposure are also similar. Second, we tested many alternative definitions of the exposure variable. In particular, we used one-mile buffers around TRI facilities instead of half-mile buffers. We also tried used 1990 and 2000 emissions only (rather than three-year averages), raw emission levels unweighted by toxicity, and a measure of "emissions" that treated each facility equally (so that communities differed only in the number of TRI sites to which they were exposed and their proximity to those communities). None of these sensitivity analyses changed the qualitative nature of the results, although the latter model did lower the magnitude of the effects, suggesting that actual pollution levels are important. Next, we tested the sensitivity of the results to the inclusion of the noisier economic data measured at the block-group level, and found that dropping them did not have much impact on these scale effects. Finally, we also estimate separate models on subsets of the data: on only those communities with no baseline exposure to estimate the effect of a new exposure; on only those communities with baseline exposure to estimate the effect of losing exposure; and on only those communities which do not change status over time to estimate the effect of baseline exposure. None of these variations changes our results. Thus, our evidence is highly robust and strongly consistent with the Tiebout hypothesis.

Migration: Composition Effects

While our theory model provides strong predictions regarding scale effects, it does not provide strong predictions about composition effects, relative to other communities, except for large changes in public goods that affect the relative rankings of the communities. Non-marginal changes in exposure caused by exiting and entering TRI facilities may well qualify as such changes. In any case, these composition effects remain of empirical interest as they may partly explain the observed correlations between race and pollution exposure in the environmental justice literature.

Ideally, we would use our data to evaluate both changes in racial composition and changes in income composition in our communities. Unfortunately, no income data is publicly available at the census block level. We do attach income data to our neighborhoods that we impute from publicly available income data for each census block's block group. However, given the very imprecise nature of this imputation, we use this imputed data as additional controls and are not comfortable using income composition as the dependent variable in our model. This limitation has implications for what our data can say about composition effects. From our Tiebout perspective, racial composition can be expected to change in response to changes in TRI emissions through less direct channels than are operating in the scale effects model. In our theory model, we highlight correlations between race and income as one channel that could lead to environmentally induced composition changes. Other potential channels include housing discrimination and correlation between race and tastes for environmental quality. The indirect nature of these channels of influence serves to further weaken the model's theoretical predictions regarding the link between TRI emissions and racial composition.

To test for these composition effects, we re-estimate the scale effects model using as dependent variables the share of each community's population made up of Non-Hispanic White, Blacks, Hispanics, and Asian/Pacific Islanders. The fit is quite good with R^2's of 0.11 to 0.42 for models with statistical controls but no fixed effects and 0.29 to 0.68 for models with fixed effects. Baseline observables are typically significant. More urban areas are becoming more White and more Hispanic but less Black and Asian. Neighborhoods with more expensive baseline housing are also becoming more White. Finally, highly White areas are generally becoming less White over time, suggesting a structural shift to less segregation, or simply randomness or entropy causing regression to the mean. These are general trends, but the effects are highly non-linear, consistent with a "tipping model" in which cut-off points for racial shares demark adjustments in opposite directions.

Table 6a presents the effects of average TRI exposure using the unweighted model. (The specification is identical to that of Equation 7, with changes in racial composition now on the left-hand side.) There is mixed evidence of a composition effect broadly consistent with intuition. Baseline and new TRI exposure cause communities to become 1 to 2 percent less white, a statistically significant change, with most of the offsetting gain made by

Hispanics. Losing TRI exposure generally has the opposite effect, but the changes are generally not statistically significant. The population-weighted regressions, reported in Table 6b, generally show the same trends but with small magnitudes and generally statistically insignificant effects. While these estimated composition effects are less robust than are those for the scale effects, they provide additional evidence in support of the general Tiebout model.

Before moving on to the Propensity Score Matching estimates, we consider the effectiveness of our fixed effects approach to controlling for unobserved heterogeneity. Underlying the regression analysis presented in Tables 5a through 6b is the assumption that the inclusion of school district and/or zip code fixed effects adequately controls for the presence of unobserved confounding variables. While the inclusion of these controls is an improvement over the existing literature, it remains an open question as to whether or not these fixed effects provide adequate control for unobservable covariates. To help evaluate the effectiveness of these controls, Table 7 compares the predicted treatment effects from a model that includes as controls only zip code fixed effects to a model that includes all of our controls in addition to the zip code fixed effect. If the zip code fixed effect is successfully controlling for unobserved confounding factors, one might expect it to also effectively control for the impact of our observed spatially varying covariates. Therefore, if the treatment effects vary little between a model which includes only the zip code fixed effects and a model that includes zip code fixed effects and the observable controls, one might have more faith in the ability of these fixed effects to control for spatially varying unobservables. Table 7 presents 95% confidence interval for the three treatment effects under these two different models. As the table shows, the zip code fixed effects appear to control quite effectively for the observable covariates, with point estimates close and confidence intervals overlapping.

Propensity Score Matching

As noted above, we find evidence of non-linear migratory responses to baseline racial composition, suggesting it may be difficult to control for these effects parametrically. These controls are important, since the estimated composition effects are sensitive to the presence of such controls. (Note, for example, the large changes in point estimates from the "No Controls" to "Basic Controls" models in Tables 6a and 6b).

To better account for this non-linearity and uncertainty about functional form, we also employ a propensity score matching model approach to non-parametrically match communities receiving a TRI "treatment" to control communities ("un-treated") with similar observable characteristics, and then compare their migration patterns. Under controlled experiments, a treatment is given randomly so that, by design, the expected values of unobserved variables are the same in the treatment and control groups. Under a natural experiment, treatment and non-treatment observations are grouped by other observed variables, and compared conditional on those variables. In our case, the three treatments are the presence of baseline TRI exposure among the set of communities that do not change exposure status over time; the move to exposure among those communities that did not experience baseline exposure; and the ending of exposure among those communities exposed in the baseline. These three treatment definitions mirror the estimated treatment effects from the simple liner models presented in tables 5 and 6.

