NON-MARKET VALUATION ACROSS SPACE AND TIME

A Dissertation
Presented to the Faculty of the Graduate School
of Cornell University
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy

by
Corey Lang
August 2011

NON-MARKET VALUATION ACROSS SPACE AND TIME
Corey Lang, Ph.D.
Cornell University 2011

My dissertation provides new evidence and a new approach to identify the extent and implications of spatial and temporal heterogeneity in the valuation of non-market goods. My research adds dimensions to existing work in economics for pricing goods or features of goods when they cannot be transacted directly in the market. In two particularly policy relevant settings, my dissertation uses unusually rich data and a fine-scale identification strategy to examine the dynamics and spatial variation of non-market values. Chapter 1 extends existing small-area estimation techniques to estimate and map spatial variation in marginal returns to household assets in a developing-country setting; substantial variation exists indicating that poverty reduction efforts of asset-specific transfer schemes would improve with a spatially targeted strategy. The second chapter examines spatial heterogeneity in willingness to pay for air quality and uses those calculations to determine the distributional impacts of the 1990 CAAA. In Chapter 3, I exploit exogenous policy variation from the 1990 Clean Air Act Amendments (CAAA) in order to identify the speed at which air quality improvements are capitalized into housing prices and how preference-based sorting is related to that speed. As a whole, my dissertation contributes to the understanding of non-market values in such a way that can improve empirical policy evaluation and policy design.

BIOGRAPHICAL SKETCH

Corey Lang received a bachelor's degree in Mathematics from Carleton College in Northfield, MN. He worked as an actuary for two years, after which he traveled in Latin America for six months. In 2005, he enrolled at Cornell University to pursue a Doctorate, which he earned in 2011. He is married and has two children.

ACKNOWLEDGEMENTS

I am truly grateful to my dissertation committee. Each member provided unique assistance that enabled me to reach this point.

The research presented in Chapter 1 was financially supported by the International Livestock Research Institute and the National Science Foundation Expeditions grant in computational sustainability (number 0832782).

TABLE OF CONTENTS

CHAPTER 1

Targeting maps: An asset-based approach to geographic targeting

Co-authors: Christopher B. Barrett and Felix Naschold

Introduction

Improved targeting of development interventions has long been recognized as central to increasing the impact from poverty reduction efforts. However, effective targeting requires reasonable estimates not just of who or where the poor are, but also of where the returns to various programs are likely to be highest. Put differently, targeting concerns "what" and "where" questions every bit as much as the more familiar "who" questions. No means currently exist, however, for estimating and comparing expected benefits across space and across alternative interventions, much less of linking such estimates to the spatial distribution of poverty. In this paper, we develop a method that, first, estimates the marginal returns to a range of assets, allowing returns to vary by household and by geography and, second, maps the estimated marginal returns to the various assets, creating a visual tool that can inform the targeting decisions of an asset transfer scheme.

This paper's motivational and methodological starting point is poverty mapping. Elbers et al. (2003) pioneered a technique that combines detailed, nationally representative household survey data with national census data to estimate poverty rates at fine levels of disaggregation for an entire country. Once estimated, the poverty rates for the different regions of a country can be used to create a poverty map, a visual representation of the spatial distribution of poverty.[1] This

[1] The resulting poverty estimates have also been used to investigate the causes of poverty (Kam et al. 2005, Okwi et al. 2007) or its consequences (Demombynes and Ozler 2005).

simple tool is popular and widely used by governments, NGOs and donors in low-income countries to guide poverty reduction efforts.[2]

Although poverty maps can facilitate policy discussions, they offer no explicit recommendation as to the best means of alleviating poverty. If a government is trying to reach a specific welfare target such as the Millennium Development Goals, poverty maps can at best guide the government to regions with high poverty rates. They do not, however, inform the critical subsequent choice of what exactly the government should do in that region.

Targeting maps address this crucial shortcoming of poverty maps by answering two general questions: 1) for a given region, which asset building activity will have the largest marginal gross-benefit? and 2) for a given type of asset building activity, in which regions are the marginal gross-benefits largest? Good answers to either or both of these questions can improve the efficacy of targeted, asset-based development programs. Answers to the first question are paramount for those wishing to cut poverty by the most efficient means possible. The second question appeals to groups interested in investments of a specific type, such as Heifer International in building livestock holdings or The Nature Conservancy in safeguarding natural resources. With scarce resources available to finance transfers, targeting maps can help identify where poverty reduction efforts are likely to generate the most bang-for-the-buck.

This approach takes as given the desirability of geographic targeting. The idea of geographic targeting is to determine a subset of geographic regions most in need and then transfer benefits first (or only) to individuals within the chosen regions. While there are several methods of targeting aid, such as a proxy-means tested targeting, community-based targeting,

[2] Tarozzi and Deaton (2009) and Elbers et al. (2008) evaluate the validity of poverty mapping methods using census data that include income measures. Tarozzi and Deaton argue that useful information is contained in the poverty estimates, but standard errors are too small and assumptions spatial homogeneity are too strong. We partially address the spatial homogeneity concern by including location-specific interaction terms in our model.

categorical or indicator targeting and self-targeting, the empirical evidence suggests that geographic targeting is particularly effective for poverty alleviation (Coady et al. 2004, Baker and Grosh 1994) and is easier and less expensive to monitor and administer than other methods (Bigman and Fofack 2000).

The major disadvantages to geographic targeting are that non-poor individuals living in targeted regions receive benefits (leakage) and poor individuals not living in targeted regions do not receive benefits (undercoverage). One remedy that is routinely applied is to combine geographic targeting with additional targeting tools to limit leakage. Coady et al. (2004) survey 122 targeted transfer programs and find the mean number of targeting tools used is more than two; for example, Mexico's celebrated PROGRESA/Opportunidades program uses four (Coady 2006). A second solution is to target more finely partitioned regions. As regions become increasingly disaggregated, within region heterogeneity decreases and targeting performance increases (Elbers et al. 2007, Baker and Grosh 1994).

In this paper, we build on the proven successes of geographic targeting and propose an enhanced, asset-based approach. We explore the possibility of transfers from an entire range of private and public assets, such as livestock, mobile phones, means of transportation, and access to roads or microfinance institutions. Our focus on assets stems from the importance of a household's asset portfolio in determining the nature, extent and persistence of poverty and vulnerability (Moser 1998, Ellis and Freeman 2004, Adato et al. 2006). Further, if and where poverty traps exist, asset transfers may push households beyond an asset poverty threshold and allow them to engineer their own escape from income poverty (Carter and Barrett 2006).

While in-kind transfers can appear paternalistic, as they constrain household choice in ways that cash transfers do not, there are several reasons why an asset-based approach could perform better than a monetary approach.[3] First, imperfect markets can make it difficult for households to procure desired assets; this is a common rationale for in-kind food or seed aid in many remote or disaster-affected regions. Second, in-kind transfers may stick to the targeted households better than cash because of the well-established endowment effects associated with physical goods but not with cash. For example, the findings of Hoffmann et al. (2009) suggest that in-kind transfers of mosquito nets would result in greater use of the nets than would equivalent cash transfers. Third, some assets – especially public goods such as paved roads – are not readily available for private purchase. Fourth, in-kind transfers often enjoy greater political support than do monetary transfers. Further, monetary transfers, due to their ready divisibility, may also be subject to a high rate of social taxation compared to a lumpy asset, perhaps undoing efforts to control leakage. And in practical terms, governments and charitable organizations routinely make in-kind transfers so improving the efficacy of such interventions is desirable even if one believes cash transfers generally preferable.

The targeting maps tool introduced in this paper improves the information set guiding geographic targeting of in-kind transfers. Given substantial spatial heterogeneity in poverty incidence and its causes (Emwanu et al. 2007, Okwi et al. 2007, Kam et al. 2005), there is little reason to believe that any single transfer form is best suited for all places in a country. Likewise, asset valuation is inevitably spatially heterogeneous, given the place-specificity of many complementary inputs – e.g., agro-ecological conditions that affect livestock value, economic activity that affects the returns to transportation infrastructure. If poverty and the returns to assets both vary markedly across space for a variety of geographic, institutional, policy and technological reasons, then it is desirable to exploit the predictable component of such variation

[3] Currie and Gahvari (2008) review the debate over monetary versus in-kind transfers, though mainly from the perspective of developed countries.

in targeting asset-based development interventions. Previous research has found considerable intra-regional variation in expected returns to different development investments, such as high yielding seed varieties and roads, in Africa and Asia (Fan and Chan-Kang 2004). By customizing asset-based interventions to specific geographic areas, significant gains could be made in cost-effectively addressing poverty. Our approach integrates spatially-explicit estimation of the marginal benefits to multiple assets into a single framework such that inter-asset comparisons of expected marginal benefits can be made for each region and linked to spatially-explicit poverty estimates. The output can then be used as one of several components informing a targeted asset transfer plan.

The method of creating targeting maps, detailed in Section 2, involves several distinct steps similar to those involved in creating a poverty map. Using detailed household survey data and spatially explicit environmental and infrastructure data, we apply multivariate regression and bootstrapping techniques to estimate the returns to various assets and to determine how the estimated returns vary across space. We then project the parameter estimates onto the broader national census data and calculate the marginal returns as a function of projected estimates and current household asset holdings, while simultaneously estimating household-specific poverty status, this latter output very similar to conventional poverty mapping. Finally, we aggregate the estimated marginal returns across households for small geographic areas and, using Geographic Information Systems (GIS), generate maps of both the magnitude and scope of estimated benefits as well as a poverty map.

In Section 2, we also discuss limitations of the methodology, largely centered around issues of endogeneity. Our estimation strategy necessarily ignores bidirectional causality between assets and welfare and unobserved household heterogeneity, both of which could bias estimates. This is a serious concern, but one that is unfortunately unavoidable in any analysis that tries to answer the questions posed above. There is no feasible way to estimate marginal returns to many assets across a large geographical space with ironclad identification. We submit that an explicit, albeit clearly imperfect decision tool is better than none at all and thus that targeting maps deliver useful information that can improve the efficacy of development interventions. While it is impossible to argue a purely causal relationship, understanding how households' asset portfolios and local environment covary with their welfare can nonetheless provide useful insights to inform development interventions. Given the considerable policy and operational importance of the questions targeting maps address, this tradeoff is attractive. Perhaps future research can ameliorate this shortcoming.

We illustrate our approach using Ugandan household survey and census data. The data are discussed in Section 3. The results, discussed in Section 4, are encouraging; estimated and projected marginal benefits to asset transfers seem reasonable and show remarkable variation across space. Our results identify promising areas to target as well as indicate key assets to use in a geographic targeting scheme. Further, our results are consistent with recent Uganda-specific research regarding transportation infrastructure (Lall et al. 2009, Raballand et al. 2009). Our findings reinforce the value of geographic targeting and the importance of spatial analysis.

Method

We estimate average expected marginal household-level benefits to various assets across geographically defined subpopulations. In the context of this paper, assets are taken as anything whose stock can affect a household's welfare.[4] We classify assets along two dimensions: private vs. public and targetable vs. non-targetable. Private and public goods follow standard definitions; public goods are non-rival and non-excludable and private goods constitute the rest. We delineate targetable from non-targetable assets based on whether an asset's quantity, quality or existence can be changed by an intervention. This classification results in four categories: private targetable assets (e.g., livestock holdings, literacy, land holdings), public targetable assets (e.g., source of potable drinking water, access to health clinics, road access), private non-targetable assets (e.g., education level of household head, gender of household head) and public non-targetable assets (e.g., rainfall, temperature). Our method estimates the returns to all types of assets, but ultimately we are only interested in those that are targetable.

The minimum data necessary to create a targeting map are a nationally representative household survey and a census taken at about the same time. In the first step of our analysis, we compare the data available in the household survey and the census to generate a set of variables that are common to both data sets, such as demographic variables, livestock holdings and durable goods. We restrict the data in this way because we must use a regression specification for the survey data that is replicable in the census for all independent variables. Additional environmental or public good variables can and should be added when available to supplement both the survey and census data.

The second step is to use the survey data to estimate the relationship between household welfare and asset holdings, which include the variables selected in the first step as well as relevant environmental and public good variables. We assume that household welfare is a

[4] For now, we remain general about the measurement of welfare, although we use expenditure data to illustrate our method.

function of asset holdings and location-specific asset returns.[5] We remain agnostic about the functional form of the asset returns equations and model the relationship between welfare and asset holdings using a second order flexible functional form. For household *i* in location *c*, we can write the general model as:

$$\begin{aligned}\ln w_{ic} &= A_{ic}'R_A(A_{ic},\overline{A}_c,B_c,Y_{ic},Z_c)+B_c'R_B(A_{ic},B_c,Y_{ic},Z_c)\\ &\quad +Y_{ic}'R_Y(A_{ic},B_c,Y_{ic},\overline{Y}_c,Z_c)+Z_c'R_Z(A_{ic},B_c,Y_{ic},Z_c)+\delta' X_{ic}+\varepsilon_{ic}\end{aligned} \tag{1.1}$$

where w_{ic} = household welfare

A_{ic} = private, targetable assets

$\overline{A}_c$ = location - specific means of A_{ic}

B_c = public, targetable assets

Y_{ic} = private, non - targetable assets

$\overline{Y}_c$ = location - specific means of Y_{ic}

Z_c = public, non - targetable assets

X_{ic} = additional controls[6]

$R_j(\bullet)$ is a vector of returns to asset type $j = A, B, Y, Z$ and is the object of estimation. The functional form of asset returns allows the expected returns to each asset to depend on the stock of every other asset. For example, the returns to a head of cattle may depend on the household head's level of education, the average number of cattle owned in that region, the existence of a nearby livestock market and/or local precipitation levels. Place-specific asset means are only interacted with household levels of the same variable (i.e., average cattle holding is interacted with each household's cattle holdings, but not with each household's pig holdings or mobile phone ownership). Further, we assume the error term is composed of a location component and a household-specific component:

$$\varepsilon_{ic} = \eta_{\mathrm{c}}(\mathrm{M}_{\mathrm{c}})+\mu_{\mathrm{ic}} = \gamma'\mathrm{M}_{\mathrm{c}}+\mu_{\mathrm{ic}} \tag{1.2}$$

[5] This specification can be thought of as permanent or structural income (Carter and May 2001, Adato et al. 2006, Naschold and Barrett forthcoming).

