

## Do People Vote with Their Feet? An Empirical Test of Tiebout's Mechanism

By H. SPENCER BANZHAF AND RANDALL P. WALSH\*

*Charles Tiebout's suggestion that people "vote with their feet" for communities with optimal bundles of taxes and public goods has played a central role in local public finance for over 50 years. Using a locational equilibrium model, we derive formal tests of his premise. The model predicts increased population density in neighborhoods experiencing exogenous improvements in public goods and, for large improvements, increased relative mean incomes. We test these hypotheses in the context of changing air quality. Our results provide strong empirical support for the notion that households "vote with their feet" for environmental quality. (JEL H41, H73, Q53, R23)*

Charles Tiebout's (1956) suggestion that people "vote with their feet" to find the community that provides their optimal bundle of taxes and public goods 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 the Tiebout model to identify general equilibrium models of household sorting (e.g., Patrick J. Bayer and Christopher Timmins 2007; Maria Marta Ferreyra 2007; Holger Sieg et al. 2004).

Given the central importance of Tiebout's insights, there have been surprisingly few direct tests of his premise. Existing tests of the Tiebout model can be grouped into two broad categories. Indirect or implicit tests, the most common, have focused on deductive implications of the model. For example, Wallace E. Oates's (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. Jan K. 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 model 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., Edward M. Gramlich and Daniel L. Rubinfeld 1982; Dennis Epple and Sieg 1999, Paul W. Rhode and Koleman S. Strumpf 2003).

Direct tests of actual migratory responses to public good provision—Tiebout's mechanism of people voting with their feet—have been less common. In this paper, we first provide a theoretical

\* Banzhaf: Department of Economics, Georgia State University, PO Box 3992, Atlanta, GA 30302 (e-mail: [hsbanzhaf@gsu.edu](mailto:hsbanzhaf@gsu.edu)); Walsh: Department of Economics and Institute of Behavioral Sciences, University of Colorado at Boulder, 256 UCB, Boulder, CO 80309 (e-mail: [walshr@colorado.edu](mailto:walshr@colorado.edu)). Support for this research was provided by the National Science Foundation (SES-03-21566). We thank Lisa Crooks and Joshua Sidon for excellent research assistance. We also appreciate the helpful comments of Patrick Bayer, Lori Benneer, Trudy Ann Cameron, Nicholas Flores, Michael Greenstone, Michael Margolis, Kenneth McConnell, Wallace Oates, David Simpson, Kerry Smith, Christopher Timmins, Kip Viscusi, Robertson Williams, Ann Wolverton, and two anonymous referees. We also thank participants of the 2004 AERE Workshop, the 2004 CU Environmental and Resource Economics Workshop, the 2005 ASSA meetings, the 2005 NBER Summer Institute, the 2005 Occasional California Workshop on Environmental and Natural Resource Economics, and seminar participants at Duke University, Georgia State University, Rice University, Tufts University, the University of Central Florida, the University of Maryland, the University of Minnesota, and the University of Wisconsin.

model that predicts relative increases in population density for neighborhoods that experience an exogenous marginal improvement in public goods. For discrete improvements, the model predicts increased relative population density in “most” cases and relative increases in mean neighborhood income for “large” improvements (we clarify “most” and “large” below). To test these effects, we use a difference-in-difference model, identifying the impact of entry and exit of facilities that are required to report their releases of chemicals for the Toxics Release Inventory (TRI), as well as changes in toxicity-weighted emissions levels, on population and demographic composition. We also test for lagged responses to differing baseline exposure levels, which were not publicly announced until 1989. As our unit of analysis, we use a set of “communities” defined randomly in space by equally spaced half-mile circles. We control for demographics and other location-specific effects using linear regression and, because migratory responses may be highly nonlinear functions of demographics, we provide additional analysis using a bias-corrected non-parametric matching estimator.

Consistent with the Tiebout model, we find clear evidence of migration correlated with TRI facility emissions and their arrival to, or exit from, a community. Furthermore, we find evidence that TRI facilities cause the composition of a community to become poorer over time. Thus, we find support for the fundamental mechanism underlying the Tiebout model: households do appear to vote with their feet in response to changes in public goods.

Our paper is not the first to look at such direct tests of Tiebout’s mechanism. Philip E. Graves and Donald M. Waldman (1991) found that the elderly retire in counties where public goods are capitalized more into wages than into land prices. Matthew E. Kahn (2000) found migration into California counties with improving air quality. These tests have identified county differences on a national or regional scale. Yet given constraints on mobility related to family, career, and other networks, residential responses to changes to public goods are most likely to occur within, rather than across, metropolitan areas. Tests at a more local scale have been more modest. Vicki Been (1997) finds little evidence of changes in demographic composition in census tracts following the closing of hazardous waste treatment storage and disposal facilities. In a carefully conducted study of four superfund sites, Trudy Ann Cameron and Ian McConnaha (2006) similarly find little evidence of a pattern in demographic responses in nearby census tracts following the cleanup of a Superfund site. Neither study examines effects on population density. Focusing largely on property value responses, Michael Greenstone and Justin Gallagher (2006) likewise find that, among census tracts hosting Superfund sites, those tracts eligible for federal cleanup experienced no consistent pattern in population density or average income, compared to host tracts ineligible for such cleanup.

Nevertheless, two main concerns raise doubts about the inference that can be drawn from these analyses. First, with the exception of Greenstone and Gallagher (2006), these studies use only a very sparse set of econometric controls. Second, all of these studies use census tracts or counties as the affected “community.” It is, however, not clear that the census tract is the appropriate “community” definition. 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. Finally, as we discuss below, tract-level demographic composition may be endogenous to public goods if the US Census draws tract boundaries to make them more homogenous. Our constructed communities are a response to these concerns.

## I. 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

simple general equilibrium model of location choice. In particular, we use a model of vertically differentiated communities introduced by Dennis Epple, Radu Filimon, and Thomas 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  $y$ . The distribution of income is given by  $f(y)$  with continuous support over the interval  $[y_l, y_h]$ . Households choose to live in one of  $J$  communities, indexed by  $j \in \{1, \dots, J\}$ , and, conditional on location choice, choose their optimal level of housing. Household preferences are represented by the indirect utility function  $V(y, P, G)$ , where  $P$  is the price of a unit of housing and  $G$  represents environmental quality. The specification of  $V(\cdot)$  also implicitly assumes the inclusion of a numeraire whose price is normalized to 1;  $V(\cdot)$  is assumed to be continuous with bounded first derivatives that satisfy  $V_y > 0$ ,  $V_G > 0$ ,  $V_P < 0$ . The associated housing demand function is assumed independent of the level of  $G$  and is given by  $D(P, y)$ , which is also continuous with bounded first derivatives  $D_y > 0$  and  $D_P < 0$ . Demand is assumed to be strictly positive and bounded from above. Each community is characterized by a continuous housing supply function  $S_j(P)$  and an exogenously determined level of environmental quality  $G_j$ . It is assumed that for each community the housing supply function incorporates a bounding price  $P_j^l$  such that  $\forall P \leq P_j^l, S_j(P) = 0$  and  $\forall P > P_j^l, 0 < S_j(P) < \infty$  (i.e., supply is zero at or below the bounding price and finite otherwise).