Under the standard matching model, observations are grouped by values of the observables (baseline racial composition, density, and other locational amenities or proxies) and, within each cell, differences in migration between treated and untreated observations are calculated. However, the number of cells required to do this can be prohibitively large. Rosenbaum and Rubin (1983) showed that, when a large number of observed variables create too many empty cells, one can instead match on the propensity score, that is, the *probability* that a community receives the TRI treatment, measured as a function of observed variables. This reduces the matching to a single dimension.

This approach relaxes the need for functional form assumptions about the controlling variables. Further, it weakens the necessary assumptions regarding the error term, requiring only that, conditional on the observables, the expected value of the error term is equal across treatments cases. This is in contrast to the classical assumption that conditional on the observables (including the treatment variable) the expected value of the error term is zero. One limitation is that this approach is only valid over the range of observable data for which both treatments and non-treatments are observed (the "region of overlapping support").[20]

[20] See Heckman, Ichimura, and Todd (1997, 1998) and Dehejia and Wahba (2002). Greenstone (2004) has recently applied this approach to air quality changes.

We estimate the propensity scores for each type of treatment using the full set of controls from the regression models. Specifically, among those communities who experienced exposure in either both 1988-1990 and 1998-2000 or neither 1988-1990 nor 1998-2000, we estimated a probit model on the probability of experiencing exposure (the "baseline" treatment). Among those who experienced no exposure in 1988-1990, we estimated a probit model on the probability of gaining exposure in 1998-2000 (the "new" treatment). Among those who did experience exposure in 1988-1990, we estimated a probit model on the probability of not having exposure in 1998-2000 (the "exit" treatment). The probits included the same controls as the above regressions plus school district fixed effects.

For each community receiving a TRI treatment, we compare its scale and composition effects to the average among all control communities within a 1-percentile caliper as measured by the predicted probability of receiving the treatment. In addition to the simple 1-percentile caliper, we also estimate the treatment effect by further restricting the match to those control communities within the same school district as the treatment community. Like the local fixed effects regression models, the latter restriction is likely to dilute the comparison between treatment and control communities, if our half-mile buffers are too small and the TRI facilities have broader impacts within the school district.

The 1-percent caliper provides a good number of matches. As shown in Table 8, it provides at least one match for approximately 98.0% of baseline treatment communities, with a median of 125 matches and a mean of 768 matches among those communities which are matched. Similarly, it provided at least one match for 99.3% of new treatment communities (with a median of 941 matches) and 98.7% of exit treatment communities (with a median of 43 matches). While there are fewer matches once the school district filter is applied, the matching results (again summarized in Table 8) are still strong. The 1-percent caliper is a compromise between the simple nearest neighbor match, which minimizes bias by taking only the very closest match, but which is inefficient because in ignores the information in other reasonably matched controls, and a wider caliper, which may be more efficient but introduces the risk of more bias by making less defensible comparisons.

Moreover, our 1-percent caliper match creates a "balanced panel," with equal expected values for all observables in the treatment and control groups. This is a necessary condition for a properly functioning propensity score match, since the theory predicts that all observables (and unobservables) are orthogonal to the treatment, after conditioning on the propensity score. In contrast, our sensitivity analyses using nearest neighbor and 5-percent caliper matches created unbalanced panels. Table 9 compares the mean of key observables between matched pairs (treatment, non-treatment) for the three different treatments under the 1-percent caliper match and the 1-percent caliper match with the school district filter. Using the simple 1-percent caliper match, for each treatment, we cannot reject the hypothesis of equal mean values of observables between treatment and control observations. For the 1-percent match within school districts, we similarly fail to reject the hypothesis under a test of equal means ignoring any covariance; under a test of mean differences equal to zero, we still fail to reject in most cases. This is a strong indication that the matching model is functioning properly.

The last two rows in each section of Tables 5 and 6 provide our results from this propensity score matching. The simple 1-percent caliper match is presented first, followed by the 1-percent caliper match with the school district filter imposed. As shown in the table, the estimated effects of TRI sites on migration are qualitatively consistent with the estimated effects from our regression models using more controls, and in many cases almost identical quantitatively. These results provide more evidence of the scale effects predicted by the model, and for the presence of composition effects posited by Been (1994, 1997) in the unweighted baseline treatment. As with the regression models, the composition effects generally have the expected sign but are statistically insignificant for the other treatments.

VI. Conclusions

Tiebout's suggestion that people vote with their feet to find the community that provides their optimal tax/public goods pair has played a central role in the theory of local public finance. More recently, a new and growing empirical literature is leveraging differences in neighborhood composition combined with the "Tieboutvian" notion that households are in a locational equilibrium to identify household preferences for spatially delineated public goods. Tiebout's original work has given rise to a large empirical literature that tests with varying results the numerous normative and positive hypothesis that have sprung from this seminal work. While this empirical literature is both deep and varied, less attention has been paid to the implications of locational equilibrium assumptions on the role that changes in local public goods can have in changing community compositions. Given the central role that this link between public good levels and community composition plays in estimating preferences for public goods in the context of locational equilibrium models, further direct investigation of this link has gained importance.

Toward this end, we use changes in the emissions of toxic air pollutants across spatially delineated "neighborhoods" to test for "environmental gentrification"—environmentally motivated migration patterns that differ systematically by income or minority status. Our analysis follows on a small number of "environmental justice" studies that have explored the link between changes in environmental quality and the prevalence of minority groups. Using a new approach to community definition that overcomes the problems associated with the use of Census tracts in conjunction with better controls for potentially confounding factors than have been used in previous studies, we provide the strongest evidence to date of the link between changes in environmental quality and local changes in community composition. We find strong evidence of scale effects of a magnitude that is both statistically significant and empirically relevant. We also find evidence of composition effects that suggests that pollution in a given location is associated with the out-migration of Whites and the in-migration of Hispanics. While we believe the un-weighted model to be the appropriate approach to testing for Environmental Gentrification, it should be noted that unlike the case with scale effects, the composition effects are weaker when the estimation is done using location specific population weights.

We believe these results to be relevant on two fronts. First, they provide direct evidence in support of migration patterns that are consistent with Tiebout's notion of households "voting with their feet." Despite its importance for a host of theoretical and applied work in local public finance, the Tiebout model surprisingly has not been well-tested in this way previously, at least for local amenities. Second, they are relevant to the literature on "environmental justice" because we provide the strongest evidence to date that correlations between race and environmental quality are likely to be at least in part the result of "Tiebout Sorting."