[6] The place specific means, $\overline{A}_c$ and $\overline{Y}_c$, are derived from the census, eliminating sampling error.

where $M_c = [\overline{A}_c, B_c, \overline{Y}_c, Z_c]$.

Our principal goal in this second step in constructing the targeting map is to accurately estimate the coefficients in the welfare-asset relationship. With all interactions included in Equation (1.1), the specification will include more than N(N+3)/2 right-hand-side variables, where N is the combined number of assets in *A*, *B*, *Y*, and *Z*. With so many variables, the likelihood of a spurious relationship is high, which would adversely affect the out-of-sample prediction.

With that in mind, we use stepwise iterative deletion (with a threshold p-value of 0.05) to drop variables from the specification.[7] This in turn can lead to other problems, specifically an important variable for an asset return function or potentially even an entire asset function could be deleted erroneously based on randomness. To alleviate this concern, we bootstrap the whole process 200 times. For each of the 200 iterations, we bootstrap the sample of households from the survey and then estimate Equation (1.1) using stepwise iterative deletion.[8] The regression uses weighted least squares (weighted by population expansion factors) with errors clustered at the enumeration area level.

Having thus estimated the shape of asset returns, in the third step we project the estimated coefficients from the first stage regressions onto the census data. Ultimately, however, we are

---

[7] This practice is common in poverty mapping (Okwi et al. 2006, Emwanu et al. 2007, Demombynes et al. 2007).

[8] Poverty mapping methods often partition the data into the smallest regions for which the survey data are statistically representative and run regressions for each of those regions separately. For example, Okwi et al. (2006) and Emwanu et al. (2007) split Ugandan data into nine strata and Demombynes and Ozler (2005) split South Africa into nine provinces. The idea behind this step is to allow coefficient estimates to vary over space. In contrast, we pool all survey data into a single regression. While in our method coefficient estimates themselves do not vary over space, asset returns can vary via the large number of place-specific interaction terms. Our motivation for this choice is to explicitly take into account the influence of place-specific characteristics on asset returns. If in contrast the geographic scope of regressions was limited, the variation in some variables, especially the place-specific variables such as climate, would necessarily also be very limited leading to biased and inconsistent parameter estimates.

not interested in the coefficient point estimates, but in the expected marginal household-level return for a given targetable asset, k:

$$\frac{\partial E[\ln w_{ic}]}{\partial A_{ick}} = A_{ic}'\frac{\partial \hat{R}_A(\bullet)}{\partial A_{ick}} + B_c'\frac{\partial \hat{R}_B(\bullet)}{\partial A_{ick}} + Y_{ic}'\frac{\partial \hat{R}_Y(\bullet)}{\partial A_{ick}} + Z_c'\frac{\partial \hat{R}_Z(\bullet)}{\partial A_{ick}} \tag{1.3}$$

For each iteration of the bootstrap, we project the coefficient estimates onto the census data and calculate the derivatives for all targetable assets.[9] Combining iterations, we calculate the mean estimated marginal return for each household.

We then aggregate households over geographically defined areas and calculate statistics fundamental to the final product. First, we compute the mean and standard error of the expected marginal benefits for every geographic area and determine which areas have average marginal benefits (AMB) that are statistically significantly greater than zero (at the 5% level).[10] The estimated average marginal returns and their statistical significance inform essential questions about the expected magnitude of average benefits associated with specific asset transfers in particular areas. Second, we calculate the proportion of households with positive expected marginal returns for every geographic area, which reflects the scope of benefits from specific asset transfers in particular areas.

Finally, using GIS, we generate maps that display and enhance the results. Unlike with poverty mapping, no one map can summarize all of the results. This product requires a series of maps. One map can display the most beneficial asset, as judged either by the highest expected

[9] Elbers et al. (2003) use a complex simulation procedure to account for the effect of the distribution of residuals on the poverty estimates. Because our focus is the derivative of Equation (1), our method departs with theirs in this respect. We still model the error, as shown in Equation 2, to accurately control for covariates.

[10] The most common way to estimate the error surrounding the poverty estimates is to use parametric bootstrapping (Elbers et al. 2003, Demombynes et al. 2007). Parametric bootstrapping projects coefficient estimates onto census households by taking random draws from the distribution defined by a single set of regression coefficient estimates and their associated covariance matrix. The poverty statuses of individual households are then averaged by geographic areas. This process is repeated many times to obtain a distribution of each area's poverty. We choose instead to bootstrap the first stage estimation in order to reduce bias in the estimates, since our method puts a greater premium on the regression coefficient estimates themselves.

average marginal returns of any asset or the highest proportion of positive expected marginal returns of any asset, for each geographic area. This map would address question one above: for a given region, which asset building activity will have the largest marginal gross benefit? Then, maps can be made for each asset, showing either the expected average marginal returns or the proportion of households with positive expected marginal returns to that asset for each geographic area. These maps would address question two above: for a given type of asset building activity, in which regions are the marginal gross benefits to such an investment highest? Two estimated objects, two broad targeting questions, and many assets make for a large number of maps, each catering to a different audience or targeting question. Combined with poverty maps that are naturally generated in the third step, one then has a powerful, visual set of tools for informing the geographic targeting of asset-based poverty reduction interventions.[11]

There are unavoidable shortcomings to this approach. First, this is a partial equilibrium analysis that cannot account for general equilibrium effects. Substantial, large-scale asset transfers could affect prices, in which case the estimated marginal benefits would be biased. For example, if too many cattle were transferred into an area, the market price of milk might decline and the benefits of owning a cow become less than estimated. However, this bias could go both ways, as substantial asset transfers could also lead to positive externalities, as would likely be the case with mobile phones or transportation infrastructure or any private asset characterized by (positive) network or technological externalities. We assume that aggregate asset transfers will typically be marginal in magnitude and therefore that partial equilibrium assumptions suffice.

In the introduction, we began to discuss endogeneity concerns. The first such concern is the dual causality between welfare and assets. Does an asset increase a household's welfare or

[11] If poverty estimates generated using a traditional poverty map method are preferred, one could just as easily combine our marginal benefit estimates with those separately estimated poverty rates.

does an increase in welfare cause a household to invest in an asset? Clearly both are plausible, and we cannot separate the two effects.

The second source of endogeneity bias comes from unobserved heterogeneity. Current asset holdings are not randomly distributed; households choose them based in part on information not available to the policy analyst. Households that perceive large returns to an asset due to such unobservables will invest in that asset, while low return households will not. This will likely bias our estimates of marginal returns upwards. The bias is less of an issue when households face constraints on their investment patterns (i.e., credit and savings constraints, missing markets for desired assets), as is often the case in low-income countries.

Additional bias could arise due to using imperfect proxies for welfare as the dependent variable. In our illustrative application, we use expenditure to proxy for welfare. This is the best available choice in the Uganda data, as in many other instances. But it is still incomplete, especially when thinking about asset investments. Some assets are acquired not because they will produce more current expenditure, but because they enhance welfare in some other way or at some future date. For example, some livestock may be held for risk prevention or social status. Further, expenditure can be correlated with asset holdings either positively (one must spend to acquire assets) or negatively (selling assets generates income which increases expenditure).

Collectively, these concerns imply that the cardinality of estimates could be biased, which would affect inter-asset comparisons and cost-benefit analysis. Unfortunately, we are unable to estimate the magnitude of the bias using only cross-sectional, observational data, and thus do not know the extent or magnitude of these effects. But at the very least, our estimates have ordinal significance for comparing the benefits of a single asset across regions. This in itself would have operational value as there exist aid organizations and government ministries that deal in only one asset and have to make intervention siting decisions routinely.

To summarize this sub-section, we deem it important to call attention to the unavoidable shortcomings of the targeting maps method. But we caution against throwing the policy analysis baby out with the statistically imperfect bathwater. We are confident that this method generates meaningful information to help fill an important void that currently plagues development policymaking and programming. Interventions today are typically planned in the absence of any empirical estimates of marginal benefits that permit comparison across space or transfer forms. Despite our method's admitted imperfections, it is a substantial improvement over the status quo.

Data

We apply our method using the 2002 Ugandan National Household Survey, the 2002 Ugandan Population and Housing Census and the 2002 Ugandan Community Survey, all administered by the Ugandan Bureau of Statistics (UBOS). The household survey and census are stratified by four regions (Central, East, North, and West) and an urban-rural split. For the purposes of this paper, we restrict our attention to rural households only (5,648 households in the survey and nearly 4.4 million in the census), due to their greater reliance on natural capital and the greater likelihood of spatially heterogeneous asset returns. The hierarchy for Ugandan administrative units, from largest to smallest, is nation, district, county, sub-county, and parish. Table 1.1 lists how many administrative units of each type exist and the average and median number of households in each unit. A parish contains less than 1,000 households on average and is roughly one-fifth the population of a sub-county. There are one or two enumeration areas

Table 1.1: Hierarchy of Ugandan administrative units and associated number of households

Administrative unit	Total units	Number of households per unit	
		mean	median
District	56	90,797	79,024
County	163	31,194	27,650
Sub-county	958	5,308	4,584
Parish	5,234	971	818

Notes: Data come from the census and include both rural and urban households.

(EA) per parish. The household survey clustered observations at the EA level and randomly sampled (usually) ten households within the EA.

We use per adult male equivalent expenditure as our key measure of welfare. The private asset variables come from the household survey and the census.[12] We use the census, the community survey and several GIS layers to create location specific public asset variables. From the census, we calculate measures of population density and ethnic diversity, as well as average asset holdings at the parish level. The community survey includes information on roads, market access and microfinance access. In addition, we use GIS to derive variables such as average distance to urban areas, average distance to freshwater and average annual rainfall and temperature, among others. Data layers for urban areas and water locations were provided by the International Livestock Research Institute (ILRI). Weather data were downloaded from www.worldclim.org at a resolution of 30 arc-seconds. These geographic variables are aggregated at the sub-county level, due to limitations with the GIS software.[13] Table 1.2 lists

[12] As stated above, we are constrained to only use variables that appear in both the census and the survey. There are several instances where potentially informative variables (e.g., mosquito net coverage of all household members) could not be included due to this limitation. This underscores the importance of planning and coordinating between household surveys and censuses.
[13] Due to the small area of some of the parishes and the relatively larger size of the weather raster data, the zonal statistics could not be calculated for all parishes.

Table 1.2: Summary statistics of all asset variables for rural Uganda

	Survey		Census	
Number of households	5648		4376978	
Monthly household expenditure (Ugandan Shilling)	118147		-	
Private, Targetable assets	Mean	St. dev.	Mean	St. dev.
Cattle (head)	1.74	12.61	1.19	12.49
Goats (head)	0.33	3.76	1.00	7.22
Pigs (head)	0.09	1.15	0.15	1.24
Chicken (head)	1.87	24.75	2.37	16.93
Land ownership (1=yes)	0.29	0.45	0.16	0.36
Motor vehicle ownership (1=yes)	0.04	0.19	0.03	0.17
Bicycle ownership (1=yes)	0.47	0.50	0.35	0.48
Mobile phone ownership (1=yes)	0.03	0.16	0.03	0.16
Proportion of household literate	0.46	0.29	0.45	0.32
Public, Targetable Assets				
Microfinance access (1=yes)	0.79	0.41	0.79	0.41
Road access index	1.09	0.28	1.11	0.30
Private, Non-targetable assets				
Household head education (years)	5.06	3.74	4.55	3.85
Public, Non-targetable Assets				
Population density (per sq. km)	289.2	454.2	396.9	875.5
Ethnic diversity of parish	0.28	0.26	0.29	0.27
Existence of market in parish (1=yes)	0.57	0.50	0.56	0.50
Average distance to an urban area in parish (km)	15.7	10.8	16.0	11.4
Average distance to freshwater in parish (km)	1.98	3.58	1.85	3.16
Average annual temperature (°C)	21.83	2.01	21.86	2.02
Average annual total precipitation (mm)	1227.5	181.6	1224.5	182.9
Average precipitation in driest month (mm)	34.1	15.4	34.3	16.1

Notes: Distance is measured as Euclidean, or straight-line, distance. Motor vehicle ownership equals one if a household owns either a car or motorcycle. Ethnic diversity is calculated (as in Easterly and Levine 1997) as the probability that two people of different ethnicity meet if randomly matched. Microfinance access is derived from the Community Survey and equals one if at least one community within a parish indicated having access to microfinance services. Road access index is derived from the Community Survey, in which community respondents rate their local roads as 0 = "no roads", 1 = "seasonal roads" and 2 = "all weather roads". Responses are averaged from all communities within a parish to form the index.

each asset variable used in the analysis, gives summary statistics for each from the survey sample and census, and defines the variables, if warranted. Cattle and chicken are the most common livestock held. Human capital is low with, on average, five years of education for the household head and less than half of the household literate. Mobile phone ownership stands at just three percent; as a result, estimates of the marginal returns to phones are likely to be inapplicable to current Uganda, given rapid mobile phone uptake in the intervening period. The statistical support is sufficiently similar for the two datasets, supporting our out-of-sample prediction. Table 1.3 lists additional variables used as controls (i.e., the matrix $\boldsymbol{X}$ in Equation (1)).

In addition to the numerical comparability of the data, geographic comparability is important. Table 1.4 reports the percentage of each administrative unit represented in the survey and community census.[14] The survey data appear well dispersed. Thus we have confidence that our estimates effectively represent many different geographies.

Results

As a first step in analyzing the results, we determine the appropriate level of aggregation for the expected marginal returns. In standard poverty mapping exercises, there is a tradeoff between geographic aggregation and precision (Elbers et al. 2003). The goal is to aggregate households into the smallest possible geographic area without sacrificing precision, which enables inter-regional comparison.

We aggregate derivatives and calculate mean marginal benefits and mean standard errors of all targetable assets at three different administrative levels: county, sub-county, and parish.