To facilitate a characterization of the equilibrium sorting of households across communities, we further assume that household preferences satisfy the “single crossing” property. This condition requires that the slope of an indirect indifference curve in the  $(G, P)$  plane be increasing in  $y$ .<sup>1</sup> 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 (uniquely identified by income level) which 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.<sup>2</sup> Unique equilibrium prices  $P_j$  and boundary incomes  $\tilde{Y}_{j,j+1}$  are implicitly defined by the equilibrium conditions of equation (1):

$$(1) \quad V(\tilde{Y}_{j,j+1}, P_j, G_j) = V(\tilde{Y}_{j,j+1}, P_{j+1}, G_{j+1}) \quad \forall j \in \{1, \dots, J-1\},$$

$$M \int_{y \in C_j} D(P_j, y) f(y) dy = S^j(P_j) \quad \forall j \in \{1, \dots, J\},$$

where  $M$  is the total mass of households 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 housing markets clear in each of the  $J$  communities, yielding  $2J-1$  equations to identify the  $2J-1$  endogenous variables.

To develop predictions from the model, we consider the case of two communities. We fix the level of environmental quality in the second community,  $G_2$ , and consider how the equilibrium outcomes (community populations, incomes, and prices) change as the level of environmental quality in Community 1,  $G_1$ , changes. Further, to simplify exposition, we assume an interior

<sup>1</sup> For a discussion of the single-crossing property in this context, see Epple and Sieg (1999).

<sup>2</sup> It is straightforward to relax this assumption by introducing heterogeneity in tastes so that there exists income heterogeneity within each community, but perfect stratification by tastes for each income (see Epple and Sieg 1999; Sieg et al. 2004). Accordingly, this assumption is not critical for the following implications of the model.

equilibrium with the population in both communities strictly positive ( $P_1 > P_1^l, P_2 > P_2^l$ ). We begin by noting that if  $G_1 < G_2$  ( $G_1 > G_2$ ) then the income sorting arising from the single-crossing assumption implies that, in terms of mean income,  $\bar{y}_1 < \bar{y}_2$  ( $\bar{y}_2 < \bar{y}_1$ ). Below, we outline the proof of the following four propositions that characterize equilibrium dynamics. A graphical treatment of the implications of these four propositions is presented in Figure 1.

PROPOSITION 1: *If  $G_1 \neq G_2$ , then*

$$\frac{dPOP_1}{dG_1} > 0, \frac{dPOP_2}{dG_1} < 0, \frac{dP_1}{dG_1} > 0, \frac{dP_2}{dG_1} < 0, \frac{d\bar{Y}_1}{dG_1} > 0, \frac{d\bar{Y}_2}{dG_1} > 0.$$

Proposition 1 follows directly from an evaluation of the comparative statics of the two-community version of equation (1). Using the implicit function theorem, first consider the impact of increasing  $G_1$  on the boundary income,  $d\tilde{Y}_{1,2}/dG_1$ :

$$(2) \quad \frac{d\tilde{Y}_{1,2}}{dG_1} = \frac{-V_{G_1}^1}{(V_y^1 - V_y^2) - f(\tilde{Y}_{1,2}) \left[ \frac{D(P_1, \tilde{Y}_{1,2})V_{P_1}^1}{\int_{y_l}^{\tilde{Y}_{1,2}} D_{P_1}(P_1, y)f(y) dy - S_{P_1}^1(P_1)} + \frac{D(P_2, \tilde{Y}_{1,2})V_{P_2}^2}{\int_{\tilde{Y}_{1,2}}^{y_h} D_{P_2}(P_2, y)f(y) dy - S_{P_2}^2(P_2)} \right]},$$

where  $V^j = V(\tilde{Y}_{1,2}, P_j, G_j)$  and subscripts denote partial derivatives.

The key to signing the derivative in equation (2) is to recognize that the single crossing property implies that  $(V_y^1 - V_y^2) < 0$  and thus  $d\tilde{Y}_{1,2}/dG_1 > 0$ .<sup>3</sup> Because of income stratification, this implies that the population in Community 1 (Community 2) increases (decreases). Also note that because the “new” households entering Community 1 and leaving Community 2 as the boundary shifts out are of higher income level than all the original households,  $\bar{Y}_1$  increases. Conversely, these households have income levels lower than all those remaining in Community 2, so there is also an increase in  $\bar{Y}_2$ . Thus, there is no prediction regarding the *relative* magnitudes of income changes associated with marginal changes in  $G_1$ .

For price impacts,

$$(3) \quad \frac{dP_1}{dG_1} = \frac{-V_{G_1}^1}{\left\{ -V_{P_1}^1 + \frac{(V_y^1 - V_y^2)}{D(\tilde{Y}_{1,2}, P_1)f(\tilde{Y}_{1,2})} \left[ \int_{y_l}^{\tilde{Y}_{1,2}} D_P(P_1, y)f(y) dy - S_{P_1}^1(P_1) \right] \right.} \\ \left. - V_{P_2}^2 \frac{\left[ \int_{y_l}^{\tilde{Y}_{1,2}} D_P(P_1, y)f(y) dy - S_{P_1}^1(P_1) \right] D(\tilde{Y}_{1,2}, P_2)}{\left[ \int_{\tilde{Y}_{1,2}}^{y_h} D_P(P_2, y)f(y) dy - S_{P_2}^2(P_2) \right] D(\tilde{Y}_{1,2}, P_1)} \right\}}$$

<sup>3</sup> By the definition of  $\tilde{Y}_{1,2}$ ,  $V(\tilde{Y}_{1,2}, P_1, G_1) = V(\tilde{Y}_{1,2}, P_2, G_2)$ . Since all those with incomes higher than  $\tilde{Y}_{1,2}$  prefer Community 2,  $V(\tilde{Y}_{1,2} + \varepsilon, P_1, G_1) < V(\tilde{Y}_{1,2} + \varepsilon, P_2, G_2)$ ,  $\forall \varepsilon > 0$ .

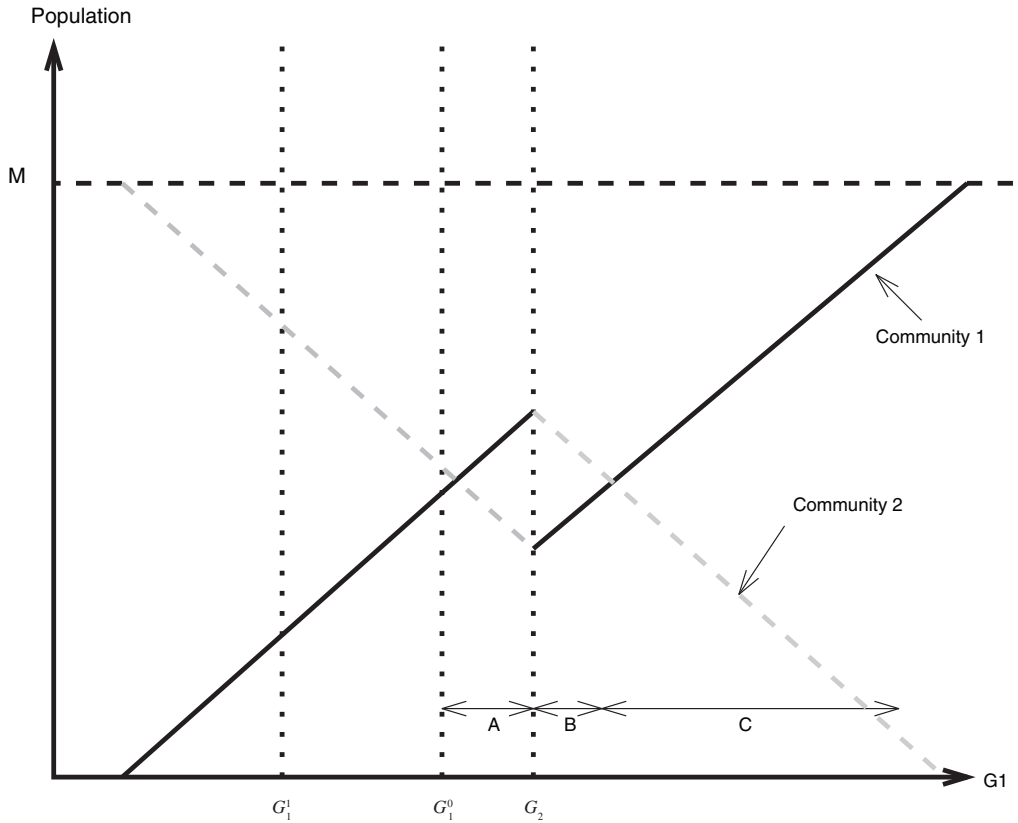


FIGURE 1. IMPACT OF CHANGES IN THE LEVEL OF PUBLIC GOODS ON COMMUNITY POPULATIONS

Notes: The level of the public good in Community 1,  $G_1$ , is increasing along the horizontal axis. The level of the public good in Community 2 is held fixed at  $G_2$ .