This finding has potentially profound implications for environmental policy. In a world where households sort in response to changes in environmental quality, the bulk of the benefits of a policy that successfully cleans up dirtier neighborhoods where the poor live may actually be captured by rich households. As the neighborhood amenity improves, wealthier households will move in, driving up rents. If the poor do not own their homes, landlords would capture the capital appreciation of the local housing, while the poor pay higher rents. This "environmental gentrification" may actually more than offset the direct gain of the environmental improvement, so that the original residents are actually worse off. Such outcomes have been demonstrated in simulation models of air quality improvements in Los Angeles (Sieg et al. 2004) and increases in protection of open space in Raleigh, NC (Walsh 2004).

Figure 1. Density of income for two household types and community income boundary.

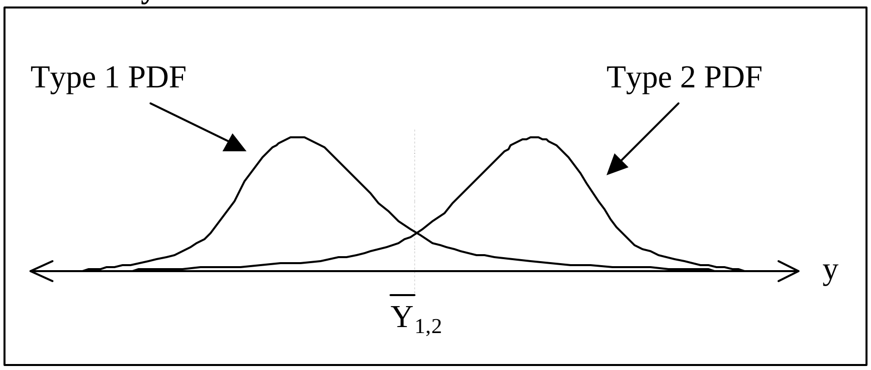

Figure 2. Shift in community income boundary after improvement in G_1.

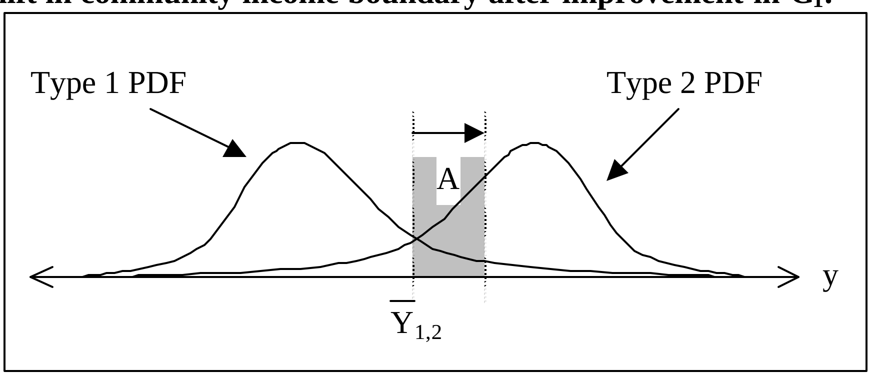

Figure 3. Sorting of Type 1 and Type 2 households into communities, with stratification by income and tastes.

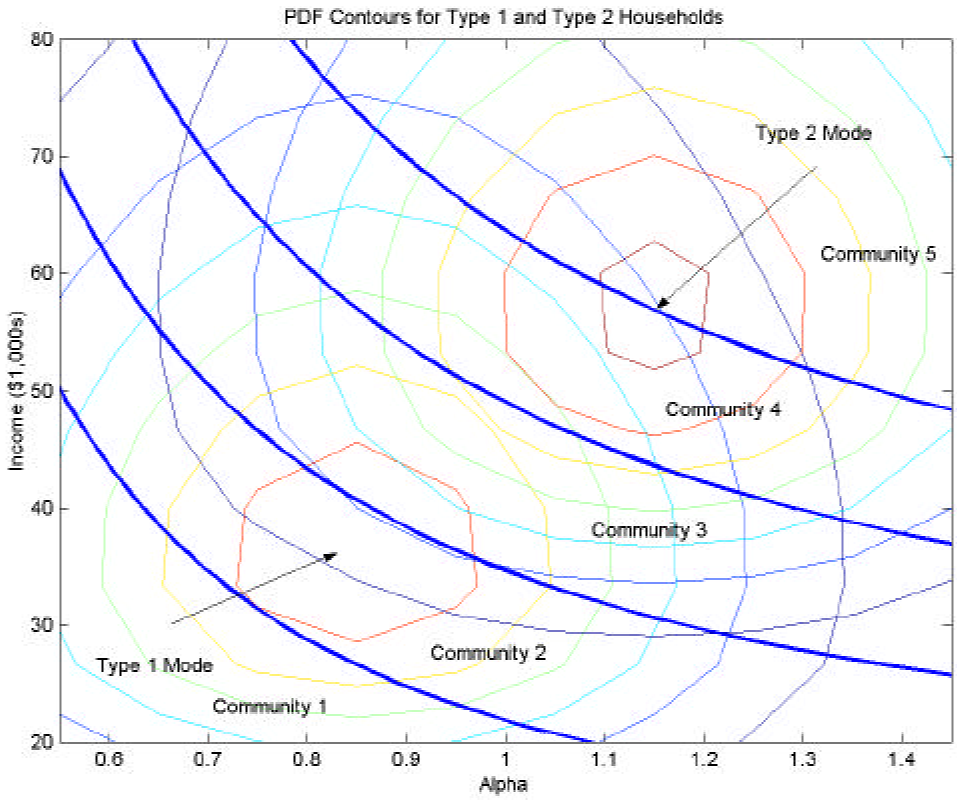

Figure 4. Shift in community income-taste loci after improvement in G_1.