[14] Due to the incomplete coverage of the Community Survey at the parish level, these variables are aggregated to both the parish and sub-county level, and the sub-county value is joined with the household data when the parish value is unavailable.

Table 1.3: Control Variables

Number of household members in school
Television
Radio
Fixed phone
Postal address
Household's living quarters are in the main part of a house
Household's living quarters are in a room of a house
Housing unit has iron roof
Housing unit has thatch roof
Housing unit has burnt brick walls
Housing unit has unburnt brick and mud walls
Housing unit has mud walls
Housing unit has a cement floor
Housing unit has a earth floor
Household lives in a detached house
Household lives in a tenament
Household cooks with charcoal
Household cooks with firewood
Household lights with electricity
Household lights with lantern
Household lights with tadooba
Household drinks water from a tap
Household drinks water from a borehole
Household drinks water from a well
Household drinks water from an open source
Household uses a private covered pit for a toilet
Household uses a shared covered pit for a toilet
Household uses an uncovered pit for a toilet
Household uses the bush for a toilet
Household bathes inside
Household bathes in outside built structure
Household bathes in outside makeshift structure
Household's kitchen is inside
Household's kitchen is outside built structure
Household's kitchen is outside makeshift structure
Sub-county borders a lake
Sub-county borders Lake Victoria

Table 1.4: Geographic coverage of data

Administrative unit	Survey	Community census	Census
Parish	11%	98%	100%
Sub-county	56%	100%	100%
County	94%	100%	100%
District	98%	100%	100%

Table 1.5: Mean standard errors of estimated average marginal returns at different levels of geographic aggregation

Asset	County	Sub-county	Parish
Cattle	0.004	0.003	0.003
Goats	0.011	0.009	0.008
Pigs	0.072	0.062	0.052
Chicken	0.003	0.002	0.002
Motorized vehicle	0.174	0.153	0.134
Bicycle	0.043	0.029	0.022
mobile phone	0.155	0.126	0.108
Proportion of household literate	0.535	0.531	0.524
Microfinance access	0.070	0.061	0.053
Road access	0.155	0.092	0.076

Table 1.5 gives the estimated mean standard errors for each targetable asset. Clearly, as the area of aggregation grows so does the standard error. This finding contrasts with the standard inverse relationship found in poverty mapping due to the difference in our method, which first estimates household level marginal returns via simulation and then aggregates over geographic areas. Our error estimates are a composite of ordinary imprecision plus inter-household variation. As the geographic scale grows, more inter-household heterogeneity is introduced and the standard errors increase. The empirical findings unequivocally indicate that parish is the appropriate level of aggregation for our estimates.

Table 1.6: Summary statistics for estimated marginal benefits to assets

Asset	Mean AMB conditional on significance (1)	Standard deviation of AMB conditional on significance (2)	Proportion of parishes with significant AMB (3)	Mean PROP (4)	Standard deviation of PROP (5)
Cattle	0.007	0.004	41.4%	81.5%	0.249
Goats	0.021	0.007	17.3%	49.0%	0.380
Pigs	0.086	0.020	0.9%	15.2%	0.228
Chicken	0.003	0.002	30.5%	57.3%	0.419
Motorized vehicle	0.536	0.235	96.2%	99.3%	0.043
Bicycle	0.086	0.051	73.2%	87.1%	0.308
mobile phone	0.422	0.109	91.5%	97.6%	0.126
Proportion of household literate	0.923	0.095	0.3%	62.1%	0.141
Microfinance access	-	-	-	2.3%	0.077
Road access	0.200	0.184	11.2%	33.4%	0.427

Notes: AMB stands for average marginal benefit and PROP stands for proportion of households with expected positive marginal benefits. "-" indicates a missing value.

Relevant summary statistics for estimated marginal benefits are presented in Table 1.6. Columns 1 and 2 give the means and standard deviations of the estimated average marginal benefit (AMB) for parishes with AMB statistically significantly greater than zero. Column 3 gives the proportions of parishes with statistically significant AMB for each asset. Columns 4 and 5 give the means and standard deviations of the proportion of households in each parish with expected marginal benefits greater than zero. The magnitudes of estimated AMB seem reasonable. For example, motorized vehicles are roughly six times more valuable than bicycles, livestock generally offer low returns. However, the AMB estimates of mobile phones and literacy seem inflated, while microfinance access appears low. The microfinance estimates are likely a function of selection bias since microfinance services are commonly targeted to poor areas. Mobile phones, at the time of the survey and census, were a scarcely owned asset and thus likely a luxury good, biasing the estimated benefits of ownership upward. For literacy, so few parishes have significant returns, that the mean estimated AMB is a collection of outliers and likely unrepresentative of the true benefits. Interestingly, the average scope of benefits for improved literacy is large at over 60%. This seemingly inconsistent result suggests that inter-household variation in marginal benefits is large, which pushes most parishes out of the statistically significant category, even though a majority of households would benefit. This duality illustrates the importance of both measures of benefits in informing targeting.

Averages are only a small part of the results. The more interesting results are the spatial distribution, heterogeneity and patterns of the estimated benefits. Figure 1.1 plots the estimated average marginal returns that are statistically significantly greater than zero for cattle, chickens, bicycles, and road access at the parish level. We see pockets of high returns, like those in the northwestern Uganda for cattle, as well as clear spatial patterns, such as the graded decline of marginal benefits to bicycles as one moves further interior from the Northeast border. For each asset, a considerable portion of the country does not exhibit statistically significant estimated returns, reflecting both relatively large standard errors and several negative point estimates. It is reasonable that some returns are actually negative because we estimate marginal returns comprehensively, including areas that are completely unsuitable for certain assets.[15] While only 11% of parishes exhibit statistically significant estimated returns for road access, we see that those significant returns are clustered in the South-central and Southern part of the country, Uganda's most urbanized parts.

[15] Fan and Chan-Kang 2004 and Kam et al. (2005) also find negative estimated returns to assets in some areas.

Figure 1.1: Examples of maps of estimated average marginal returns that are significantly greater than zero for the given asset.

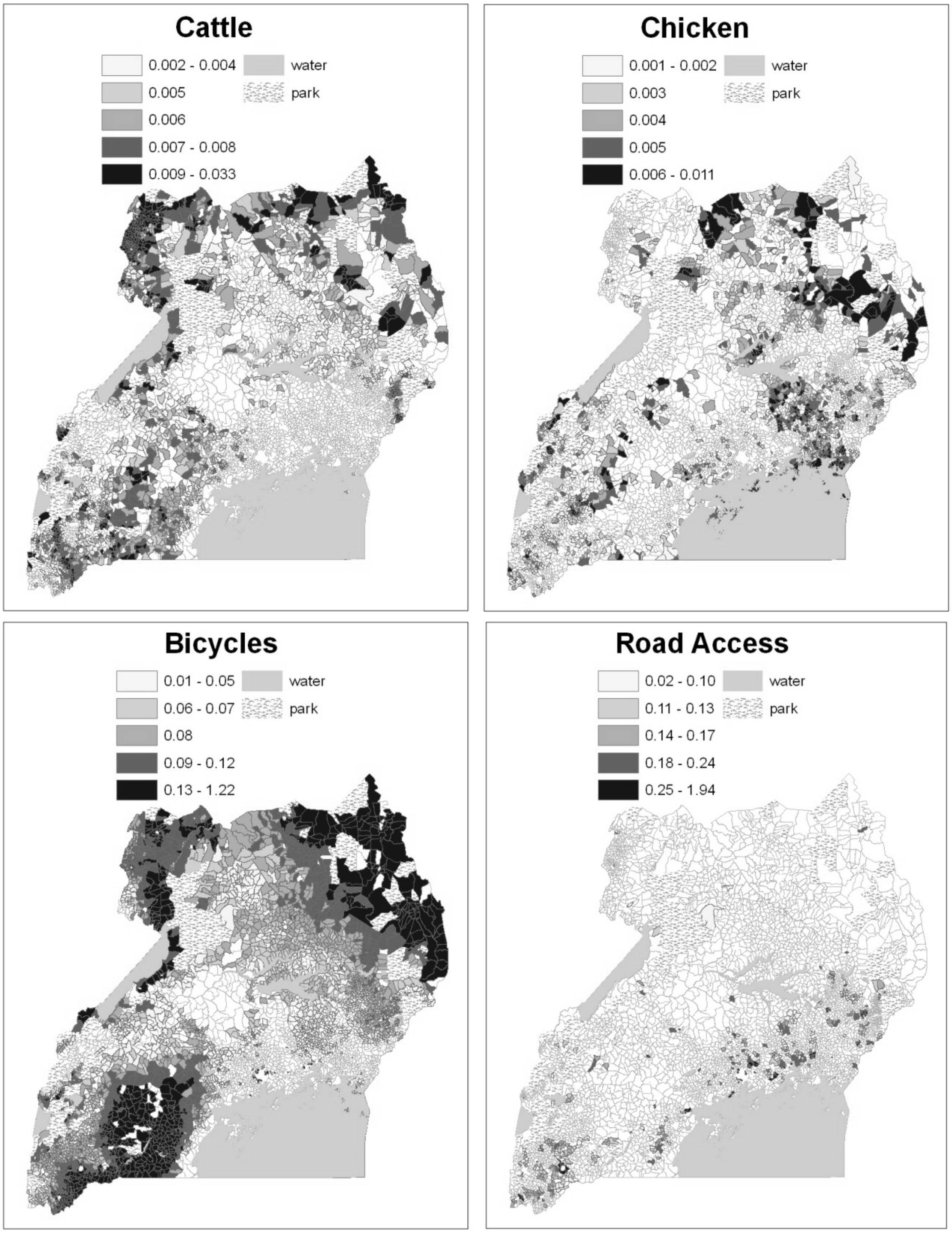


Figure 1.2 plots the proportion of households with estimated marginal returns greater than zero for cattle, chickens, bicycles, and road access. The maps mostly reinforce the information displayed in Figure 2. Areas with high significant returns also have a large proportion of households with positive expected returns (e.g., cattle in the northwest). While the near-monochromatic map for bicycles offers little information in terms of geographic targeting, it suggests that outside of the urban areas in South-central Uganda, positive marginal benefits are near universal.

Next, we identify which asset offers the largest estimated benefits for each parish (Figure 1.3). Motor vehicles and mobile phones dominate these maps, which is not surprising given the large value and expense of motor vehicles and the scarcity and luxury status of mobile phones at the time these data were collected. However, preferred asset maps can be created with any set of assets desired, excluding those that are infeasible (e.g., due to expense) or about which there exists doubt as to the validity of the estimates. To this end, we also generate maps of the assets with maximum returns, limited to livestock assets only, in the bottom panel of Figure 3. Cattle, goats and chickens all have substantial presence on these maps indicating that each species is valuable, but differentially across space, resulting in geographically heterogeneous preferences.

Beyond looking at estimated marginal benefits of an asset, we examine how those benefits relate to existing holdings of that asset and to the poverty headcount rate by parish.[16] A negative correlation between benefits and holdings suggests untapped potential, perhaps indicating the presence of a market failure. If positive, on the other hand, then asset investments

[16] The poverty headcount rate is the percentage of the population that is poor. In Uganda, a household is deemed poor if their estimated monthly expenditure falls below the expenditure thresholds set by Emwanu et al. (2007). As a check on our method, we compare our poverty estimates to those previously estimated for Uganda using the same data from Emwanu et al. (2007), who estimated the poverty headcount rate at the sub-county level. The correlation between the two estimated poverty headcount rates is 0.85; the rank correlation is 0.83.

Figure 1.2: Examples of maps of proportion of households with estimated positive marginal return for the given asset.