Again, it is straightforward to sign equation (3) and demonstrate that  $dP_1/dG_1 > 0$ . Similar analysis demonstrates that  $dP_2/dG_1 < 0$ .

**PROPOSITION 2:** *When  $G_1 = G_2$  there is a unique equilibrium price,  $\bar{P}_{G_1 = G_2}$ , and a continuum of equilibrium household sortings.*

Proposition 2 follows from the fact that when  $G_1 = G_2$  the communities are identical and an interior solution requires that  $P_1 = P_2$ . Further, single crossing no longer implies any particular sorting of individuals. However, whatever the sorting of individuals, the sum of the demand across both communities must equal the sum of supply across both communities. Thus, we have the aggregate market clearing condition of equation (4):

$$(4) \quad M \int_{y \in [y_l, y_n]} D(y, \bar{P}_{G_1 = G_2}, G_j) f(y) dy = S^1(\bar{P}_{G_1 = G_2}) + S^2(\bar{P}_{G_1 = G_2}).$$

Our assumptions on demand, supply, and the income distribution imply that there is a unique price that solves equation (4). However, because in this special case *all* households are indifferent

between communities, in equilibrium we can arbitrarily switch sets of households between the two communities as long as the aggregate housing demand in each set is identical.

**PROPOSITION 3:**  $\lim_{G_1 \rightarrow G_2} P_1 = \lim_{G_1 \rightarrow G_2} P_2 = \bar{P}_{G_1=G_2}$ ,  $\lim_{G_1 \rightarrow G_2^-} POP_1 > \lim_{G_1 \rightarrow G_2^+} POP_1$ ,  $\lim_{G_1 \rightarrow G_2^-} POP_2 < \lim_{G_1 \rightarrow G_2^+} POP_2$ .

In other words, as the quality of  $G_1$  surpasses  $G_2$ , the price in Community 1 increases smoothly, while the population drops discontinuously.

To see this, we begin by considering two related equilibria under the case of  $G_1 = G_2$  and  $P_1 = P_2 = \bar{P}_{G_1=G_2}$ . First, consider the case with complete positive income sorting. In this case, all households with incomes less than some income level, denoted  $\tilde{Y}^-$ , locate in Community 1 and all others locate in Community 2. The second equilibrium is identical, except that the income sorting is reversed and all households with income greater than some level  $\tilde{Y}^+$  locate in Community 1 and all others locate in Community 2. Note that, given the assumption that household housing demands are independent of  $G$ ,  $\tilde{Y}^-$  and  $\tilde{Y}^+$  are uniquely determined by the market clearing conditions and will typically differ. Given our assumptions regarding well-behaved utility and demand functions,  $\tilde{Y}_{1,2} = \tilde{Y}^-$ ,  $P_1 = P_2 = \bar{P}_{G_1=G_2}$  is the limiting equilibrium as  $G_1$  approaches  $G_2$  from below, and  $\tilde{Y}_{1,2} = \tilde{Y}^+$ ,  $P_1 = P_2 = \bar{P}_{G_1=G_2}$  is the limiting equilibrium as  $G_1$  approaches  $G_2$  from above. Thus,  $P_1$  and  $P_2$  are continuous functions of  $G_1$  and  $\lim_{G_1 \rightarrow G_2^-} P_1 = \lim_{G_1 \rightarrow G_2^+} P_2 = \bar{P}_{G_1=G_2}$ .

We now have only to show that the continuity of prices at  $G_1 = G_2$  implies that  $\lim_{G_1 \rightarrow G_2^-} POP_1 > \lim_{G_1 \rightarrow G_2^+} POP_1$ . To see this, note that as we approach  $G_1 = G_2$  from either side, prices are arbitrarily close to  $\bar{P}_{G_1=G_2}$ . At the limit, populations are determined by the market clearing condition evaluated at  $\bar{P}_{G_1=G_2}$ . But, because income sorting is completely reversed as  $G_1$  exceeds  $G_2$ , when the limit is approached from the left, the incomes of households choosing Community 1 are lower than when the limit is approached from the right. Moreover, per capita housing and land demand is higher for richer households. Thus, when evaluated at  $\bar{P}_{G_1=G_2}$ , the market clearing condition implies that Community 1's population is higher when approaching the limit from the left than when approaching from the right. Thus, over some range of  $G_1$ , increases in  $G_1$  that take it above  $G_2$  lead to a decrease in Community 1's population.

A trivial corollary to Proposition 3 is that any change in  $G_1$  that reverses the ordering of the two communities in terms of their public good levels (taking  $G_1$  from a level below  $G_2$  to a level above  $G_2$ ) will lead to an unambiguous increase in the relative mean incomes of the two communities.

**PROPOSITION 4:** *If  $dP_1/dG_1$  and  $dP_2/dG_1$  are bounded away from zero, then even in cases where improvements in  $G_1$  cause it to exceed the level of  $G_2$ , for sufficiently large increases in  $G_1$ , this increase will still result in an increase in  $POP_1$ .*

Proposition 4 follows directly from the fact that increases in  $G_1$  can drive the price in Community 2 arbitrarily close to  $P_2^l$ , thus pushing its population toward zero and the population of Community 1 arbitrarily close to  $M$ .

To clarify the implications of these propositions, Figure 1 graphs the populations of Community 1 and Community 2 as a function of  $G_1$ , holding  $G_2$  constant. As  $G_1$  increases, the population of Community 1 rises everywhere but at  $G_1 = G_2$ , as shown by the solid line in the figure, while the population in Community 2 falls everywhere but at  $G_1 = G_2$ , as shown by the dashed line. Populations always sum to  $M$ . Now, consider the case where  $G_1$  is improved from an initial value of  $G_1^0$ . Small improvements that keep  $G_1$  below  $G_2$  (Region A) result in an increase in the population of Community 1 relative to that of Community 2 and, as discussed above,

indeterminate changes in *relative* incomes. Once  $G_1$  increases above  $G_2$  (Regions B and C), the income sorting is reversed and there is an unambiguous increase in Community 1's income relative to Community 2. However, for small increases above  $G_2$  (Region B), there is actually a drop in Community 1's population relative to Community 2. For larger increases in  $G_1$  (into Region C), there is again an increase in the relative population of Community 1. It should also be noted that this example was specifically constructed to guarantee the existence of Region B. If, instead,  $G_1$  had initially been at level  $G_1^1$ , then *any* increase in  $G_1$  would lead to a relative increase in Community 1's population.

To summarize, the model predicts that marginal increases in  $G_1$  will lead to a relative increase in the population of Community 1 "almost everywhere" (i.e., at all but one point). Large increases in  $G_1$  are likely to have the same effect, with the one exception being the case where  $G_1$  just surpasses  $G_2$ , so that the income effect of housing demand dominates the population effect. The model also predicts that large increases in  $G_1$  will be associated with increases in Community 1's average income relative to that of Community 2.