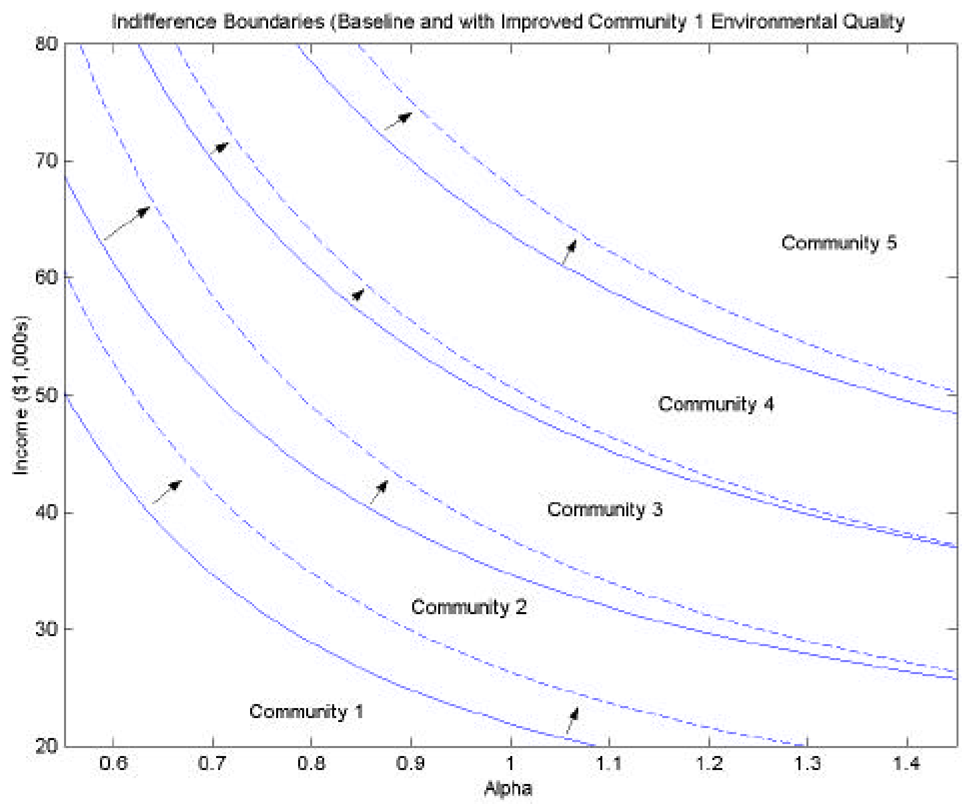

Figure 5.

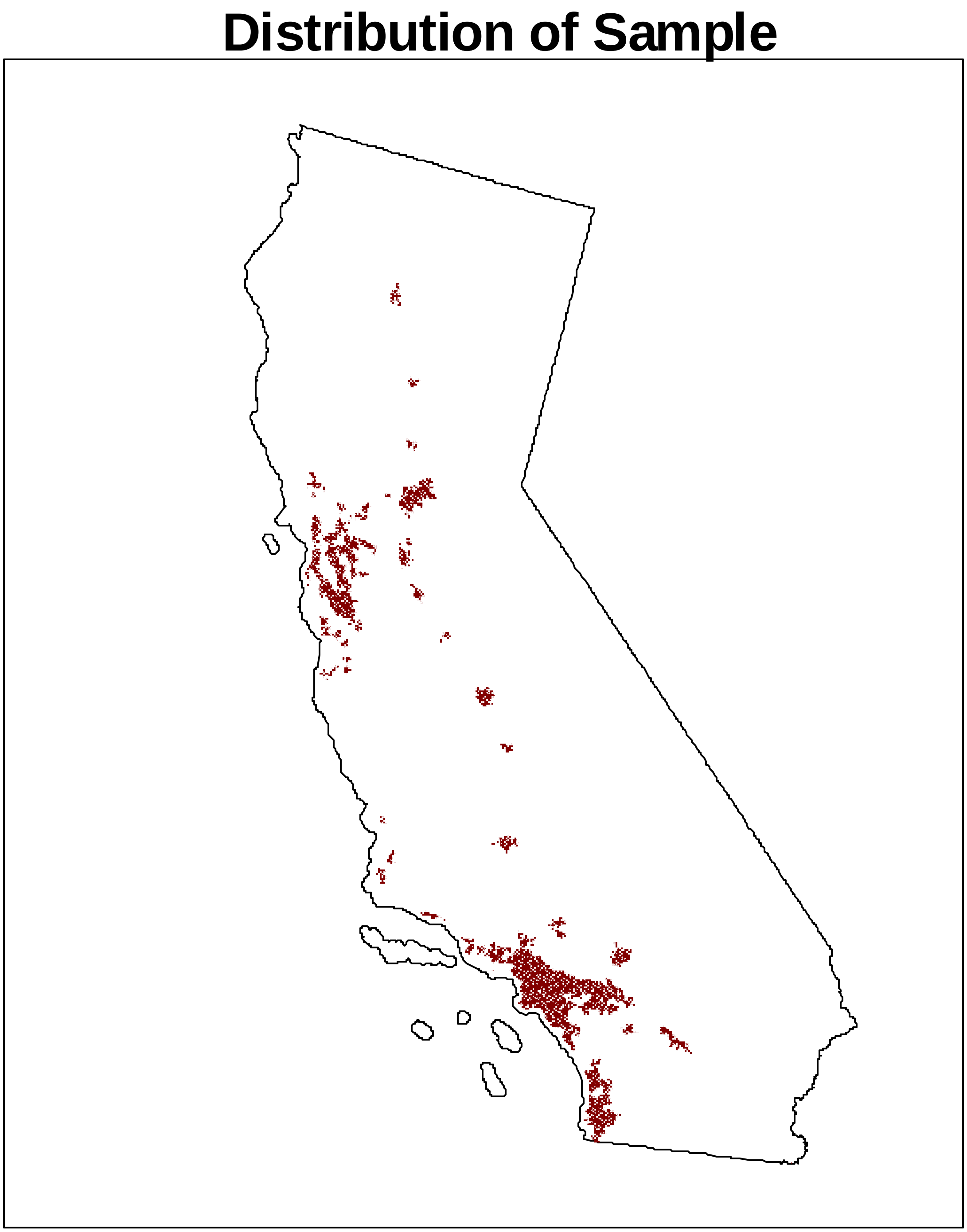

Figure 6.