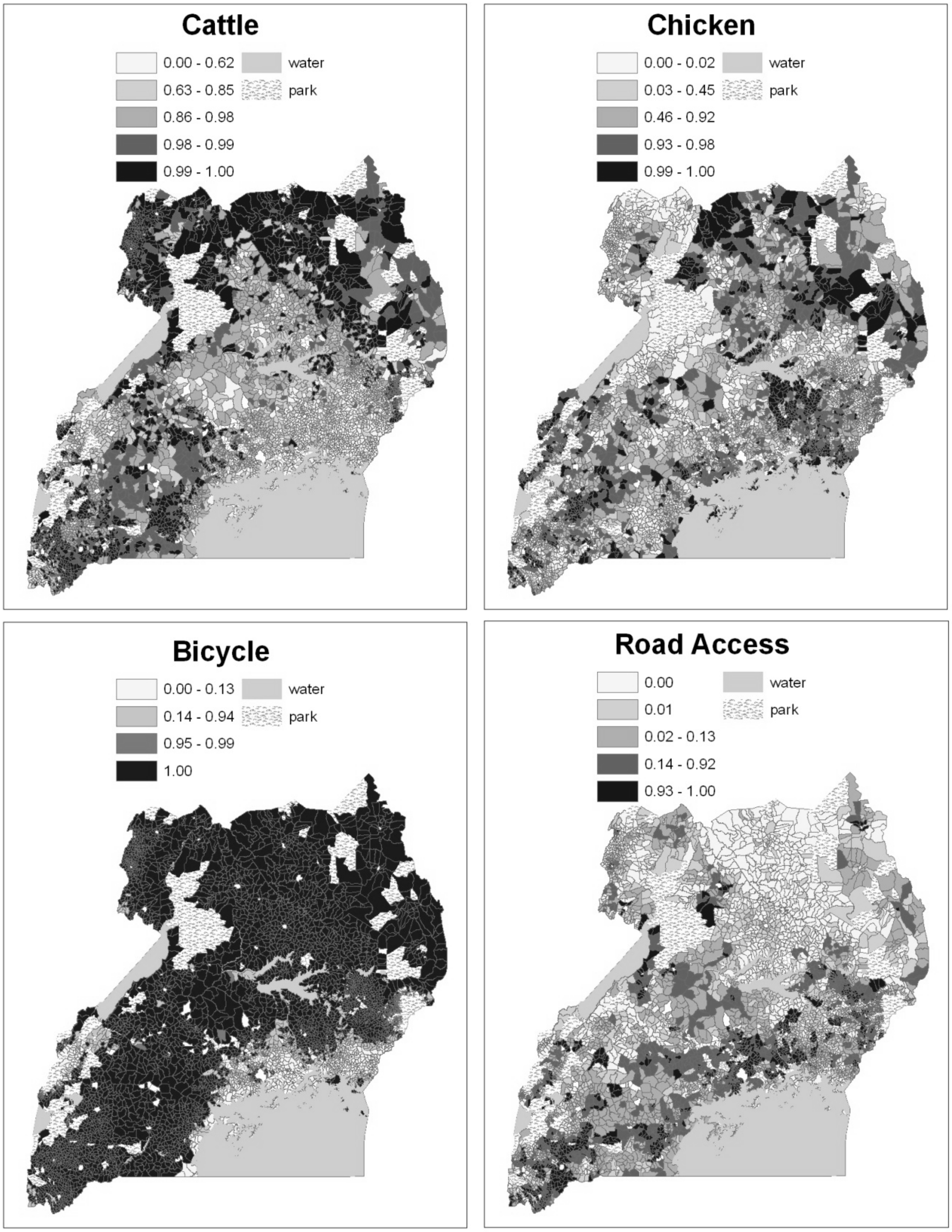
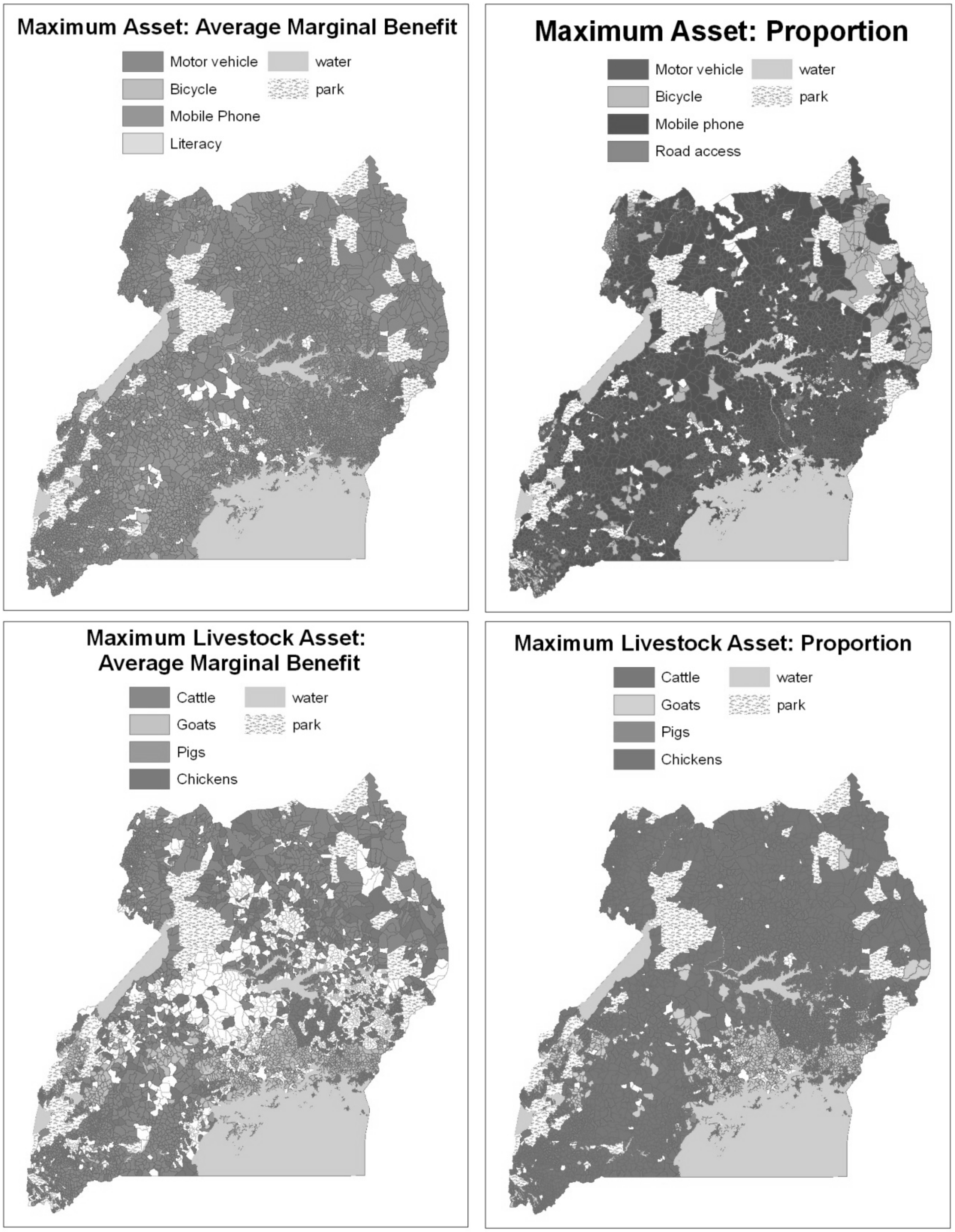


Figure 1.3: Maximum asset returns



are already in line with returns, but further investments could still improve welfare. The correlation between benefits and poverty is particularly relevant for designing the distribution of aid as it reveals prospective tradeoffs and synergies between the objectives of efficiency (i.e., maximizing total expected benefits) and equity (i.e., targeting the poor).

Table 1.7 shows the correlations of statistically significant estimated AMB and the proportion of households with positive marginal returns with mean asset holdings and the estimated poverty rate for all targetable assets. The results suggest that goats and chickens may be difficult for households to procure in areas of high estimated returns. Further, the magnitude and scope of benefits to cattle, bicycle and motor vehicle transfers suggest consistency between poverty reduction and efficiency goals.

Table 1.7: Correlation of estimated average marginal benefits that are significantly greater than zero and proportion of households receiving a positive expected benefit with average asset holdings and the poverty rate

Asset	Asset holdings		Poverty rate	
	Average significant benefit	Proportion positive benefit	Average significant benefit	Proportion positive benefit
Cattle	-0.02*	0.06*	0.05*	0.2*
Goats	-0.12*	-0.2*	-0.04	-0.59*
Pigs	0.6*	-0.16*	-0.37*	-0.37*
Chicken	-0.04*	-0.24*	-0.04*	-0.01*
Motorized vehicle	-0.36*	-0.57*	0.24*	0.11*
Bicycle	-0.21*	0.15*	0.04	0.25*
mobile phone	0.29*	0.06*	-0.1	0.08*
Proportion of household literate	0.01*	0.94*	0.25*	-0.8*
Microfinance access	-	-0.1*	-	0.11*
Road access	0.39*	0.74*	-0.24*	-0.36*

Notes: * indicates significance at 5% level. "-" indicates a missing value.

The analysis thus far has centered on estimated marginal gross returns; information about the costs of supplying different assets has been conspicuously absent. In order to address this deficiency and to enable explicit benefit-cost comparisons (albeit simplistically and incompletely), we compare estimated benefits with actual costs for all livestock assets.[17] Costs are based on the mean price of livestock purchased or sold, as reported in the household survey (costs of other assets are unavailable in the data). The cost data do not include the marginal costs of maintaining stocks; total costs of acquiring and holding an animal would be higher. Because we are unsure over what time horizon the stream of benefits would accrue and what discount rate is appropriate, we report only the expected increase in expenditure for a single month. Table 1.8 presents the findings.

While crude and simplistic, our approach underscores the considerable marginal returns to investment in rural Uganda. Although only pigs pass a cost-benefit comparison outright, the other livestock assets would surely pass if the timeframe was extended in accordance with an animal's expected lifespan. For instance, chicken would pass with a time horizon of three

Table 1.8: Simplified cost-benefit analysis

Asset	Cost	Expected marginal monthly benefit	
		Median	95th percentile
Cattle	214,112	3,074	8,851
Chicken	3,308	1,216	2,864
Goats	16,301	10,200	18,490
Pigs	19,788	43,631	78,401

Notes: All numbers are in Ugandan Shillings. Expected benefits are only calculated for parishes that have significantly greater than zero average marginal benefits.

[17] The expected household marginal benefit was calculated using the approximation for log-linear models that a marginal return of α would increase the household's expenditure by α%.

months, and cattle would pass for time horizons of 6.7 years, just two years in areas with expected returns on the high end of the distribution.[18]

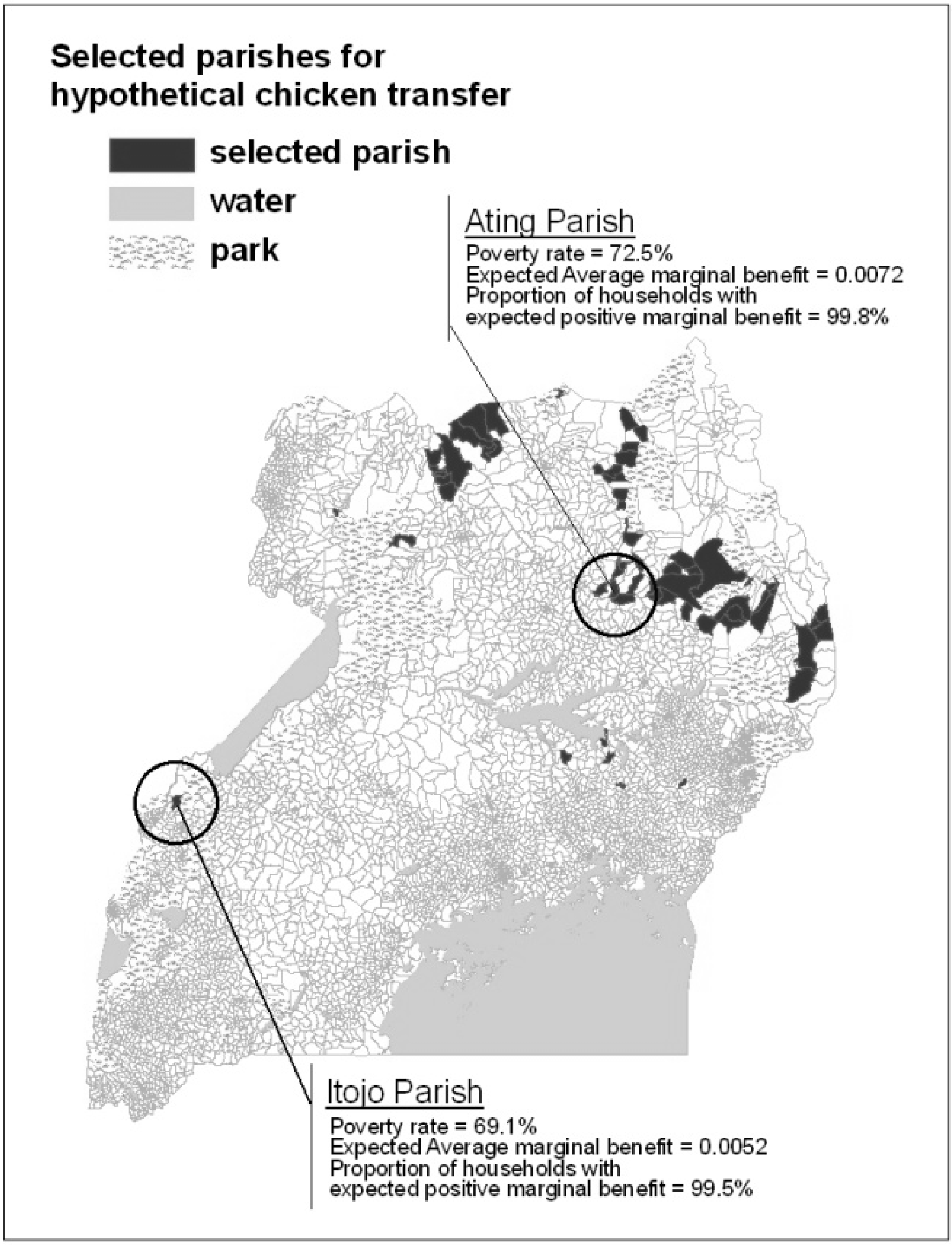
While detailed exploration of the behavioral and institutional reasons for these findings is beyond the scope of this methodological paper, the results clearly underscore apparent underinvestment in productive assets in rural Uganda. Targeting maps of this sort can help development agencies identify best bet forms for asset transfers in a specific area, given such apparent underinvestment. Such targeting maps are perhaps especially useful for geographic targeting of a specific asset transfer program (e.g., livestock or bicycles), since the costs of provision typically vary only modestly across space for a given asset.

As the final step in illustrating the potential utility of targeting maps, we detail a hypothetical chicken transfer program. We choose chickens because they are inexpensive, do relatively well in the cost-benefit comparison and one can easily imagine a development agency implementing such a scheme. We select candidate parishes based on the following three criteria: 1) expected AMB greater than 0.005 and statistically significant, 2) at least 80% of households have positive expected marginal benefit to chickens, and 3) a poverty headcount rate greater than 50%. A total of 58 parishes meet these criteria and are mapped in Figure 1.4. Of those, we highlight two parishes that show particular promise for this sort of development intervention, Itojo parish in the southwest and Ating parish in the northeast, based on high levels of both expected returns and poverty. This sort of simple – and very useful –geographic targeting guidance can be easily repeated for any asset included in the estimation.

Having now presented the results, and given our concerns about endogeneity and other prospective biases in our estimates, we now validate our results using recent empirical research

[18] These calculations assumed a 5% annual discount rate.