### A. Empirical Strategy

Of course, whether the changes in the pollution emissions and the presence of polluting facilities studied here constitutes a "large" or a "small" change is an empirical question.<sup>4</sup> We thus take these hypotheses to the data and test for "scale effects" (increases in relative population following improvements in public goods) and income effects.

Our analysis is at a much finer level of aggregation than has previously been considered in the literature, focusing on communities defined by half-mile-diameter circles. To test for scale and income effects, we relate 1990–2000 changes in community populations (in level terms and percentage terms) and changes in income to changes in exposure to air pollution. Environmental quality is measured as the toxicity-weighted exposure to air pollution released from sites listed in the Toxics Release Inventory.

We employ a difference-in-difference design for the effect of 1990–2000 changes in pollution on 1990–2000 changes in these demographic measures. One potential weakness with this analysis is that if firms site their facilities based on local demographics or other confounding factors, the relationship between changes in pollution and changes in demographics would be endogenous. To address this problem, we also identify lagged demographic responses to TRI sites that predated the demographic changes, and can therefore be treated as exogenous. We discuss these and other identification issues in more detail following a description of the data.

## II. 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 Toxics Release Inventory of the US Environmental Protection Agency (EPA).

<sup>4</sup> In previous empirical work with additional co-authors (Sieg et al. 2004), we estimated a structural model, similar in construction to the theoretical model presented here, using cross-sectional data for the Los Angeles Metropolitan Area. That model suggests that 1990–1995 air quality changes would have induced 90 percent of the "communities" (school districts in that model) to change their relative rank, holding air quality in all other communities at baseline levels. This suggests that our model's empirical predictions for income effects from nonmarginal changes are likely to be empirically relevant for the TRI closings studied here.



### A. 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. Been (1997) found that, nationally, one-fifth of tracts had changed boundaries between decennial censuses. Moreover, 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 because it gives more integrity to the concept of community, such gerrymandering may also bias empirical estimates. 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 population, census tracts range greatly in size. For example, in California, tracts range from less than a tenth of a square mile to more than a thousand square miles. This large degree of spatial heterogeneity creates problems when estimating migration models, as any impact of a polluting facility may be diluted when averaging over a large area, but not in the smaller areas to which these larger areas may be compared. Finally, previous research on the correlation between demographics and environmental quality has shown that results can be quite sensitive to community definitions (Douglas L. Anderton et al. 1994; Robert Hersh 1995). Census tracts may be too aggregate a unit and, in any case, preclude sensitivity analysis along these lines.

For these reasons, we take a different approach to neighborhood definitions. We define neighborhoods as a set of half-mile-diameter circles (alternatively one-mile-diameter 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.

The specifics of the data construction are as follows. First, to keep the data task manageable, we restrict our analysis to California. California is attractive because of its racial heterogeneity and because of its relative size. 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.<sup>5</sup>

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 circles a half-mile (one mile) in diameter that are evenly distributed across the state. The set of circles that fall within the Census’s 1990 urban area boundaries is then selected, and all circles that lie across water are dropped. This process yields 6,218 “communities” within one-mile circles and 25,166 “communities” based on half-mile circles. Figure 2 shows the distribution of communities across the study area.

### B. 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, in terms of both

<sup>5</sup> Because of the size of California, the distance between lines of longitude varies by approximately 13 percent as one moves from its southern border to the northern border. To avoid an uneven sampling density between northern and southern portions of the state, it is necessary to account for this variation. In particular, the grids are constructed using the following factors: 1 degree of latitude = 69.172 miles and 1 degree of longitude =  $\cos(\text{latitude}) * 69.172$ .



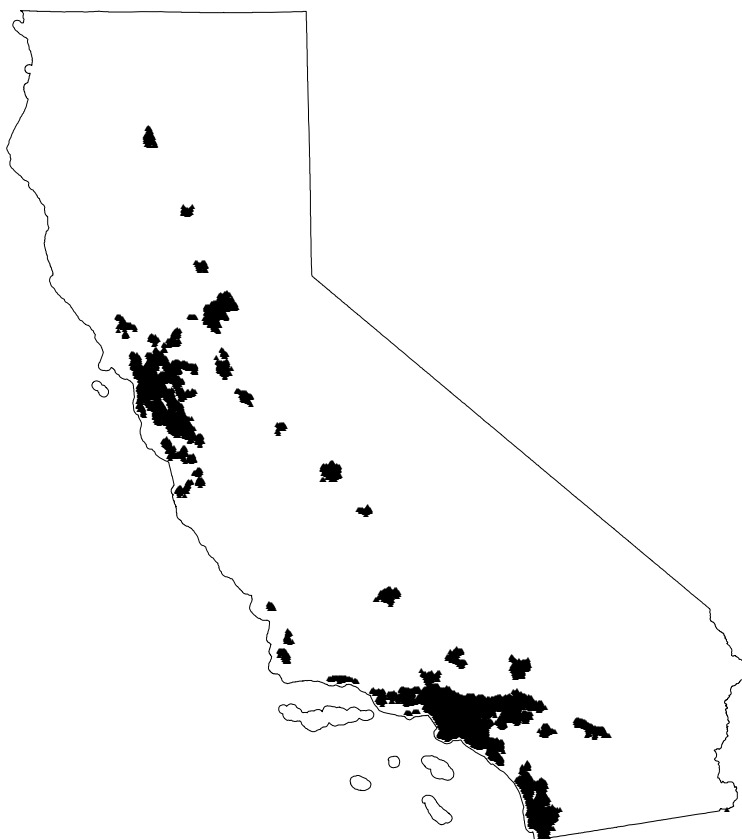


FIGURE 2. LOCATION OF "COMMUNITIES"

individuals and 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 are then 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.<sup>6</sup> Even for our half-mile communities, most blocks are assigned entirely to a single community, and 99 percent are assigned to five or fewer. Table 1 summarizes the opposite mapping, the number of blocks assigned to each community. Because of the splitting of blocks between 1990 and 2000, there are more blocks assigned per circle in 2000 than in 1990. The fiftieth percentile ranges from a low of 10 blocks per circle for half-mile circles (1990) to 38 for one-mile circles (2000).

Table 2 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 (approximately 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 1,162.

<sup>6</sup> 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.

TABLE 1—ASSIGNMENT OF CENSUS BLOCKS AND TRI EMISSIONS TO CIRCLE COMMUNITIES

|                               | Half-mile circles | One-mile circles |
|-------------------------------|-------------------|------------------|
| Count                         | 25,166            | 6,218            |
| Blocks per circle (1990)      |                   |                  |
| 25th percentile               | 4                 | 11               |
| 50th percentile               | 10                | 29               |
| 75th percentile               | 19                | 55               |
| Max                           | 132               | 383              |
| Blocks per circle (2000)      |                   |                  |
| 25th percentile               | 6                 | 17               |
| 50th percentile               | 13                | 38               |
| 75th percentile               | 22                | 64               |
| Max                           | 136               | 408              |
| Circles with TRI exposure     |                   |                  |
| ¼-mile buffer                 | 3,109             | 1,295            |
| ½-mile buffer                 | 5,179             | 1,795            |
| TRI sites for exposed circles |                   |                  |
| ¼-mile buffer                 |                   |                  |
| 25th percentile               | 1                 | 1                |
| 50th percentile               | 2                 | 2                |
| 75th percentile               | 3                 | 4                |
| Max                           | 19                | 25               |
| ½-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               |

### C. TRI Data

As a measure of pollution exposure, we use the EPA's Toxics Release Inventory to find releases of air pollution at facilities throughout California. These sites may be perceived as a disamenity through several channels. First, they may simply be visually unattractive sites. Second, their air pollution may have an unpleasant odor or may cause respiratory irritation. Finally, since 1987, facilities handling more than 10,000 pounds each year of certain hazardous chemicals have been required to report these emissions. These data first became publicly available in 1989 and are routinely publicized by environmental activists. Because these data first became available immediately prior to our 1990–2000 study period, we might also expect to find lagged migratory responses to earlier emissions.