Mapping TRI Site Emmisions to Neighborhoods

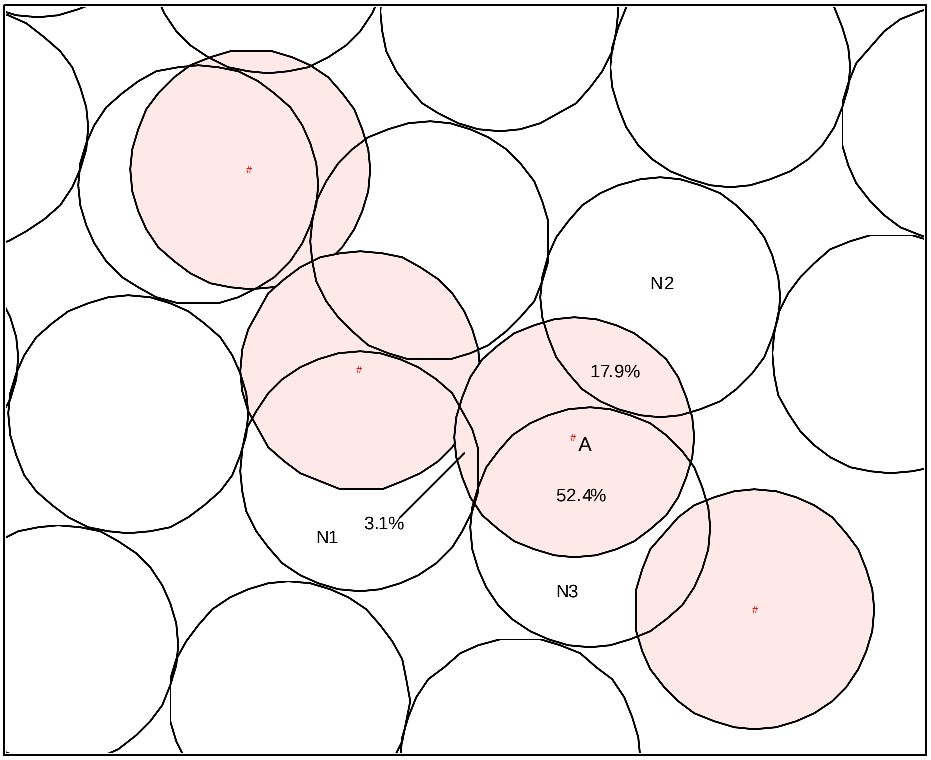

Table 1. Simulated Community Demographics Before and After Improvements to G_1.

	Ex Ante Demographics				Ex Post Demographics				Change			
Community	Population	Average Income	Type 1 Share	Type 2 Share	Population	Average Income	Type 1 Share	Type 2 Share	Population	Average Income	Type 1 Share	Type 2 Share
1	8.83	$28,126	97.06%	2.94%	15.88	$31,859	94.14%	5.86%	7.05	$3,733	-2.91%	2.91%
2	18.04	$37,345	85.31%	14.69%	14.28	$40,513	80.04%	19.89%	-3.76	$3,168	-5.27%	5.20%
3	22.52	$46,152	58.39%	41.61%	20.61	$47,155	56.04%	43.96%	-1.91	$1,003	-2.35%	2.35%
4	24.75	$56,799	34.83%	65.17%	23.89	$56,252	33.24%	66.76%	-0.86	-$547	-1.59%	1.59%
5	25.86	$68,857	16.51%	83.49%	25.34	$68,968	16.26%	83.70%	-0.52	$111	-0.25%	0.21%

Table 2. Assignment of Census Blocks and TRI Emissions to Circle-Communities

	Half Mile Circles	One Mile Circles
Count	25,166.00	6,218.00
1990 Blocks per Circle		
25th percentile	4	11
50th percentile	10	29
75th percentile	19	55
Max	132	383
2000 Blocks per Circle		
25th percentile	6	17
50th percentile	13	38
75th percentile	22	64
Max	136	408
Circles with TRI Exposure		
1/4 Mile Buffer	3,109	1,295
1/2 Mile Buffer	5,179	1,795
# TRI Sites for Exposed Circles		
1/4 Mile Buffer		
25th percentile	1	1
50th percentile	2	2
75th percentile	3	4
Max	19	25
1/2 Mile Buffer		
25th percentile	1	1
50th percentile	2	2
75th percentile	4	5
Max	27	34
Circles per School District		
25th percentile	45	14
50th percentile	93.5	27
75th percentile	169	47
Max	2,352	620
Circles per Zip Code		
25th percentile	11	3
50th percentile	21	6
75th percentile	35	9
Max	190	49

Table 3. Descriptive Statistics of the Data, for half-mile circle communities

Baseline Demographic Data (1990)	**Mean**	**Standard Deviation**
Population (Density)	772	930
Share Black	0.05	0.11
Share Hispanic	0.19	0.20
Share Asian	0.08	0.10
Share Other Minority	0.01	0.02
Pct Households with single-parent families	0.08	0.07
Mean Rental Rate ($)	689	263
Mean Housing Value ($)	229,872	138,199
Share Owning their Home	0.66	0.27
Pct Employed	0.94	0.05
Pct of Employed in Manufacturing, if Employed	0.15	0.08
Pct Not Graduating from High School	0.10	0.07
Pct with Bachelor degree	0.49	0.14
Median Household Income ($)	46,461	21,551
Changes in Demographics (1990-2000)		
Population	92	256
Percentage Change in Population	0.09	0.67
Share Non-Hispanic White	-0.09	0.13
Share Black	0.01	0.06
Share Hispanic	0.05	0.11
Share Asian	0.03	0.08
Change in Percentage of Single-Parent Households	0.01	0.06
TRI Data		
Share with Baseline TRI Exposure (1988-1990)	0.10	NA
Share with new TRI Exposure (1998-2000)	0.01	NA
Share Losing TRI Exposure (1998-2000)	0.04	NA
Baseline Emissions	300,714	4,718,020
Baseline Emissions, among those Exposed	3,006,542	1.46e7
Locational Data		
1990 FBI Crime Index	0.08	0.28
Change in Crime Index	-0.03	0.14
Distance to Coast	47.2	45.3
Degrees Latitude	35.41	2.03
School or Zip Code Fixed Effects	NA	NA

Table 4a Baseline Exposure as function of Demographic Composition (Tobit Models)

	Model 1	Model 2	Model 3
Share Black	16201 ***	19807 ***	16672 ***
Share Hispanic	28457 ***	33272 ***	31053 ***
Share Asian	10870 ***	14228 ***	2151 **
Population Density	--	-4 ***	-6 ***
KM to Coast	--	-66 ***	--
Degrees Latitude	--	216	--
Fixed Effects	No	No	School District
Constant	-38750 ***	-41569 ***	-38952 ***
Pseudo-R^2	0.02	0.03	0.06

Dependent Variable: 1990 Hazard-Weighted TRI Exposures.
*Significant at 10%; **Significant at 5%; ***Significant at 1%.