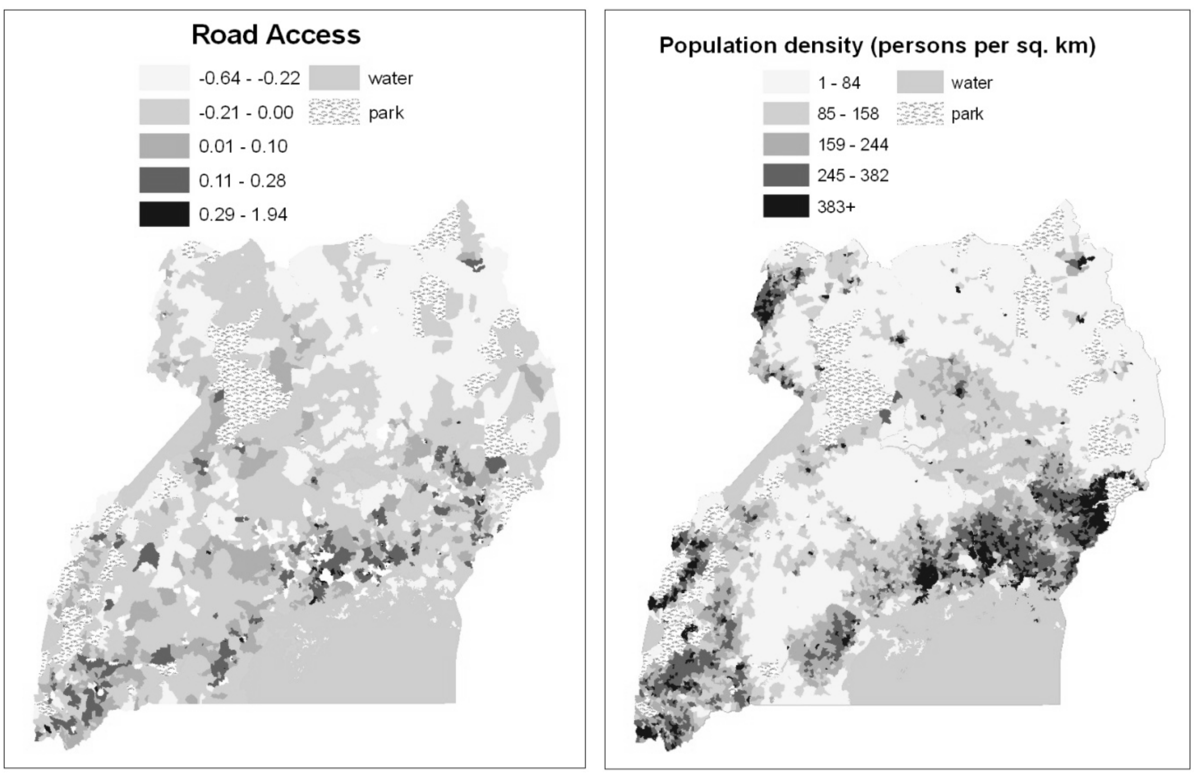
Figure 1.4: Sample targeting exercise



Notes: Parishes are selected by meeting three criteria: 1) estimated poverty rate greater than 50%, 2) expected AMB greater than 0.0005, and 3) Proportion of households with expected positive marginal benefits greater than 80%.

on road access in Uganda. Raballand et al. (2009) examine the roads investment strategy and find that the goal of extending road access to within 2 km of every household is misguided and that larger gains exist in improving and maintaining existing roads. Lall et al. (2009) estimate a locational choice model for industry and find strong agglomeration forces that suggest investment in rural infrastructure is unlikely to benefit the rural poor in terms of job creation. Our results are consistent with these findings. Figures 1 and 2 and the correlations in Table 7 indicate that the benefits to additional road investments are largest primarily surrounding urban areas. Figure 1.5 offers further results. The first panel maps the estimated average marginal benefits to road access (regardless of significance); the second panel is a map of population density. The patterns in the two panels are similar. Given the differences in methods and data

Figure 1.5: Estimated average marginal benefits to road access and population density



between our work and that of Raballand et al. (2009) and Lall et al. (2009), the consistency in results suggests that our method generates sensible results. We found no other comparable empirical studies of spatial distribution of returns in Uganda against which we could compare our results.

Conclusions

This paper presents a novel method that has the potential to enhance the efficacy of geographically targeted asset transfer schemes. We add to the substantial literature on small area estimation, moving beyond estimating poverty so as to begin to identify the best means of alleviating it. Development agencies and government ministries need to know not only where the poor reside, but also what forms of transfers are most likely to help move them out of poverty. Our method first estimates the marginal returns to various assets and then creates a series of maps that can address a variety of questions regarding the magnitude and scope of benefits and the efficient spatial allocation of development programs. The results produced using Ugandan data are promising; estimated and projected asset returns seem reasonable and show substantial variation across space. When combined with a simultaneously generated poverty map, a potentially powerful geographic targeting tool emerges.

Continued work with additional inputs is needed to complement targeting maps. First, even if a policy maker has a targeting map in hand, there are still unanswered questions about the net benefits to and final effects of various asset transfers. We addressed some of these concerns with a limited benefit-cost analysis. A more thorough analysis for all assets with more precise information on procurement and maintenance costs, as well as asset lifespan, is a natural and straightforward exercise for agencies intending to implement a transfer scheme using targeting maps as an input.

Second, targeting maps are not an end in themselves. They estimate marginal returns, which is only an intermediate step to an end goal of poverty reduction. A natural extension of the targeting maps method is to use panel data to determining the expected impacts of an asset transfer program on poverty (or on other outcome variables of interest). Further, optimization algorithms could be constructed to maximize expected poverty reduction given a fixed budget and spatial constraints to transfers (e.g., due to logistical concerns).

The maps and other results produced in this paper serve mainly to demonstrate the potential usefulness of this method. Our hope is that the method can be eventually implemented in development programming, complementing the well-established use of poverty maps in low-income countries. The promise of these methods might also help encourage organizers of household surveys and censuses to better coordinate future questionnaires with poverty maps and targeting maps in mind.

REFERENCES

Adato, M., Carter, M. R. and May, J., 2006. "Exploring poverty traps and social exclusion in South Africa using qualitative and quantitative data" *Journal of Development Studies*, 42(2), 226-247.

Baker, J. L. and Grosh, M. E., 1994. "Poverty Reduction Through Geographic Targeting: how well does it work?" *World Development*, 22(7), 983-995.

Bigman, D. and Fofack, H., 2000. "Geographical Targeting for Poverty Alleviation: An introduction to the special issue" *World Bank Economic Review*, 14(1), 129-145.

Coady D,2006 "The welfare returns to finer targeting: The case of the Progresa program in Mexico" *International Tax and Public Finance*, 13, 217-239.

Coady D, Grosh, M and Hoddinott J., 2004 "Targeting Outcomes Redux" *World Bank Research Observer*, 19(1), 61-85.

Currie, J and Gahvari, F, 2008 "Transfers in cash and in-kind: Theory meets the data" *Journal of Economic Literature*, 46(2), 333-383.

Demombynes, G. and Ozler, B., 2005. "Crime and local inequality in South Africa" *Journal of Development Economics*, 76, 265-292.

Demombynes, G., Elbers, C., Lanjouw, J. O., and Lanjouw, P., 2007. "How good a map? Putting small area estimation to the test" World Bank working paper 4155.

Easterly, W. and Levine, R., 1997. "Africa's growth tragedy: Policies and ethnic divisions" *Quarterly Journal of Economics*, 112(4), 1203-1250.

Elbers, C., Fujii, T., Lanjouw, P., Ozler, B., and Yin, W., 2007. "Poverty alleviation through geographic targeting: how much does disaggregation help?" *Journal of Development Economics*, 88, 198-213.

Elbers, C., Lanjouw, J. O., and Lanjouw, P., 2003. "Micro-level estimation of poverty and inequality" *Econometrica*, 71(1), 355-364.

Elbers, C., Lanjouw, J. O., and Lanjouw, P., 2005. "Imputed welfare estimates in regression analysis" *Journal of Economic Geography*, 5, 101-118.

Elbers, C., Lanjouw, P., and Leite, P.G., 2008. "Brazil within Brazil: Testing the poverty map methodology in Minas Gerais" World Bank Policy Research Working Paper 4513.

Ellis, F and Freeman, A, 2004 "Rural livelihoods and poverty reduction strategies in four African countries" *Journal of Development Studies*, 40(4), 1-30.

Emwanu, T., Okwi, P. O., Hoogeveen, J. G., Kristjanson, P., and Henninger, N., 2007. *Nature, distribution and evolution of poverty and inequality in Uganda 1992-2002.* Uganda Bureau of Statistics and the International Livestock Research Institute.

Fan, S and Chan-Kang, C, 2004. “Returns to investment in less-favored areas in developing countries: a synthesis of evidence and implications for Africa” *Food Policy*, 29, 431-444.

Hoffmann, V, Barrett, C and Just, D, 2009 “Do free goods stick to poor households? Experimental evidence on insecticide treated bednets” *World Development*, 37(3), 607-617.

Kam, S., Hossain, M., Lal Bose, M., and Villano L.S., 2005. “Spatial patterns of rural poverty and their relationship with welfare-influencing factors in Bangladesh” *Food Policy*, 30, 551-567.

Lall, S.V., Schroeder, E. and Schmidt, E., 2009. “Identifying spatial efficiency-equity tradeoffs in territorial development policies: Evidence from Uganda”, World Bank Policy Research Working Paper 4966.

Moser, C, 1998 “The asset vulnerability framework: Reassessing urban poverty reduction strategies” *World Development*, 26(1), 1-19.

Naschold, F. and Barrett, C.B., Forthcoming. “Do short-term observed income changes overstate structural mobility?” *Oxford Bulletin of Economics and Statistics*.

Okwi, P.O., Ndeng’e, G., Kristjanson, P. Arunga, M., Notenbaert, A. Omolo, A., Henninger, N., Bensen, T., Kariuki, P., and Owuor, J., 2007. “Spatial determinants of poverty in rural Kenya”. *Proceedings of the National Academy of Sciences of the United States of America*, 104 (43), 16769-16774.

Raballand, G., Macchi, P., Merotto, D., and Petracco, C., 2009. “Revising the roads investment strategy in rural areas: An application for Uganda”, World Bank Policy Research Working Paper 5036.

Tarozzi, A. and Deaton, A., 2009. “Using Census and Survey data to estimate poverty and inequality for small areas”, *Review of Economics and Statistics*, 91(4), 773-792.

CHAPTER 2

Spatial and Social Disparities in the Benefits from Air Quality Improvements

Co-authors: Antonio Bento and Matthew Freedman

Introduction

Reducing the ambient concentration of particulate matter is a major U.S. public policy issue. Excess exposure to this type of solid and liquid pollutants raises major public health concerns, especially related to the well-being of sensitive populations such as asthmatics, children, and the elderly (Neidell, 2004; Currie and Neidell, 2005; Currie et al., 2009). Elevated concentrations of particulates in the air can also affect public welfare through decreased visibility, damage to animals, crops, vegetation, and buildings.

These various concerns have provided a rationale for public policymakers to pass (and, over the years, amend) the Clean Air Act (CAA). Yet the CAA and its various amendments remain some of the most controversial pieces of federal legislation, with costs well documented but overall benefits and the distribution of benefits poorly understood.

Under the 1990 Clean Air Act Amendments (CAAA), the U.S. Environmental Protection Agency (EPA) refined its particulate policy to regulate particulates less than 10 micrometers in diameter (PM_{10}). This change was implemented due to a growing body of scientific evidence suggesting that the greatest health concerns stemmed from PM_{10}, which can penetrate into sensitive regions of the respiratory tract. The new standard requires the three-year geometric average of PM_{10} concentration for each monitor in a county to be less than 50 $\mu g/m^3$. It further required that the 24-hour average not exceed 150 $\mu g/m^3$. To implement these standards, the U.S. EPA delegates regulatory power to regional air quality management districts.

This paper exploits the structure and implementation of the 1990 CAAA to measure the capitalization of air quality on housing values and to provide evidence of the benefits resulting from the reductions in PM_{10} that occurred between 1990 and 2000. Our empirical strategy relies on a novel instrumental variables approach that explicitly reflects the behavior of the local regulators when implementing and enforcing the CAAA.

For counties with multiple monitors, regulators must keep all monitors in attainment with the federal standards in order for the county to be in attainment. That is, all that is needed for a county to be out of attainment is for a single monitor to have readings above the standard. Therefore, and perhaps not surprisingly, local regulators do not allocate their time and enforcement efforts evenly across monitors. Indeed, as documented in Aufhammer et al. (2009), regulators concentrate their efforts around the dirtier monitors. Aufhammer et al. (2009) provide convincing evidence that the drops in PM_{10} near dirty monitors located in counties out of attainment was -5.4 $\mu g/m^3$ per year. In contrast, the drops among monitors in attainment located in counties out of attainment were essentially the same as the drops in PM_{10} among monitors in attainment located in counties in attainment. Thus, there may be substantial heterogeneity in drops of PM_{10} within counties out of attainment. We investigate the impacts of air quality improvements on housing values taking advantage of monitor-level variation in PM_{10}, evaluating the impacts of air quality improvements on housing values within rings of varying radii around monitors.

A large body of literature, inspired by the seminal work of Rosen (1974), has relied on hedonic price methods to estimate the effects of air quality on housing values. Earlier work typically relied on cross-sectional data that correlate the value of housing transactions for a selected number of metropolitan areas with air quality. Such studies often find no effect of air quality on housing values (sometimes even estimates of the incorrect sign), in part because they suffer from serious identification problems.

More recently, Chay and Greenstone (2005) addressed many of these concerns by exploiting the structure of the 1970 CAAA to create a county level instrument of the attainment status for TSP. The present study differs from earlier work in several ways. First, in contrast to Chay and Greenstone (2005), we explore the regulatory behavior and heterogeneity in changes in PM_{10} to instrument for changes in PM_{10} based on both county and monitor non-attainment status. Monitor non-attainment status effectively captures the behavior of local regulators who concentrate most of their efforts around the dirtiest areas in non-attainment counties. Our more localized estimates using census tracts and rings around each monitor shed light on the heterogeneous impacts of the legislation across space and, correspondingly, socio-economic groups. We build on these results to measure the distribution of benefits resulting from the air quality improvements for households living within 0-1, 1-3, 3-5, 5-10 and 10-20 miles of the monitors. We estimate these based on the county level and monitor level models and discuss the implications of these models for properly measuring the incidence of the policy. Finally, we use these estimates to update the calculations of the overall benefits of the CAAA reported in Chay and Greenstone for the 1990s and discuss the distribution of benefits.