For our measure of exposure, we use a toxicity-weighted index of all emissions in the TRI<sup>7</sup> based on weights developed by the EPA and available in its Risk Screen Environmental Indicators

<sup>7</sup> The list of reporting chemicals 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. This judgment should not make much

TABLE 2—DESCRIPTIVE STATISTICS OF THE DATA FOR HALF-MILE CIRCLE COMMUNITIES

| <i>Baseline demographic data (1990)</i>              | <i>Mean</i> | <i>Standard deviation</i> |
|--|-------------|---------------------------|
| 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                      |
| Percentage households with single-parent families    | 0.08        | 0.07                      |
| Mean rental rate (\$)                                | 689         | 263                       |
| Mean housing value (\$)                              | 229,872     | 138,199                   |
| Share owning their homes                             | 0.66        | 0.27                      |
| Percentage employed                                  | 0.94        | 0.05                      |
| Percentage of employed in manufacturing, if employed | 0.15        | 0.08                      |
| Percentage not graduating from high school           | 0.10        | 0.07                      |
| Percentage with bachelor's degree                    | 0.49        | 0.14                      |
| Average household income (\$)                        | 46,461      | 21,551                    |
| <i>Changes in demographics (1990–2000)</i>           |             |                           |
| Population   | 92          | 256                       |
| Income   | 23,035      | 24,086                    |
| <i>TRI data</i>                                      |             |                           |
| 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                    |
| <i>Locational data</i>                               |             |                           |
| 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                        |

model (RSEI).<sup>8</sup> Because of the TRI reporting threshold and their self-reported nature, these data all involve some measurement error, which is likely to have a conservative effect on our results, biasing them to zero (see Scott de Marchi and James T. Hamilton (2006) for an evaluation of the data). In sensitivity analyses, we also consider nonweighted emissions and the simple presence of a polluting site.

The latitude and longitude for each facility were taken from a recent careful quality control analysis by the EPA.<sup>9</sup> 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 3 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 unshaded 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 3, 3.1 percent of the emissions from TRI site A are assigned to community N1, 17.9 percent to community N2, and 52.4 percent to community N3. In this way,

difference, as the new pollutants are less hazardous and the correlation between 2000 emissions with and without the new chemicals is 0.94.

<sup>8</sup> Information about this model is available at <http://www.epa.gov/opptintr/rsei/>.

<sup>9</sup> This quality control analysis provides a predicted accuracy for each site's location data. Fifteen sites are dropped because of poor-quality geo-coding data.

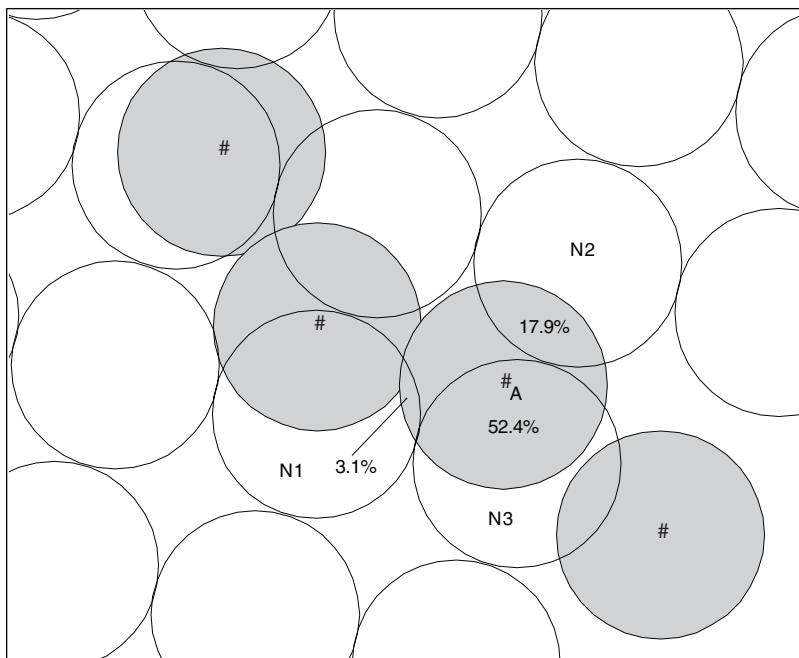


FIGURE 3. HALF-MILE BUFFERS (SHADED) AND CIRCLE COMMUNITIES (UNSHADED) AROUND FOUR TRI SITES

Note: # designates TRI sites.

we can consistently aggregate emissions levels from multiple TRI sites into a total exposure in each community.

Table 1 summarizes the assignment of all TRI sites, active at some time during the 1988–2000 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 fiftieth percentile is two sites, with the maximum exposure going as high as 34 sites in the case of one-mile circles and one-mile TRI buffers. On a community basis, Table 2 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. The table also reports mean toxicity-weighted exposure among all communities and among those exposed.

#### D. Additional Spatial Variables

Several factors, including other spatially distributed amenities, are likely to drive sorting across communities and should be controlled for as well. As controls, we include crime rates and, in some models, coarse locational measures, 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, allowing us to hold district level quality and

changes in quality constant in our comparisons. 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 remaining 3 percent within three or four. Table 1 shows the distribution of the number of communities falling within each school district. Zip codes are even more local measures. Here, we assign each community to one of the 883 zip codes in our study area based on the zip code of the community centroids. Table 1 reports the distribution of communities across zip codes. The median zip code is assigned 21 half-mile circles and 6 one-mile circles.

### III. Estimation and Results

Using these data, we test for differential changes in community population and income that are induced by changes and baseline differences in TRI emissions. Our primary results center on models using half-mile-diameter communities and half-mile-diameter buffers around TRI facilities. We consider respective one-mile diameters in sensitivity analyses.

#### A. Estimation Strategy

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 baseline TRI exposure, changes in TRI exposure, and additional controls.<sup>10</sup> To test for income effects, we use changes in mean income as the dependent variable. As discussed below, we also consider a nonparametric matching estimator as an alternative to linear regression.

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 status from exposed to not exposed (or vice versa), plus continuous measures of the change in emissions 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. These lagged responses may be particularly important when we consider that emissions levels were not reported prior to 1989.

The model for the scale effect regression analysis is presented in equation (5):

$$(5) \quad \Delta POP_i = \delta_0 + \delta_{BL} I_i^{BL} + \delta_{NEW} I_i^{NEW} + \delta_{EXIT} I_i^{EXIT} + \delta_{y_i} y_i^{1990} + \delta_{\Delta y_+} (\Delta y_i | \Delta y_i > 0) \\ + \delta_{\Delta y_-} (\Delta y_i | \Delta y_i < 0) + \delta_D D_i + \delta_L L_i + u_i,$$

where  $\Delta POP$  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;  $\Delta y_i | \Delta y_i > 0$  is the change in toxicity-weighted

<sup>10</sup> 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.

exposure, if positive; and  $\Delta y_i / \Delta y_i < 0$  is the change in toxicity-weighted exposure, if negative.  $D_i$  represents demographic variables and  $L_i$  represents locational variables. For regression models of income effects,  $\Delta POP_i$  is simply replaced by  $\Delta INC_i$ .