Table 4b Baseline Composition as function of Exposure (OLS).

	Dependent Variable: Share White	Dependent Variable: Share Black	Dependent Variable: Share Hispanic	Dependent Variable: Share Asian
Presence of TRI Site	-0.128 ***	0.006	0.127 ***	-0.006 *
Toxicity-Weighted TRI Emissions	-3.1e-6 ***	2.0e-8	3.8e-6 ***	-7.0e-7 ***
Density	-6.2e-5 ***	2.1e-6	5.2e-5 ***	7.0e-6 ***
R^2	0.54	0.33	0.51	0.40

All Models contain School District Fixed Effects.
*Significant at 10%; **Significant at 5%; ***Significant at 1%.

Table 5a. Estimated Scale Effects: Unweighted.

	Average Effect of Baseline TRI Exposure	Average Effect of New TRI Exposure	Average Effect of Exiting TRI Exposure	R^2
Population Levels				
-No Controls	-30***	-13	43***	0.00
-Basic Controls	-54***	-35**	39***	0.07
-School District FE	-59***	-35**	42***	0.11
-Zip Code FE	-71***	-36**	45***	0.26
-Propensity Score Matching	-24***	-38***	45***	--
-Matching within School Dist	-26***	-28**	29**	--
% Change in Population				
-No Controls	-15.6***	-5.3	7.1**	0.00
-Basic Controls	-11.7***	-7.3**	5.0**	0.04
-School District FE	-10.3***	-8.3**	6.1**	0.09
-Zip Code FE	-12.0***	-9.3***	6.3***	0.19
-Propensity Score Matching	-5.0**	-10.0**	4.5**	--
-Matching within School Dist	-6.1***	-11.2**	4.2	--

*Significant at 10%; **Significant at 5%; ***Significant at 1%.
See Equation 8 for definition of the treatment effects.
Standard Errors for Propensity Score Matching models based on bootstrap with 300 draws.

Table 5b. Estimated Scale Effects: Population-Weighted.

	Average Effect of Baseline TRI Exposure	Average Effect of New TRI Exposure	Average Effect of Exiting TRI Exposure	R^2
Population Levels				
-No Controls	-46*	-18	81***	0.00
-Basic Controls	-81***	-39***	71***	0.18
-School District FE	-84***	-31	78***	0.25
-Zip Code FE	-108***	-42**	78***	0.58
-Propensity Score Matching	-43**	-95***	86***	--
-Matching within School Dist	-11	-26*	88***	--
% Change in Population				
-No Controls	-2.6	0.8	3.0**	0.00
-Basic Controls	-3.6**	-0.7	2.6*	0.05
-School District FE	-4.0***	-1.0	3.0**	0.10
-Zip Code FE	-4.7***	-1.6	2.9***	0.24
-Propensity Score Matching	-2.2**	-2.6**	3.6***	--
-Matching within School Dist	-1.1	-0.8	6.8***	--

*Significant at 10%; **Significant at 5%; ***Significant at 1%.
See Equation 8 for definition of the treatment effects.
Standard Errors for Propensity Score Matching models based on bootstrap with 300 draws.

Table 6a. Estimated Composition Effects: Unweighted

	Average Effect of Baseline TRI Exposure	Average Effect of New TRI Exposure	Average Effect of Exiting TRI Exposure	R^2
Share White				
-No Controls	-0.9	-1.9**	-1.0	0.00
-Basic Controls	-1.8***	-1.6**	0.4	0.21
-School District FE	-1.1*	-1.0	0.1	0.35
-Zip Code FE	-1.3**	-1.6**	0.3	0.45
-Propensity Score Matching	-2.1***	-0.5	0.2	--
-Matching within School Dist	-1.3**	-0.7	-0.1	--
Share Black				
-No Controls	-0.8**	1.0**	0.1	0.00
-Basic Controls	0.4*	0.7**	-0.1	0.17
-School District FE	0.0	0.4	0.0	0.29
-Zip Code FE	0.0	0.5	0.2	0.39
-Propensity Score Matching	0.4	0.5	0.2	--
-Matching within School Dist	0.4	1.1*	-0.3	--
Share Hispanic				
-No Controls	1.2**	1.8**	0.3	0.00
-Basic Controls	1.5***	1.0	-0.6	0.23
-School District FE	1.3***	0.7	-0.6	0.35
-Zip Code FE	1.4***	1.1*	-1.0**	0.44
-Propensity Score Matching	2.0***	0.3	-0.6	--
-Matching within School Dist	0.9*	-0.3	-0.6	--
Share Asian				
-No Controls	0.7*	-0.7	0.5	0.00
-Basic Controls	0.0	-0.0	0.4	0.11
-School District FE	-0.0	-0.0	0.5	0.31
-Zip Code FE	-0.1	0.2	0.5*	0.41
-Propensity Score Matching	-0.2	-0.3	0.3	--
-Matching within School Dist	0.0	-0.0	1.0**	--

*Significant at 10%; **Significant at 5%; ***Significant at 1%.
See Equation 8 for definition of the treatment effects.
Standard Errors for Propensity Score Matching models based on bootstrap with 300 draws.