A third set of studies relies on structural methods to examine the distributional impacts of the CAAA. For example, Sieg et al. (2004) and Tra (2010) measured the distribution of benefits to Los Angeles residents that resulted from the improvements in Ozone concentrations. These structural studies calculate the general equilibrium WTP for air quality improvements by decomposing it into two components: the partial equilibrium WTP and the WTP for relocation.

Like Sieg et al. (2004), and Tra (2010), we examine the distribution of the benefits of the 1990 CAAA to different groups of the population. There are two key differences between our estimates and that Sieg et al. (2004) and Tra (2010). First, they report general equilibrium estimates, where the general equilibrium estimate of the WTP for air quality improvements is the sum of the WTP in partial equilibrium and the WTP that results from the relocation of households. Because we do not estimate a structural model, we are unable to calculate the WTP that results from the relocation of households in response to air quality improvements. However, by relying on an instrumental variables approach, we are able to properly identify the partial equilibrium WTP for air quality improvements whereas Sieg et al. and Tra calculate their partial WTP based on cross sectional data and assumptions about determinants of heterogeneity of preferences. To the extent that the bulk of the general equilibrium WTP reported in these studies comes from the partial equilibrium WTP, we believe it is crucial to get these estimates precisely estimated.

The rest of the paper is organized as follows. The next section provides an overview of the CAAAs and describes local regulator behavior in response to the policy. Section 3 describes the data we use in this study and provides some descriptive statistics. We describe our identification strategy in Section 4 before detailing our empirical model in Section 5. After we present our main results in Section 6, we discuss the welfare implications of our findings in Section 7. Section 8 concludes.

The Clean Air Act Amendments and basic aspects of PM_{10} regulation

Particulate Matter (PM) is a term used for a class of solid and liquid air pollutants. Total suspended particulates (TSPs) include particles less than 100 microns in diameter. The 1971 CAA authorized the EPA to enforce a National Ambient Air Quality Standard (NAAQS) for TSPs.

There are two types of standards: primary and secondary standards. Primary standards set limits to protect public health, including the health of "sensitive" populations such as asthmatics, children, and the elderly. Secondary standards set limits to protect public welfare, including protection against decreased visibility, damage to animals, crops, vegetation, and buildings[19]. Each standard is defined in terms of an annual and 24-hour benchmark averages. From April 30th 1971 until July 1st 1987, the primary annual standard for TSPs was 260 $\mu g/m^3$ for the 24-hour average and 75 $\mu g/m^3$ for the annual average. The secondary standard for TSPs was 50 $\mu g/m^3$ for the 24-hour average and 60 $\mu g/m^3$ for the annual average (National Archives and Records Administration, 1987).

If a county exceeded the primary annual standard for one year or the primary 24-hour standard for more than a single day per year it was considered to be in violation of the standard. Under the provisions of the CAA, the EPA can move to designate a county "non-attainment." After a lengthy review process, a non-attainment county was required to submit, in a state implementation plan (SIP), the strategy that it intends to use to become in attainment with the NAAQS. If the deficiency remains uncorrected, or if the EPA "finds that any requirement of an approved plan (or approved part of a plan) is not being implemented," the county is given 18 months to correct the deficiency. If the deficiency continues to be uncorrected, the EPA administrator may impose sanctions on the county in violation, including the withholding of federal highway funds, and the imposition of technological "emission offset requirements" on new or modified sources of emissions within the county (National Archives and Records Administration, 2005). In the first stage of the sanction process, only one of the sanctions is applied at the discretion of the EPA administrator; if the county continues to be in violation 6

[19] See U.S. Environmental Protection Agency (2005) for further discussion.

months after the first sanction, then both are applied. These sanctions are enforced not at the state level, but at the political subdivisions that "are principally responsible for such deficiency" (National Archives and Records Administration, 1987).

In 1987, the EPA refined their particulate policy to regulate particulates less than 10 micrometers in diameter (PM_{10}). The new primary standard required the three-year geometric average of PM_{10} concentration for each monitor in a county to be less than 50 $\mu g/m^3$. It further required via a secondary standard that the 24-hour average concentrations at a monitor do not exceed 150 $\mu g/m^3$. This change was implemented because a growing body of scientific evidence indicated that the greatest health concern from particulate matter stemmed from PM_{10}, which can penetrate into sensitive regions of the respiratory tract.[20]

Data

This section discusses the sources and relevant features of the air quality data, the regulatory data, and the housing and population data in turn. Further, we specify the various levels of geographic aggregation our analysis uses, ranging from individual air quality monitors and census tracts to county averages, and the various advantages and disadvantages of conducting an analysis at each level.

The PM_{10} concentrations data were obtained from the Air Quality Standards (AQS) database, which is maintained by the EPA. For each monitor, the database includes the annual mean concentration, the highest concentration recorded in any 24-hour period, the geospatial coordinates of the monitor, and several reliability measures.

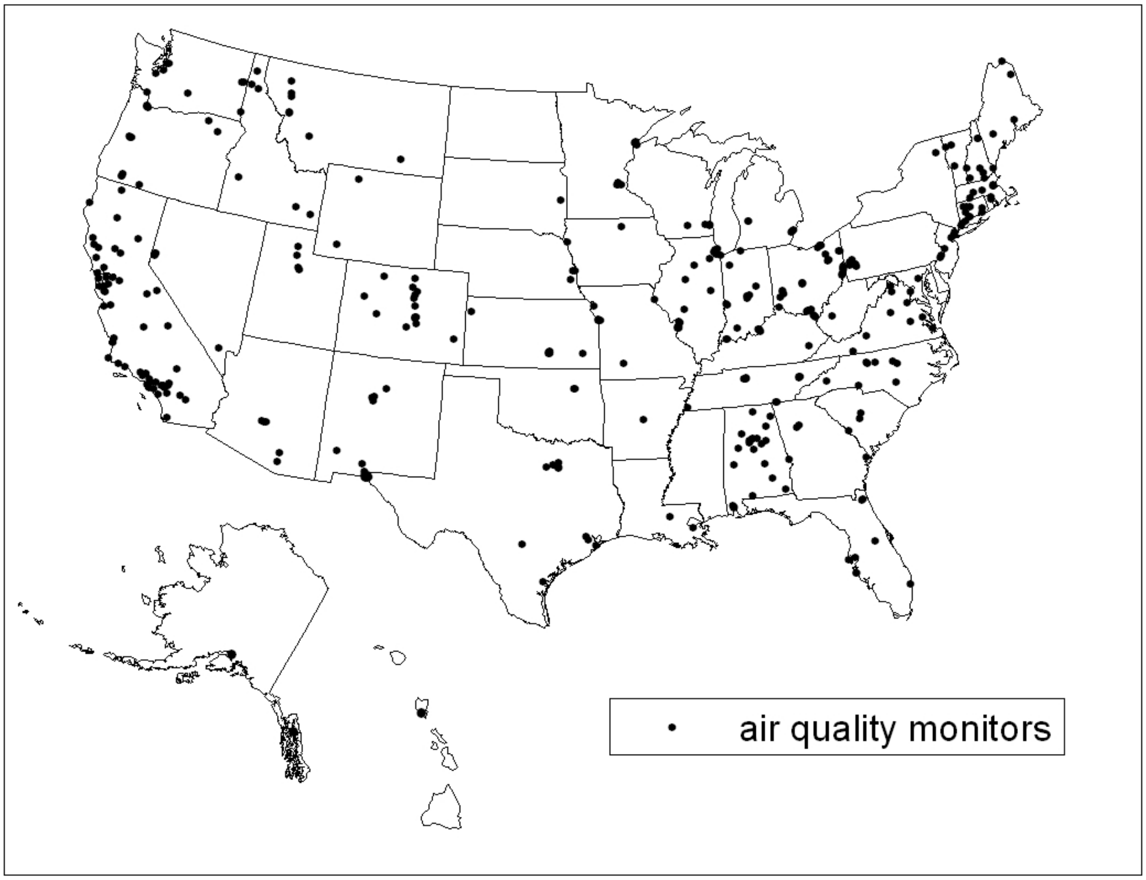
---

[20] For a concise analysis of the health effects from exposure to PM_{10}, see Hall et al. (1992). For an analysis of the impact of air pollution on infant health, see Currie and Neidell (2005) and Chay and Greenstone (2003)

For the purposes of our analysis, we restrict monitors to those that are sufficiently reliable and have readings in necessary years. Title 40 Part 58.12 and Title 40 Part 50 Appendix K of the Code of Federal Regulations (CFR) prescribe the monitoring frequencies for PM_{10} monitors, as well as the criteria for establishing whether a monitor is "representative" and therefore should be included in the analysis. In the AQS data, a criteria flag is set based on data completeness criteria so that if it is set to "Y", then the assumption can be made that the data represent the sampling period of the year. These summary criteria are based on 75% or greater data capture and data reported for all four calendar quarters in each year. Additionally, we exclude monitor-year observations that are affected by "extreme natural events" beyond human influence, occurrences that are noted in the AQS data. Further, we require that monitors have at least one reading in each of the following sets of years: 1989-1990, 1991-1996, and 1999-2000. This enables us to match concentration levels with decadal census data and construct instruments from mid-decade observations.[21] The reliability and timing requirements place significant demands on the set of monitors, and as a result our sample is only a fraction of all monitors that had observations during 1989-2000. However, the PM_{10} trends observed in our sample of monitors are consistent with the full sample of monitors, and thus we feel our sample is representative.[22] Figure 2.1 shows the geographic distribution of the 375 monitors that are included in our sample.[23] The 230 counties containing these monitors encompass approximately one-third of the U.S. population.

[21] If a monitor has a valid observation from 1990 (2000), then that observation is attached to the 1990 (2000) census data. If a monitor does not have a valid observation from 1990 (2000), but does from 1989 (1999), then the 1989 (1999) observation is attached to the 1990 (2000) census data.
[22] The Appendix details the construction of our sample, and we examine the representativeness of our sample. Table A1 shows that 1990 PM_{10} levels are higher on average for included monitors, but the decadal changes are insignificantly different. Further, we perform robustness checks of our main results that relax reliability requirements, and results do not differ substantially.
[23] Three additional monitors satisfied the conditions for inclusion, but were excluded because they were located in the same tract as another monitor and had fewer valid observation days than that other monitor.

Figure 2.1: PM_{10} monitors included in the analysis

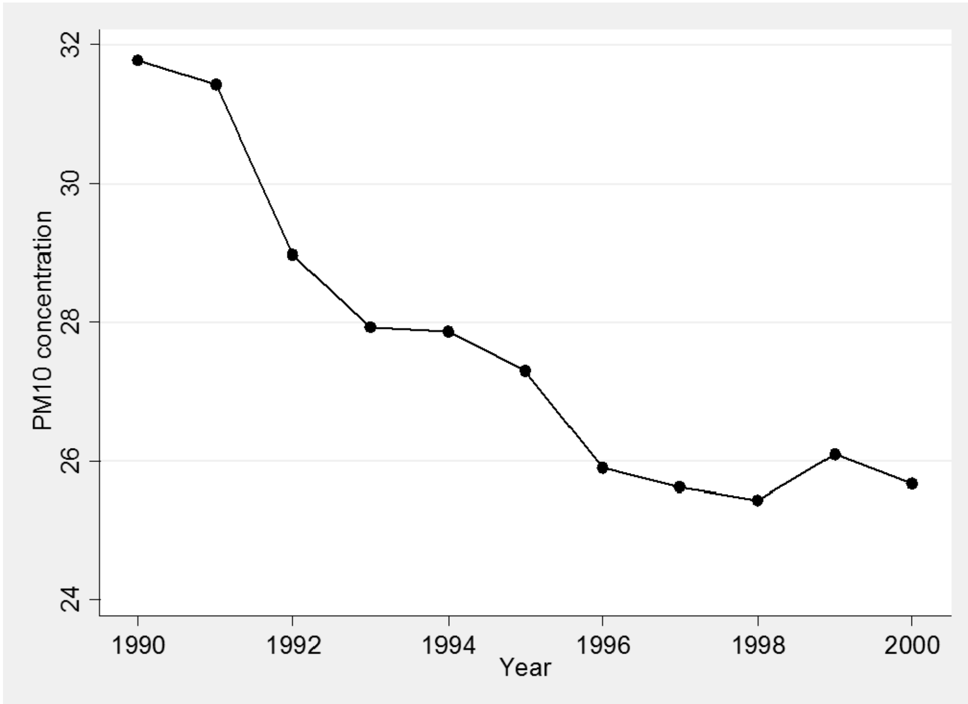


Notes: Monitor records and locations come from the Air Quality Standards database. Shown in the figure are the 375 monitors included in the analysis. See text for details on the inclusion rules.

In the spirit of Chay and Greenstone (2005), we use the weighted annual mean PM_{10} concentration as the key measure of air quality for our estimation. Figure 2.2 shows the trend of PM_{10} over the 1990s; in aggregate, concentrations declined by 19%, which is consistent with the findings of Auffhammer et al. (2009).

There are two primary ways of matching air quality data to human behaviors or outcomes. First, air quality can be averaged over a given area. For instance, Chay and Greenstone average at the county level and Bayer et al. (2010) average at the MSA level.

Figure 2.2: Average PM10 concentrations 1990-2000



Notes: Sample includes all 375 monitors that are included in the analysis.

Second, air quality can be interpolated using a weighted average of proximate monitors. For example, Tra (2010) use distance-weighted average of the nearest three monitors. A variant of interpolation is matching individuals or areas to the nearest monitor (the weighting matrix equaling one for the closest monitor and zero for all other monitors). Sieg et al. (2004) employ this approach.

We use both of these matching strategies in our empirical approach. First, we aggregate air quality to the county level. Consistent with Chay and Greenstone (2005), we construct a county mean concentration as the weighted average of the monitor-specific concentrations, with weights being the number of valid days each monitor was active.