To test our model's predictions, we focus on three treatment effects. The "average effect of baseline TRI exposure" estimates the average differential effect on a neighborhood's 1990–2000 population change from being exposed to TRI emissions in 1990. The "average effect of new TRI exposure" estimates the average differential effect on a previously unexposed neighborhood of becoming exposed to TRI emissions. And the "average effect of exiting TRI exposure" estimates the average differential effect on a previously exposed neighborhood of losing all its TRI exposure. These treatment effects are calculated as combinations of the estimated coefficients on both indicator and continuous variables. Specifically,

$$(6) \quad \begin{aligned} \text{Average baseline treatment} &= \hat{\delta}_{BL} + \hat{\delta}_y \left( \frac{1}{N_{BL}} \sum_{i \in BL} y_i^{1990} \right), \\ \text{Average new treatment} &= \hat{\delta}_{NEW} + \hat{\delta}_{\Delta y+} \left( \frac{1}{N_{NEW}} \sum_{i \in NEW} \Delta y_i \right), \\ \text{Average exit treatment} &= \hat{\delta}_{EXIT} + \hat{\delta}_{\Delta y-} \left( \frac{1}{N_{EXIT}} \sum_{i \in EXIT} \Delta y_i \right), \end{aligned}$$

where, for example,  $N_{BL}$  is the number of communities with baseline exposure. Thus, the estimated effect of average baseline TRI exposure, relative to no exposure, is the estimated coefficient on the indicator variable for exposure, plus the estimated coefficient on the continuous measure of exposure times the average baseline exposure among those communities with baseline 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.

We estimate four basic regression models with different levels of controls for confounding factors (the  $D$  and  $L$  variables). Our first model includes no controls. While clearly lacking any pretense to identifying causality, this model does give a signal as to the correlation between migration and changed exposure. Our second model controls for the baseline demographic variables listed in Table 2, plus squares of these terms. As an important spatial amenity, it also includes the FBI crime rate imputed from overlapping political jurisdictions and 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. Finally, all models are estimated with and without baseline population weights for the communities.

Before turning to the results, we consider some of the potential limitations associated with these data and with our experimental design, and we discuss their implication for interpreting our results. In most cases, the potential limitations will tend to bias our results toward zero, so that our tests can be described as conservative. First, note that we are essentially comparing treatment communities within our TRI buffers to control communities outside them, but 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 our control communities, again diluting the identified differential. Moreover, this latter effect is accentuated by our local fixed effects, since the



nearest control communities are the most likely to also be affected by the TRI facility. Thus, we essentially identify only the *differential* effect of the polluting facilities on the nearest (and most affected) communities relative to the next-nearest.

Second, as noted above, TRI emissions are censored based on threshold quantities of individual chemicals handled by the facilities. This suggests that some low-level emissions go undetected. To the extent that this in turn means that some communities diagnosed as controls are in fact exposed, the estimated differential is again diluted. 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.

Finally, consider the potential issues related to endogeneity of the treatments. Our baseline treatment has the strongest claim to exogeneity, since historical exposure cannot be the consequence of future changes in local demographics. However, two potential problems arise with endogeneity for our new treatment and exit treatment. First, 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. Similarly, firms' decisions about opening and closing polluting facilities may be made partly in response to changes in local labor market conditions. Our response to these concerns is the inclusion of the local spatial fixed effects for school districts or zip codes. The relevant spatial scale for considering economic conditions and labor market opportunities in a locational choice is likely much wider than the scale of environmental amenities. If labor and other economic opportunities are roughly equal within a given school district or zip code, but if the environmental disamenities (actual pollution exposure, smell, or sight) of a polluting facility differ by our half-mile communities within those areas, restricting comparisons to being within the school district or zip code through fixed effects will effectively eliminate the endogeneity problem.<sup>11</sup>

A more troublesome problem with endogeneity arises for our new and exit treatments if polluting firms make entry and exit decisions based on the demographics of the community affected by the disamenity per se. Note, however, that the relevant issue is not the baseline demographics of the affected community but, rather, firms' response to *forecasted changes* in demographics from 1990 to 2000.

### B. Scale Effects

The results from our scale-effect models are presented in Tables 3A and 3B. Both the weighted and the unweighted 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 nonlinear adjustments to baseline racial composition.

Table 3A presents the estimated scale effects associated with toxic emissions from TRI sites from the unweighted regressions. The table includes estimates for models with changes in population level and percentage changes in population as the dependent variable, for each of the three

<sup>11</sup> As a plausibility check on this strategy, we compared 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 fails to control for observed confounding factors, one might call into question the assumption that it controls for the effect of *unobserved* spatially varying covariates. We find that the zip code fixed effects control quite effectively for the observable covariates, with point estimates on the treatment effects close and confidence intervals overlapping for the two specifications.

TABLE 3A—ESTIMATED SCALE EFFECTS: UNWEIGHTED<sup>a</sup>

|  | Average effect of<br>baseline TRI exposure | Average effect of<br>new TRI exposure | Average effect of<br>exiting TRI exposure | R <sup>2</sup> |
|--|--|---------------------------------------|---|----------------|
| <i>Effect on population levels</i>               |  |                                       |   |                |
| No controls                                      | -30 (<0.01)                                | -13 (0.37)                            | 43 (<0.01)                                | 0.00           |
| Basic controls                                   | -54 (<0.01)                                | -35 (<0.01)                           | 39 (<0.01)                                | 0.07           |
| School district fixed effects                    | -59 (<0.01)                                | -35 (<0.01)                           | 42 (<0.01)                                | 0.11           |
| Zip code fixed effects                           | -71 (<0.01)                                | -36 (<0.01)                           | 45 (<0.01)                                | 0.26           |
| Matching estimator                               | -32 (<0.01)                                | 27 (0.16)                             | 31 (<0.01)                                | —              |
| <i>Effect on percentage change in population</i> |  |                                       |   |                |
| No controls                                      | -15.6 (<0.01)                              | -5.3 (0.29)                           | 7.1 (0.04)                                | 0.00           |
| Basic controls                                   | -10.7 (<0.01)                              | -7.3 (0.11)                           | 5.0 (0.09)                                | 0.04           |
| School district fixed effects                    | -10.3 (<0.01)                              | -8.3 (0.07)                           | 6.1 (0.04)                                | 0.09           |
| Zip code fixed effects                           | -12.0 (<0.01)                              | -9.3 (0.05)                           | 6.3 (0.04)                                | 0.19           |
| Matching estimator                               | -10.7 (<0.01)                              | -12.11 (0.05)                         | 4.3 (0.04)                                | —              |

Notes: Standard errors for matching models based on Abadie and Imbens (2006). “Basic controls” include latitude, distance to coast, 1990 crime rates, changes in crime rates, 1990 population density, mean rental payments, mean housing values, the percent of households owning their home, the percent employed in manufacturing, the percent without a high school diploma, the percent with a bachelor’s degree, household income, the percent of the population that is black, Hispanic, Asian, or other race, and the percent of households with single-parent families, and squares of these demographic variables. Fixed effect models are the same with the exception that the relevant fixed effects replace latitude and distance to coast.

<sup>a</sup> See equation (6) for definition of the treatment effects. *p*-values for treatment effect in parentheses.

treatment effects. Both the change in level and percentage change models provide statistically significant, policy-relevant evidence of migratory scale effects consistent with the Tiebout hypothesis. Focusing on the percentage change model, on average, baseline exposure to TRI emissions is associated with relative population declines that range from 10 to 16 percent, depending on the model. Likewise, the appearance of new toxic emissions in a previously untreated neighborhood is associated with population declines between 5 and 9 percent. 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 percent.

The unweighted models take communities as their unit of analysis. 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. From a policy perspective, however, we might be equally interested in understanding the average effect of these changes on the population. To better understand how populations are behaving, we rerun these regressions weighting by the baseline population. These weighted regressions are reported in Table 3B. 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, since the weighting scheme underweights 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 theoretical model.