Table 6b. Estimated Composition Effects: Population-Weighted

	Average Effect of Baseline TRI Exposure	Average Effect of New TRI Exposure	Average Effect of Exiting TRI Exposure	R^2
Share White				
-No Controls	1.3***	-0.0	-1.1***	0.00
-Basic Controls	0.3	-0.2	-0.0	0.36
-School District FE	0.0	-0.4	0.2	0.52
-Zip Code FE	-0.5*	-1.0***	0.3	0.68
-Propensity Score Matching	-0.3*	0.8	0.3	--
-Matching within School Dist	0.0	-0.3	1.5***	--
Share Black				
-No Controls	-1.8***	-0.8**	0.5***	0.00
-Basic Controls	0.1	0.4	0.1	0.42
-School District FE	0.1	0.7***	-0.0	0.56
-Zip Code FE	0.3**	0.6***	-0.0	0.68
-Propensity Score Matching	1.0***	0.4	-0.5***	--
-Matching within School Dist	0.3	1.2***	-1.2***	--
Share Hispanic				
-No Controls	1.2***	2.8***	0.6*	0.00
-Basic Controls	-0.2	0.7	0.6**	0.39
-School District FE	-0.1	0.4	0.3	0.52
-Zip Code FE	-0.0	0.7**	0.2	0.65
-Propensity Score Matching	-0.6	-0.1	0.8***	--
-Matching within School Dist	-0.0	0.2	0.6	--
Share Asian				
-No Controls	-0.7**	-1.9***	-0.0	0.00
-Basic Controls	-0.3	-0.9***	-0.6***	0.28
-School District FE	-0.1	-0.7**	-0.5***	0.49
-Zip Code FE	0.1	-0.4	-0.4**	0.63
-Propensity Score Matching	-0.1	-1.1***	-0.6***	--
-Matching within School Dist	-0.2	-1.2***	-0.8***	--

*Significant at 10%; **Significant at 5%; ***Significant at 1%.
See Equation 8 for definition of the treatment effects.
Standard Errors for Propensity Score Matching models based on bootstrap with 300 draws.

Table 7. Sensitivity of zip-code fixed effect models to inclusion of other observables: 95% confidence intervals for each case.

Outcome Variable	Population Weighted?	Base Treatment Without Controls	Base Treatment With Controls	New Treatment Without Controls	New Treatment With Controls	Exiting Treatment Without Controls	Exiting Treatment With Controls
?Pop	Y	(-125, -64)	(-136, -79)	(-75, 9)	(-81, -3)	(56, 110)	(53, 103)
	N	(-99, -56)	(-97, -46)	(-68, -7)	(-69, -4)	(33, 76)	(22, 67)
%?Pop	Y	(-.060, -.015)	(-.070, -.025)	(-0.41, .022)	(-.047, .015)	(.012, .052)	(.009, .049)
	N	(-.212, -.090)	(-.179, -.061)	(-.159, -.001)	(-.169, -.018)	(.033, .145)	(.010, .116)
?%White	Y	(-.005, .007)	(-.010, .001	(-.006, .011)	(-.017, -.002)	(-.009, .002)	(-.002, .008)
	N	(-.010, .015)	(-.023, -.002)	(-.017, .014)	(-.030, -.003)	(-.021, .001)	(-.007, .012)

Table 8. Matches in the Propensity Score Matching Estimator

Treatment	Restricted to within school district?	# / Pct of Treatment Communities Successfully Matched	Median Number of Matched Controls, if matched	Mean Number of Matched Controls, if matched
Base	N	1442 / 98.0%	125	768
Base	Y	1052 / 71.0%	5	22
New	N	274 / 99.3%	941	1626
New	Y	245 / 88.8%	20	58
Exit	N	1144 / 98.7%	43	40
Exit	Y	648 / 55.9%	2	4

Table 9. P-values for test of balanced panels (equality of key observables between treatment and control communities).

Panel I. Matches within all of study area

	After Matching (Base Treatment)		After Matching (New Treatment)		After Matching (Exit Treatment)	
Observable	Equal Means	Mean Differences	Equal Means	Mean Differences	Equal Means	Mean Differences
Share Black	0.70	0.69	0.83	0.83	0.90	0.90
Share Hispanic	0.92	0.90	0.86	0.84	0.90	0.89
Share Asian	0.24	0.23	0.91	0.91	0.44	0.43
Pop Density	0.42	0.39	0.68	0.67	0.67	0.66

Panel II. Matches restricted to within school districts.

	After Matching (Base Treatment)		After Matching (New Treatment)		After Matching (Exit Treatment)	
Observable	Equal Means	Mean Differences	Equal Means	Mean Differences	Equal Means	Mean Differences
Share Black	0.61	0.50	0.30	0.10	0.62	0.51
Share Hispanic	0.41	0.19	0.42	0.16	0.73	0.05
Share Asian	0.33	0.23	0.94	0.92	0.73	0.67
Pop Density	0.17	0.08	0.58	0.38	0.17	0.08

"Equal Means" refers to test of equality of means for treatment and control distributions ignoring covariance—that is, for t=treatment case and c=control case, $\sigma = \sqrt{s_t^2 / n_t + s_c^2 / n_c}$. "Mean Differences" refers to test of paired differences equal to zero (i.e. $\sigma = s_\Delta$).

References

Anderton, Douglas L., Andy B. Anderson, John Michael Oakes, and Michel R. Fraser. 1994. "Environmental Equity: The Demographics of Dumping." *Demography* 31(2): 229-48.

Arora, Seema, and Timothy N. Cason. 1999. "Do Community Characteristics Influence Environmental Outcomes? Evidence from the Toxics Release Inventory." *Southern Economic Journal* 65(4): 691-716.

Ash, Michael, and T. Robert Fetter. 2002. "Who Lives on the Wrong Side of the Environmental Tracks? Evidence from the EPA's Risk-Screening Environmental Indicators Model." Working Paper #50, Political Economy Research Institute, U of Massachusetts at Amherst.

Bayer, Patrick, Robert McMillan, and Kim Reuben. 2005. "An Equilibrium Model of Sorting in an Urban Housing Market." Mimeo. New Haven, CT: Yale University.

Been, Vicki. 1994. "Locally Undesirable Land Uses in Minority Neighborhoods: Disproportionate Siting or Market Dynamics?" *Yale Law Journal* 103: 1383-1422.

Been, Vicki. 1997. "Coming to the Nuisance or Going to the Barrios? A Longitudinal Analysis of Environmental Justice Claims." With Francis Gupta. *Ecology Law Quarterly* 34: 3-56.