Our second and central analysis focuses on individual monitor readings and matches census tracts to the closest monitor.[24] By not averaging monitor readings, valuable within-county variation in air quality can be used to identify the price-pollution gradient. Using the 88 of our sample counties with at least two monitors (this includes 60% of our sample monitors), we performed a variance decomposition to determine how much of the total variation in PM_{10} concentrations is within versus between counties. The results, shown in Table 2.1, demonstrate that 34.5% of the variance was within-counties for 1990 levels (the remaining 65.5% of variation is attributed to between counties) and that 38.2% of the variance was within-counties for 1990-

Table 2.1: Variance decomposition of PM10 and socioeconomic characteristics within and between counties

	Within	Between
PM_{10}, 1990	34.5	65.5
Median house value, 1990	46.8	53.2
Median rent, 1990	59.5	40.5
Median family income, 1990	81.8	18.2
Share white, 1990	71.0	29.0
Share college grad, 1990	84.5	15.5
Population density, 1990	42.1	57.9
PM_{10}, 2000-1990	38.2	61.8
Median house value, 2000-1990	83.4	16.6
Median rent, 2000-1990	72.7	27.3
Median family income, 2000-1990	88.2	11.8
Share white, 2000-1990	89.9	10.1
Share college grad, 2000-1990	90.4	9.6
Population density, 2000-1990	87.9	12.1

Notes: The PM10 decomposition was done using 233 of the sample monitors that are located in multiple-monitor counties. The housing and population decomposition was done with the 230 counties (and all of the 24,207 tracts contained within them) that are included in our main analysis.

[24] Interpolation is not feasible because our IV strategy depends on specific monitor readings.

2000 changes. This implies significant local heterogeneity exists given that some of these counties are thousands of miles apart. This in turn suggests that measuring valuation of air quality at a fine resolution may yield additional insights because aggregating air quality readings may mask the air quality that individuals truly face. In addition, because many counties are large in area, with some homes 20 or more miles from an air quality monitor, we reduce measurement error by only matching housing units to nearby monitors.

We obtained the county attainment designations from the annual CFR. Since the primary and secondary standards are identical for PM_{10}, we have a single indicator variable for each county and year. While the EPA designates each county in the United States as attainment or non-attainment, not all counties contain air quality monitors that meet our time and reliability requirement, which is necessary to be included in our analysis.

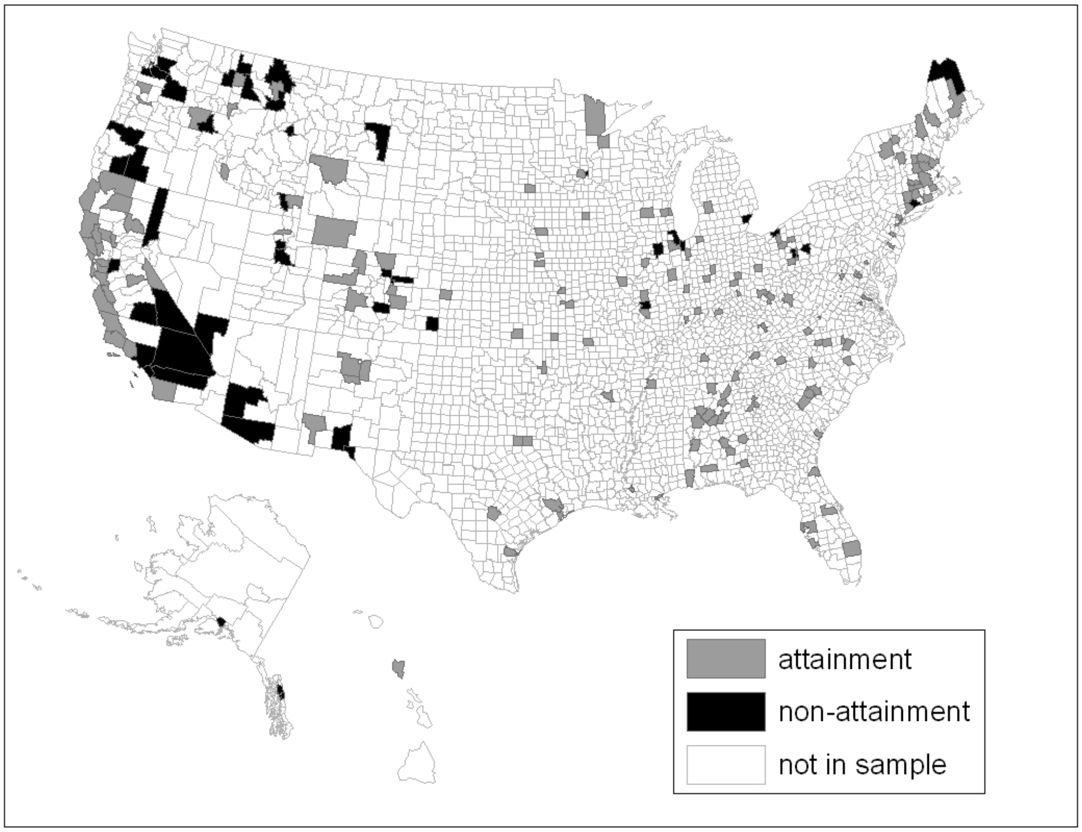
For the purpose of disaggregated, monitor-level analysis, we assign an attainment status for each monitor using the same threshold rules as the EPA's county designation. If in year t a monitor's annual PM_{10} concentration is greater than 50 $\mu g/m^3$ or its 24-hour concentration exceeds 150 $\mu g/m^3$ more than once then that monitor is designated non-attainment in year $t+1$. [25]

Figure 2.3 displays the 1990 attainment status for each county in our sample. The spatial distribution of non-attainment counties confirms widely held beliefs of which areas are most polluted. The Southwestern US, particularly Los Angeles, mountain cities like Denver and Salt Lake, and rust belt cities (Chicago, Detroit, Cleveland, and Pittsburg) are all included in the non-attainment group. While some counties are persistently non-attainment through the 1990s, individual monitors show much more variation.

---

[25] Though each standard was created for different purposes, non-attainment under either standard is expected to result in future pollution reductions as a result of regulatory action. Chay and Greenstone (2005) also use both standards to predict changes in air quality.

Figure 2.3: Geographic distribution of included counties and 1990 attainment status



Notes: 230 counties contain at least one monitor that meets the inclusion requirements, and are thus included in the analysis. Of these, 50 were non-attainment in 1990.

We obtained all of the housing and population data from the GeoLytics Neighborhood Change Database. This dataset aggregates decennial US Census microdata to normalized tract boundaries such that the data is directly comparable across time periods. For the years 1990 and 2000, we obtained tract-level data for the median owner occupied housing value, median rental rate, housing characteristics (e.g., share of owner occupied housing units with 2 bedrooms, share of owner occupied housing units built more than 50 years ago) and socioeconomic composition (e.g., share Latino, median family income). Table 2.2 gives a complete list of variables used in the analysis and their means at the tract level. Using Geographic Information Systems (GIS), we

Table 2.2: Summary statistics

	County		Tract	
	1990	2000	1990	2000
Median housing value (2000$)	133,075	131,697	110,030	111,455
Median rent (2000$)	522	492	771	536
Median family income (2000$)	45,669	49,452	37,586	40,819
PM10 concentration	31	24	32	25
Total housing units	167,801	185,561	1,635	1,736
% of housing units occupied	89.7	90.8	89.8	90.1
% of housing units owner occupied	64.5	66.2	52.2	52.6
% of houses heated by coal	0.4	0.1	0.3	0.1
% of houses heated by wood	6.8	3.1	3.7	1.5
% of houses without a kitchen	1.2	1.6	1.4	2.1
% of houses with full plumbing	99.3	99.4	99.2	99.2
% of houses with 2 bedrooms*	24.7	22.5	31.8	30.0
% of houses with 3 bedrooms*	49.7	49.5	44.4	43.2
% of houses with 4 bedrooms*	17.5	19.0	13.7	14.7
% of houses with 5+ bedrooms*	4.1	4.5	3.7	4.1
% of houses that are single, detached units*	81.5	81.9	76.9	77.3
% of houses that are single, attached units*	3.9	4.3	5.3	5.4
% of houses that are mobile homes*	9.4	8.8	7.1	6.8
% of houses built 5-10 years ago*	9.1	7.8	5.9	3.9
% of houses built 10-20 years ago*	22.1	14.8	15.4	9.6
% of houses built 20-30 years ago*	16.0	18.0	13.6	13.4
% of houses built 30-40 years ago*	16.0	13.0	16.1	11.5
% of houses built 40-50 years ago*	8.7	13.3	11.7	15.6
% of houses built 50+ years ago*	18.2	22.9	31.5	40.3
% less than high school education	23.2	18.1	31.3	26.3
% with college degree	20.4	24.3	15.3	18.5
% Black	10.2	11.4	14.1	15.8
% Latino	7.2	10.2	11.7	16.3
% under age 5	8.8	6.6	9.0	6.8
% over age 65	12.7	12.7	14.2	12.9
% foreign born	5.5	8.0	7.0	10.4
% of households headed by a female	22.1	24.0	31.1	32.6
% living in same house as 5 years ago	52.0	53.5	49.3	49.6
% unemployed	6.8	6.2	9.4	9.0
% employed in manufacturing	19.2	16.6	18.1	15.5
% poor	13.3	12.7	20.7	20.1
% receiving public assistance	7.6	8.1	12.0	12.4
Population density (per sq. mile)	1,324	1,392	3,120	3,175
Total population in sample	94,784,035	106,679,918	1,364,128	1,443,936

Notes: Sample includes 230 counties and 352 tracts that contain air quality monitors included in the analysis. 1990 housing prices are adjusted using the CMHPI. Variables marked with a * are specific to owner-occupied housing units and are included only in homeowner-specific analyses. Analogous variables for rented housing units are available and are included in the renter specifications.

matched each monitor to a single census tract, and thus create the sample of 352 tracts (23 monitors match to tracts with missing data).

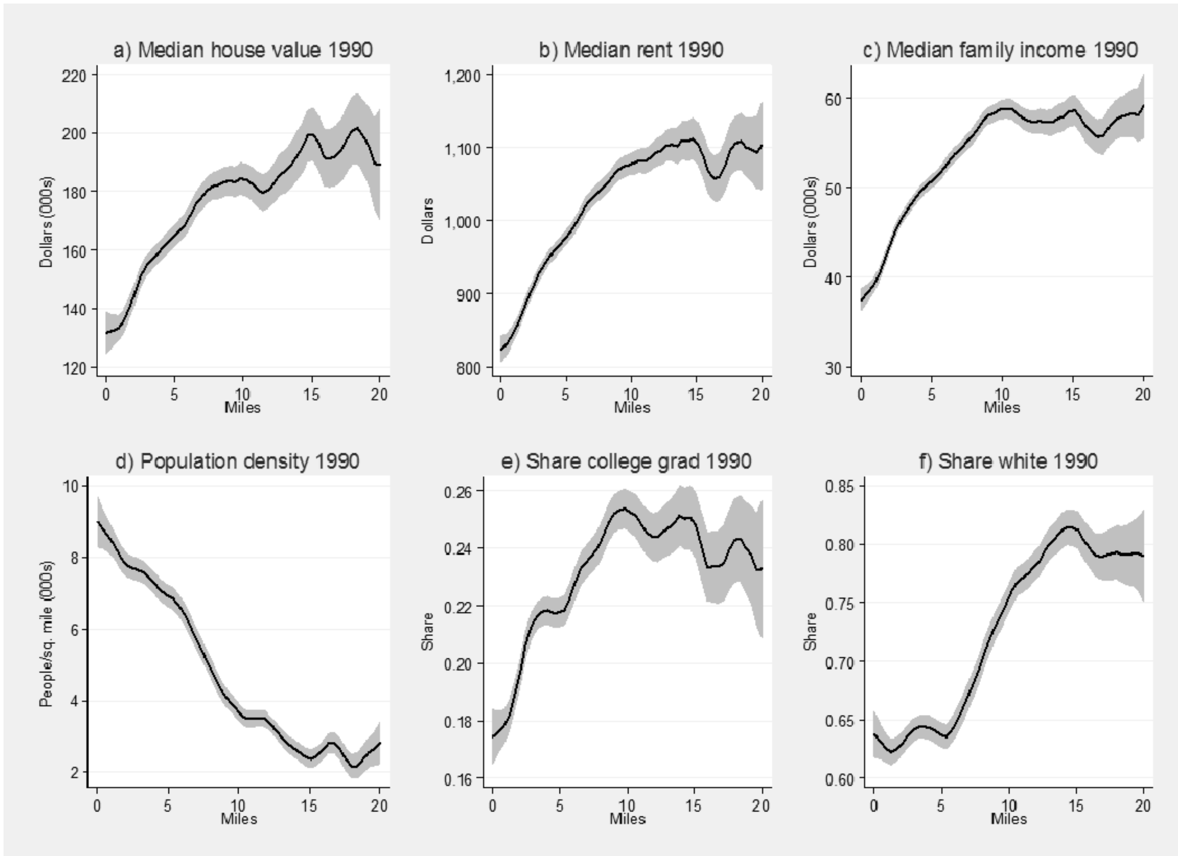
To match with our county average PM_{10} measure, we additionally aggregate the housing and population data to the county level, using population weights. Since a weighted mean of median house value does not equal a county mean, we utilize additional census data reported at the county level to provide the most accurate measure for median house value, as well as median income and median rent.[26] This data is then matched to the county level air quality data to form the sample of 230 counties. Table 2.2 additionally reports means at the county level.