To verify the robustness of these results, we employ a large number of sensitivity analyses. First, as an alternative to changes in population, which could be partly affected by family size and age, we used changes in the number of households as the dependent variable, continuing to find qualitatively similar and statistically significant effects. Second, we replicated the reported models with one-mile-diameter (versus half-mile) communities. For baseline exposure,

TABLE 3B—ESTIMATED SCALE EFFECTS: POPULATION-WEIGHTED<sup>a</sup>

|  | Average effect of<br>baseline TRI exposure |         | Average effect of<br>new TRI exposure |         | Average effect of<br>exiting TRI exposure |         | $R^2$ |
|--|--|---------|---------------------------------------|---------|---|---------|-------|
| <i>Effect on population levels</i>               |  |         |                                       |         |   |         |       |
| No controls                                      | -46  | (0.07)  | -18                                   | (0.43)  | 81  | (<0.01) | 0.00  |
| Basic controls                                   | -81  | (<0.01) | -39                                   | (<0.01) | 71  | (<0.01) | 0.18  |
| School district fixed effects                    | -84  | (<0.01) | -31                                   | (0.13)  | 78  | (<0.01) | 0.25  |
| Zip code fixed effects                           | -108                                       | (<0.01) | -42                                   | (0.07)  | 78  | (<0.01) | 0.58  |
| Matching estimator                               | -43  | (<0.01) | 42                                    | (<0.01) | 46  | (<0.01) | —     |
| <i>Effect on percentage change in population</i> |  |         |                                       |         |   |         |       |
| No controls                                      | -2.6                                       | (0.11)  | 0.8                                   | (0.50)  | 3.0                                       | (0.02)  | 0.00  |
| Basic controls                                   | -3.6                                       | (0.01)  | -0.7                                  | (0.58)  | 2.6                                       | (0.03)  | 0.05  |
| School district fixed effects                    | -4.0                                       | (<0.01) | -1.0                                  | (0.38)  | 3.0                                       | (0.01)  | 0.10  |
| Zip code fixed effects                           | -4.7                                       | (<0.01) | -1.6                                  | (0.19)  | 2.9                                       | (<0.01) | 0.24  |
| Matching estimator                               | -1.1                                       | (0.16)  | -3.9                                  | (0.01)  | 1.7                                       | (<0.01) | —     |

Notes: Standard errors for matching models based on Abadie and Imbens (2006). “Basic controls” include latitude, distance to coast, 1990 crime rates, changes in crime rates, 1990 population density, mean rental payments, mean housing values, the percent of households owning their home, the percent employed in manufacturing, the percent without a high school diploma, the percent with a bachelor’s degree, household income, the percent of the population that is black, Hispanic, Asian, or other race, and the percent of households with single-parent families, and squares of these demographic variables. Fixed effect models are the same with the exception that the relevant fixed effects replace latitude and distance to coast.

<sup>a</sup> See equation (6) for definition of the treatment effects.  $p$ -values for treatment effect in parentheses.

the effects are of greater magnitude (even in percentage terms) 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.

Third, 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 used 1990 and 2000 emissions only (rather than three-year averages), raw emissions 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.

Fourth, we also estimated 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 that do not change status over time, to estimate the effect of baseline exposure. None of these variations changed our results. Fifth and finally, to relax the restriction that treatment effects are the same throughout California, we considered separate analyses of the LA metro area alone and the San Francisco Bay area alone, but again found similar results. Thus, our evidence is highly robust and strongly consistent with the Tiebout hypothesis.

### C. Income Effects

Although our theory model provides strong predictions regarding scale effects, it predicts income effects only for large changes in public goods that affect the relative rankings of the communities. Nonmarginal changes in exposure caused by exiting and entering TRI facilities may well qualify as such changes. In any case, these income effects remain of empirical interest.

To test for income effects, we reestimate the scale effects model using as dependent variables the change in average household income from 1990 to 2000. The fit is reasonably good with  $R^2$ s of 0.31 and 0.42 for models with statistical controls but no fixed effects, and 0.41 to 0.60 for models with fixed effects. Table 4A presents the effects of average TRI exposure using the unweighted model. Baseline TRI exposure causes communities to have a differential growth in average income of about \$2,000 or \$3,000 less than that experienced by the controls. Consistent with the small number of communities gaining exposure, the new exposure treatment generally does not have a statistically significant effect, although in the specifications where it does, the effect is in the expected direction of lower income growth. As expected, losing TRI exposure generally has the opposite effects, and these effects are statistically significant for most models. Our results are not as robust to weighting by population, as reported in Table 4B, with some of the parameter estimates becoming statistically insignificant, although again where they are significant they are consistent with expectations. We also subjected these models to the same sensitivity tests described above for the scale effects, and again found qualitatively similar results. Consistent with our predictions for “large” changes in public goods, these results provide additional evidence in support of Tiebout’s mechanism.

#### D. Nearest-Neighbor Matching

As noted above, we find evidence of nonlinear migratory responses to baseline racial composition, suggesting it may be difficult to control for these effects parametrically. To better account for the uncertainty about functional form, we also nonparametrically match each community receiving a TRI “treatment” to a set of control communities with similar observable characteristics, and then compare their migration patterns. Symmetrically, we match each control community to a set of treatment communities with similar characteristics. 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. In a quasi experiment, treatment and nontreatment 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 linear models presented in Tables 3 and 4.

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 for the treatment and control cases. This is in contrast to the classical assumption that, conditional on the observables, the expected value of the error term is zero. (See James J. Heckman, Hidehiko Ichimura, and Petra Todd 1997, 1998; Rajeev H. Dehejia and Sadek Wahba 2002.)

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 outcomes between treated and untreated observations are calculated. However, the number of cells required to do this can be prohibitively large. Paul R. Rosenbaum and Donald B. Rubin (1983) showed that when a large number of observed variables create too many empty cells, one can, instead, match on a univariate index representing overall distance in the space of observables between a treatment observation and its matched control(s).

We estimate the Euclidian distance between communities in the space of the same set of observables from the regression models. The universe of potential control communities is defined to eliminate confounding treatments. For each community receiving one of these respective TRI

TABLE 4A—ESTIMATED INCOME EFFECTS: UNWEIGHTED<sup>a</sup>

|                               | Average effect of<br>baseline TRI exposure | Average effect of<br>new TRI exposure | Average effect of<br>exiting TRI exposure | R <sup>2</sup> |
|-------------------------------|--|---------------------------------------|---|----------------|
| No controls                   | -7,619 (<0.01)                             | -7,652 (<0.01)                        | 1,899 (<0.01)                             | 0.01           |
| Basic controls                | -2,618 (<0.01)                             | 624 (0.52)                            | 1,344 (0.06)                              | 0.31           |
| School district fixed effects | -2,458 (<0.01)                             | -277 (0.75)                           | 1,693 (0.01)                              | 0.41           |
| Zip code fixed effects        | -2,194 (<0.01)                             | -189 (0.82)                           | 1,416 (0.03)                              | 0.50           |
| Matching estimator            | -3,182 (0.01)                              | -6,115 (<0.01)                        | 530 (0.33)                                | —              |

<sup>a</sup> See footnote for Table 3.