Brooks and Sethi. 1997. "The Distribution of Pollution: Community Characteristics and Exposure to Air Toxics." *JEEM* 32: 233-250.

Brueckner, Jan K. 1982. "A Test for Allocative Efficiency in the Local Public Sector." *Journal of Public Economics* 14: 311-32.

Cameron, Trudy Ann, and Graham D. Crawford. 2004. "Superfund Taint and Neighborhood Change: Ethnicity, Age Distributions, and Household Structure." Mimeo.

Daniels, Glynis and Samantha Friedman. 1999. "Spatial Inequality and the Distribution of Industrial Toxic Releases: Evidence From the 1990 TRI." *Social Science Quarterly* 80(2): 244-62.

Dehejia, Rajeev H., and Sadek Wahba. 2002. "Propensity Score-Matching Methods for Nonexperimental Causal Studies." *Review of Economics and Statistics* 84(1): 151-161.

Epple, Dennis, Radu Filimon, and Thomas Romer. 1984. *Journal of Public Economics* 24: 281-308.

Epple, Dennis, and Holger Sieg. 1999. "Estimating Equilibrium Models of Local Jurisdictions." *Journal of Political Economy* 107: 645-681.

Ferreyra, Maria M. 2005. "Estimating the Effects of Private School Vouchers in Multi-District Economies." Mimeo. Pittsburgh, PA: Carnegie Mellon University.

Goldman, Benjamin A., and Laura Fitton. *Toxic Wastes and Race Revisited*. Washington, DC: Center for Policy Alternatives.

Gramlich, Edward M., and Daniel L. Rubinfeld. 1982. "Micro Estimates of Public Spending Demand Functions and Tests of the Tiebout and Median-Voter Hypotheses." *Journal of Political Economy* 90(3): 536-560.

Graves, Philip E., and Donald M. Waldman. 1991. "Multimarket Amenity Compensation and the Behavior of the Elderly." *American Economic Review* 81(5): 1374-1381.

Greenstone, Michael. 2004. "Did the Clean Air Act Amendments Cause the Remarkable Decline in Sulfur Dioxide Concentrations?" *Journal of Environmental Economics*, forthcoming.

Hamilton, James T. 1995. "Testing for Environmental Racism: Prejudice, Profits, Political Power?" *Journal of Policy Analysis and Management* 14: 107-132.

Heckman, James J., Hidehiko Ichimura, and Petra Todd. 1997. Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program. *Review of Economic Studies* 64: 605-654.

Heckman, James J., Hidehiko Ichimura, and Petra Todd. 1998. Matching as an Econometric Evaluation Estimator. *Review of Economic Studies* 65: 261-294.

Hersh, Robert. 1995. "Race and Industrial Hazards: A Historical Geography of the Pittsburgh Region, 1900-1990." RFF Discussion Paper 95-18. Washington, DC: Resources for the Future.

Kahn, Matthew E. 2000. "Smog Reduction's Impact on California County Growth." *Journal of Regional Science* 40(3): 565-582.

Kriesel, Warren, Terence J. Centner, and Andrew G. Keller. 1996. "Neighborhood Exposure to Toxic Releases: Are there Racial Inequities?" *Growth and Change* 27: 479-99.

Mohai, Paul, and Bunyan Bryant 1992. "Environmental Racism: Reviewing the Evidence." In *Race and the Incidence of Environmental Hazards: A Time for Discourse*, edited by Bunyan Bryant and Paul Mohai. Boulder, CO: Westview Press.

Morello-Frosch, Rachel, Manuel Pastor, and James Sadd. 2001. "Environmental Justice and Southern California's 'Riskscape': The Distribution of Air Toxics Exposures and Health Risks Among Diverse Communities." *Urban Affairs Review* 36(4): 551-578.

Oates, Wallace E. 1969. "The Effects of Property Taxes and Local Public Spending on Property Values: An Empirical Study of Tax Capitalization and the Tiebout Hypothesis." *Journal of Political Economy* 77(6): 957-971.

Oates, Wallace E. 2005. “The Many Faces of the Tiebout Model.” Working Paper.

Rinquist, Evan. 1997. "Equity and Distribution of Environmental Risk: The Case of TRI Facilities." *Social Science Quarterly* 78: 811-829.

Rosenbaum, P. and D. Rubin. 1983. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* 70: 41-55.

Rubinfeld, Daniel L. 1987. "The Economics of the Local Public Sector." In, *Handbook of Public Economics*, ed. by Alan J. Auerbach and Martin Feldstein. Amsterdam: North-Holland, pp. 571-645.

Sadd, James L., Manuel Pastor, J. Thomas Boer, and Lori D. Snyder. 1999. "'Every Breath You Take…:' The Demographics of Toxic Air Releases in Southern California." *Economic Development Quarterly* 13(2): 107-123.

Sieg, Holger, V. Kerry Smith, H. Spencer Banzhaf, and Randy Walsh. 2004. "Estimating the General Equilibrium Benefits of Large Changes in Spatially Delineated Public Goods." *International Economic Review* 45(4): 1047-1077.

Tiebout, Charles. 1956. "A Pure Theory of Local Expenditures." *Journal of Political Economy* 64: 416-424.

Timmins, Christopher. 2003. "If You Can't Take the Heat, Get Out of the Cerrado... Recovering the Equilibrium Amenity Cost of Non-Marginal Climate Change in Brazil." Mimeo. Durham, NC: Duke University.

United Church of Christ. 1987. *Toxic Wastes and Race in the United States*. Report of the Commission for Racial Justice.

Walsh, Randall P. 2004. "Analyzing Open Space Policies in a Locational Equilibrium Model with Endogenous Landscape Amenities." Mimeo. Boulder, CO: University of Colorado.

Wolverton, Ann. 2002. "Does Race Matter? An Examination of a Polluting Plant's Location Decision." Mimeo. Washington, DC: National Center for Environmental Economics, US EPA.

Wolverton, Ann. 2003. "The Demographic Distribution of Pollution: Does Neighborhood Composition Affect Plant Pollution Behavior?" Mimeo. Washington, DC: National Center for Environmental Economics, US EPA.