TABLE 4B—ESTIMATED INCOME EFFECTS: POPULATION-WEIGHTED<sup>a</sup>

|                               | Average effect of<br>baseline TRI exposure | Average effect of<br>new TRI exposure | Average effect of<br>exiting TRI exposure | R <sup>2</sup> |
|-------------------------------|--|---------------------------------------|---|----------------|
| No controls                   | -3,951 (<0.01)                             | -5,205 (<0.01)                        | 784 (0.18)                                | 0.01           |
| Basic controls                | -362 (0.33)                                | 216 (0.77)                            | -35 (0.94)                                | 0.42           |
| School district fixed effects | 179 (0.62)                                 | -937 (0.18)                           | -22 (0.96)                                | 0.51           |
| Zip code fixed effects        | 222 (0.55)                                 | -795 (0.23)                           | 126 (0.76)                                | 0.60           |
| Matching estimator            | -2,433 (<0.01)                             | -2,015 (<0.01)                        | 30 (0.93)                                 | —              |

<sup>a</sup> See footnote for Table 3.

treatments, we compare its scale and income effects to the average among four nearest (most “similar”) control communities. The difference, the estimated treatment effect, corresponds to the first terms in equation (6) (without the continuous terms). To account for the potential bias associated with a comparison with control communities having slightly different observables, we adjust for these differences using linear regression.<sup>12</sup>

The last row in each section of Tables 3 and 4 reports the sample average treatment effect from the nearest-neighbors matching estimator. Standard errors are computed using the approach proposed by Alberto Abadie and Guido Imbens (2006). Qualitatively, most outcomes are unchanged from the regression models and are consistent with our derived hypotheses, with one exception. For the “new exposure” treatment, the outcome for the raw scale effect (but not percentage changes) reverses signs.<sup>13</sup> Overall, however, the results provide more evidence of the scale and income effects predicted by the model.<sup>14</sup>

#### IV. Conclusions

Tiebout’s suggestion that people vote with their feet to find the community that provides their optimal tax and public goods pair has played a central role in the theory of local public finance and is central to the study of numerous policy issues including school quality, pollution, housing

<sup>12</sup> In particular, we estimate a linear regression of the respective outcomes on characteristics among control communities. When matching, we then net out the effect predicted to be due to the differences in observable variables. That is, we net out  $\beta(\mathbf{X}_{treat} - \mathbf{X}_{control})$ , where  $\mathbf{X}$  represents the vector of  $y_i$ ,  $\mathbf{D}_i$ , and  $\mathbf{L}_i$ .

<sup>13</sup> This anomaly is eliminated when focusing on the average effect among treated communities only, in contrast to the average effect on the entire sample. It is also eliminated when using any number of matches that meet a strict criterion of “similarity,” rather than always four. Both of these qualifications suggest the sign change is likely the result of the small number of “new” treatments available for matching to the control communities.

<sup>14</sup> In addition to the simple nearest-neighbors estimates reported here, we also estimate the treatment effect by further restricting the match to those control communities located within the same school district as the treated community, an approach analogous to our fixed effects regression models. The results remain unchanged when we restrict the matches to those within the same school district.

policy, and taxation. Despite its importance, this model has been subjected to few direct tests of its basic mechanism: households adjusting locations in responses to changes in these amenities.

Toward this end, we use changes in the emissions of toxic air pollutants across spatially delineated neighborhoods to test for environmentally motivated migration patterns associated with increased demand for land in improving neighborhoods, and which may alter the demographic mix between richer and poorer households. 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 demographics. We find strong evidence of scale effects of a magnitude that is both statistically significant and empirically relevant. We also find evidence of income effects that suggest that pollution in a given location is associated with the emigration of richer households and/or immigration of poorer households.

Our results are consistent with the hypotheses generated from a simple Tiebout model and affirm Tiebout's hypothesis that households do "vote with their feet" in response to local public goods. They also are consistent with recent findings on the potential gentrification effects of exogenous improvements in local amenities (Sieg et al. 2004; Jacob Vigdor 2002).

## REFERENCES

- Abadie, Alberto, and Guido W. Imbens.** 2006. "Large Sample Properties of Matching Estimators for Average Treatment Effects." *Econometrica*, 74(1): 235–67.
- 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.
- Bayer, Patrick, and Christopher Timmins.** 2007. "Estimating Equilibrium Models of Sorting Across Locations." *Economic Journal*, 117(518): 353–74.
- Been, Vicki, and Francis Gupta.** 1997. "Coming to the Nuisance or Going to the Barrios? A Longitudinal Analysis of Environmental Justice Claims." *Ecology Law Quarterly*, 24: 1–56.
- Brueckner, Jan K.** 1982. "A Test for Allocative Efficiency in the Local Public Sector." *Journal of Public Economics*, 19(3): 311–31.
- Cameron, Trudy Ann, and Ian T. McConnaha.** 2006. "Evidence of Environmental Migration." *Land Economics*, 82(2): 273–90.
- De Marchi, Scott, and James T. Hamilton.** 2006. "Assessing the Accuracy of Self-Reported Data: An Evaluation of the Toxics Release Inventory." *Journal of Risk and Uncertainty*, 32(1): 57–76.
- Dehejia, Rajeev H., and Sadek Wahba.** 2002. "Propensity Score-Matching Methods for Nonexperimental Causal Studies." *Review of Economics and Statistics*, 84(1): 151–61.
- Epple, Dennis, Radu Filimon, and Thomas Romer.** 1984. "Equilibrium Among Local Jurisdictions: Toward an Integrated Treatment of Voting and Residential Choice." *Journal of Public Economics*, 24(3): 281–308.
- Epple, Dennis, and Holger Sieg.** 1999. "Estimating Equilibrium Models of Local Jurisdictions." *Journal of Political Economy*, 107(4): 645–81.
- Ferreira, Maria Marta.** 2007. "Estimating the Effects of Private School Vouchers in Multidistrict Economies." *American Economic Review*, 97(3): 789–817.
- 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): 536–60.
- Graves, Philip E., and Donald M. Waldman.** 1991. "Multimarket Amenity Compensation and the Behavior of the Elderly." *American Economic Review*, 81(5): 1374–81.
- Greenstone, Michael, and Justin Gallagher.** 2006. "Does Hazardous Waste Matter? Evidence from the Housing Market and the Superfund Program." MIT Department of Economics Working Paper 05–27.
- Heckman, James J., Hidehiko Ichimura, and Petra E. Todd.** 1997. "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme." *Review of Economic Studies*, 64(4): 605–54.
- Heckman, James J., Hidehiko Ichimura, and Petra E. Todd.** 1998. "Matching as an Econometric Evaluation Estimator." *Review of Economic Studies*, 65(2): 261–94.



- Hersh, Robert.** 1995. "Race and Industrial Hazards: An Historical Geography of the Pittsburgh Region, 1900–1990." Resources for the Future Discussion Paper 95–18.
- Kahn, Matthew E.** 2000. "Smog Reduction's Impact on California County Growth." *Journal of Regional Science*, 40(3): 565–82.
- 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): 957–71.
- Oates, Wallace E.** 2005. "The Many Faces of the Tiebout Model." In *The Tiebout Model at Fifty: Essays in Public Economics in Honor of Wallace Oates*, ed. William A. Fischel, 21–45. Cambridge, MA: Lincoln Institute of Land Policy.
- Rhode, Paul W., and Koleman S. Strumpf.** 2003. "Assessing the Importance of Tiebout Sorting: Local Heterogeneity from 1850 to 1990." *American Economic Review*, 93(5): 1648–77.
- Rosenbaum, Paul R., and Donald B. Rubin.** 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika*, 70(1): 41–55.
- Rubinfeld, Daniel L.** 1987. "The Economics of the Local Public Sector." In *Handbook of Public Economics*. Vol. 2, ed. Alan J. Auerbach and Martin Feldstein, 571–645. Amsterdam: North-Holland.
- 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): 1047–77.
- Tiebout, Charles.** 1956. "A Pure Theory of Local Expenditures." *Journal of Political Economy* 64(5): 416–24.
- Vigdor, Jacob L.** 2002. "Does Gentrification Harm the Poor?" *Brookings-Wharton Papers on Urban Affairs*: 133–73.