Just as we examined within-county variance of air quality, the GeoLytics data allow us to explore how much housing and socioeconomic characteristics vary within counties, as compared to between counties. Table 2.1 reports the results of a variance decomposition performed on the sample counties for a subset of the variables used on our analysis for 1990 levels and decadal changes. For every variable, within-county variance is substantial, nearly always greater than 50%, and surpasses 80% for many decadal changes.[27]

In Figure 2.4, we further explore within-county heterogeneity by plotting non-parametric relationships between a tract median housing prices, median income, population density, share college educated, and share white and that tract's distance to the closest air quality monitor in our sample. Figure 2.4 demonstrates systematic variation; as the distance from a monitor increases, median housing and rental values, median incomes, share college educated and share white all increase, whereas population density decreases. Figure 2.4 underscores the fact that the monitors are placed in tracts that are systematically different than other tracts and the county as a

[26] We downloaded data on median housing value, median rent and median income for all counties for 1990 and 2000 from USA counties, an aggregation of decennial censuses as well as other data (available at http://censtats.census.gov/usa/usa.shtml). As a robustness check, we also calculated the 'median' variables in the same manner as the buffers (i.e. the population weighted average of the median) and the results were very similar.
[27] A variance decomposition done with all US counties gives similar numbers.

Figure 2.4: The relationship between select socioeconomic characteristics and distance from monitors



Notes: The sample of 23,008 census tracts consists of tracts within included counties whose centroid is less than 20 miles from a sample monitor.

whole. This is a direct consequence of the EPA's requirement that monitors be located in densely populated areas.

If valuation of air quality differs by observable characteristics, then aggregating individuals over too large an area may suppress this heterogeneity in valuation. Clearly this is a concern for our analysis; the variance decomposition in Table 2.1 showed that as much as 90% of the variation can be concealed by using counties as the unit of analysis. In part motivated by this fact, a majority of the literature on hedonic valuation of housing attributes uses census tracts or even individual houses as the unit of analysis (e.g., Harrison and Rubinfeld 1978, Palmquist 1984, Greenstone and Gallagher 2008). Measuring valuation at the tract level is thus the natural starting point for our disaggregated analysis.

There are advantages and disadvantages with both the tract level and county level analysis, and neither is clearly superior. The disaggregated, tract level analysis focuses on valuation of air quality for an unrepresentative population. Because individuals that live close to a monitor are systematically different in terms of wealth, education and other characteristics that may be correlated with valuation of air quality, the estimates garnered at fine levels of disaggregation are difficult to generalize to the population as a whole.

Despite this drawback, there are two significant reasons to prefer the disaggregated analysis. First, at the tract level, we have a much more precise measure of the air quality that individuals in the sample actually face. There is considerable local heterogeneity in air quality, and Auffhammer et al. (2009) show that the CAAA work through local, monitor-level channels. As we aggregate our sample and increase the distance from monitors, we compromise the validity of the air quality measure included in our analysis. Thus, at fine levels of disaggregation we have more confidence that we are estimating capitalization accurately even if it is for an unrepresentative population. Second, this "unrepresentative population" is of inherent interest to those concerned about environmental justice and adverse distributional effects of large scale policy. By analyzing disadvantaged populations living in areas with the worst air quality, we can ascertain if these equity concerns are justified for the CAAA.

While neither the tract nor the county level analysis is perfect, both are informative. However, they are two end points of a spectrum of choices. In an effort to bridge the dichotomous choice of tract or county, we additionally create intermediate units of analysis by constructing concentric ring buffers around each monitor at distances of 0-1, 1-3, 3-5, 5-10 and 10-20 miles. The ring concept closely mirrors Greenstone and Gallagher (2008), who examine valuation of Superfund cleanup for census tracts containing the Superfund site, tracts neighboring the site, and all tracts within a 2- and 3-mile buffer of the site. Given the patterns shown in Figure 2.4, each of these rings will comprise a different slice of the population, with different housing prices, incomes, human capital, and racial composition. The results from estimating the same valuation model on each of these subpopulations will be suggestive of how each subpopulation values air quality improvements. For example, if estimated valuation increases with distance, then we could infer that valuation is correlated with income, education or race. To be clear, measurement error associated with air quality increases as the distance from the monitor increases, and interpretation of valuation estimates from large rings should be appropriately nuanced. That being said, aggregating air quality to the county level introduces even more measurement error. An additional reason to look at a larger scale than the tract is that the 23 monitors that could not be matched to tracts due to missing data can be added to the sample size for rings.

Our rings are constructed such that county lines and lines equidistant with other monitors in the same county truncate the rings. Consistent with Greenstone and Gallagher (2008) and Banzhaf and Walsh (2008), we aggregate housing and socioeconomic data for all tracts falling within a given ring, using weights equal to each tract's land area within the relevant ring multiplied by its population.[28]

[28] When tracts are aggregated to a ring level, median variable (i.e., median house value, median rent, median family income) lose their "median" and are weighted averages of medians. In this case, of course, there is no alternative data source from which to draw these median variables, as was done for the county level.

In prior research, distance is primarily a means of measuring exposure, be it to Superfund sites (Cameron and McConnaha, 2006; Greenstone and Gallagher, 2008; Gamper-Rabindran et al., 2011), toxic emissions (Banzhaf and Walsh, 2008), or power plants (Davis, 2011). However, all of those examples consider very localized disamenities. Whereas PM_{10}, while by no means a global pollutant, tends to be in similar concentrations for nearby areas. Thus, distance takes on a different role in our analysis. Instead of proxying for varying degrees of exposure, our rings identify different subpopulations that tend to have similar air quality and socioeconomic characteristics. Our objective is to build up our disaggregated, tract level model to the county level in an attempt to decompose the county valuation estimates by different subpopulations. Cumulatively, the rings cover 92% of total county population , and thus by estimating valuation for each of these rings, we can determine which populations in a county are driving the aggregate valuation. Further, it will allow us to comment on the incidence of the policy, which in turn permits us to relate our findings to recent structural work on the impact of environmental policies (Sieg et al., 2004; Tra, 2010).

Because the scope of our study is national and we compare appreciation rates across many cities and regions, it is necessary to control for local housing market trends. If housing price trends across regions are correlated with patterns of air quality improvements, it could bias our estimates of the effects of pollution reductions on home values. We institute a novel method to control for housing market trends that uses data from Freddie Mac on local house prices. In particular, we use Freddie Mac's conventional mortgage home price index (CMHPI), which gives quarterly estimates of home prices.[29]

[29] The CMHPI is an index of all mortgages purchased or securitized by Fannie Mae or Freddie Mac in a given year. The mortgages are only for single family homes and do not include mortgages insured by the federal government or

Freddie Mac offers MSA-specific indices for 11 large MSAs (Boston, Chicago, Dallas-Ft. Worth, Detroit, Los Angeles, Miami, New York, Philadelphia, San Francisco, Seattle, and Washington, DC) as well as indices for each state. For the purposes of our analysis, if a tract is within one of the 11 MSAs, then it is matched to that index; otherwise, it is matched to the state index. For each of these indices, we calculate a decadal appreciation rate from 1990 to 2000 using an average of the four quarterly estimates from those years. We then use those rates to convert housing prices in 1990 to year 2000 dollars. Given the adjustment for metro or statewide house price appreciation, the dependent variable represents appreciation relative to the local housing market trend. Thus, while the scope of our study is national, we can compare housing price changes from one part of the country to another because all price changes are relative to a smaller market. Importantly, though, to the extent that changes in PM_{10} levels affect housing prices metro-wide, our estimates of the effects of air pollution on home values will be attenuated.[30]

The 1990 Clean Air Act Amendments as a Quasi-Experiment

Attempts to estimate the extent to which improvements in air quality are capitalized into housing values are complicated by the fact that unmeasured or unobserved changes over time in local characteristics may affect both pollution and housing prices in a given area. For example, expansions in local transportation infrastructure or increases in overall economic activity could affect both housing prices and pollution levels. Such correlations, which we generally expect to

---

jumbo loans. While the CMHPI is based only on repeat transactions and therefore does not incorporate the prices of new homes, it does include appraisals associated with refinanced mortgages.

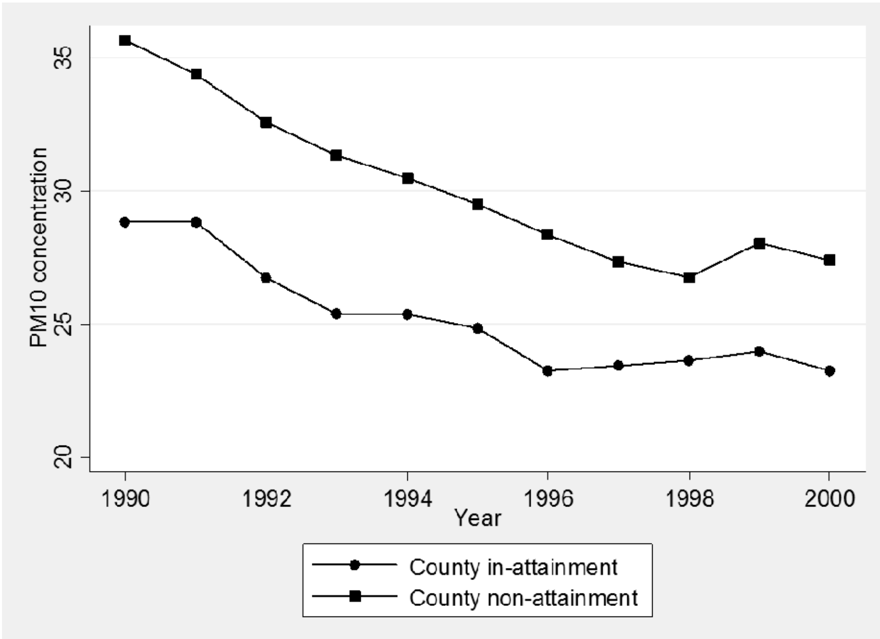
[30] Instead of using Freddie Mac's CMHCPI data, we have also attempted to use state and MSA time trends. However, since we often have only a small number of monitors per MSA or state, location-specific time trends absorb much of the relevant variation and render it difficult to distinguish the effect of changes in pollution of house values.

bias the coefficient on pollution toward zero, are at the root of the simultaneity problem that calls for an instrumental variable strategy. Our approach exploits the 1990 CAAA and its implications for local official behavior to address the simultaneity that would otherwise exist between house prices and pollution.

Our identification strategy follows that of Chay and Greenstone (2005), who instrument for changes in pollution at the county level between 1970 and 1980 using county nonattainment status in the mid-1970s. Like that of Chay and Greenstone (2005), our IV approach relies on the assumption that, conditional on other observable neighborhood and housing characteristics, nonattainment status only affects house prices through its impact on local pollution levels. We build on Chay and Greenstone (2005) not only by examining the effects of more recent legislation regarding air pollution, but also by exploiting heterogeneity within counties in pollution levels and socio-economic characteristics as well as in officials' behavior.

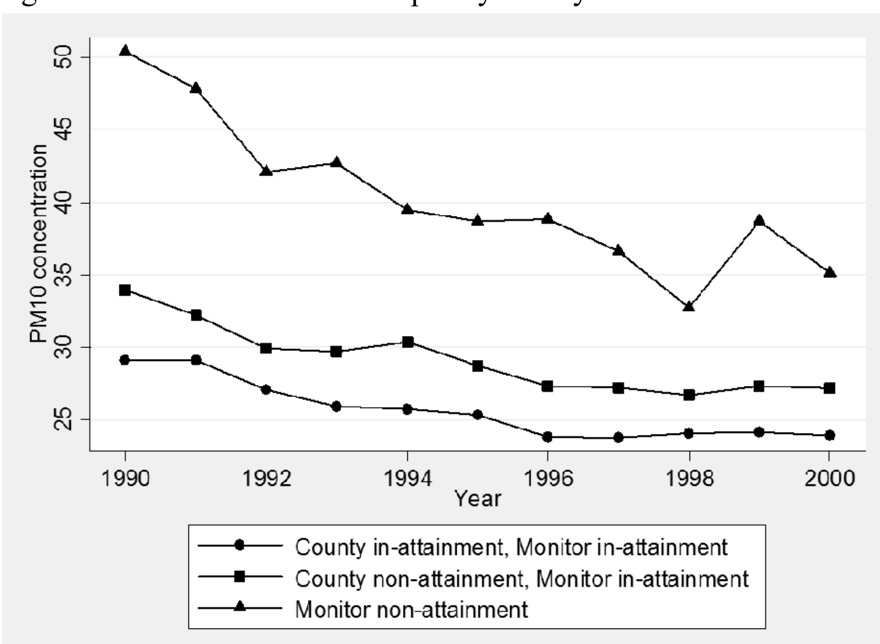
Chay and Greenstone (2005) evaluate how changes in TSPs at the county level induced by the 1970 and 1977 Clean Air Acts affect house prices at the county level. Specifically, they use the county level attainment designation in 1975 to instrument for 1970-1980 changes in air quality. They convincingly demonstrate that the county level designation is strongly correlated with reductions in TSPs. This relationship holds true in the 1990s as well. Figure 2.5a shows PM_{10} concentrations over the course of the 1990s for counties designated in attainment all years between 1992 and 1997 and for counties designated out of attainment at least one year between 1992 and 1997. Concentrations declined by 5.1 $\mu g/m^3$ in counties in attainment, but declined 8.0 $\mu g/m^3$ in out of attainment counties, suggesting county designation may be a good predictor of changes in PM_{10}.

Fig 2.5a: PM10 concentrations split by county attainment status



Notes: Monitor Sample includes all 375 monitors that are included in the analysis. A monitor is classified as 'county non-attainment' if it is located in a county that is non-attainment at somepoint during 1992-1997.

Fig 2.5b: PM10 concentrations split by county and monitor attainment status



Notes: Monitor Sample includes all 375 monitors that are included in the analysis. A monitor is classified as 'monitor non-attainment' if it exceeds either of the EPA standards at somepoint during 1992-1997. A monitor is classified as 'county non-attainment' if it is located in a county that is non-attainment at somepoint during 1992-1997, but is not non-attainment itself. All but one monitor designated 'monitor non-attainment' are located in non-attainment counties.

One of the major contributions of this paper is to allow for heterogeneous treatment effects. Figure 2.5b shows PM_{10} reductions for monitors in attainment located in counties in attainment, for monitors in attainment located in counties out of attainment, and for monitors out of attainment.[31] The decline in concentrations for in-attainment monitors located in attainment or nonattainment counties is similar to that shown in Figure 2.5a, 5.1 $\mu g/m^3$ and 7.0 $\mu g/m^3$, respectively. The levels of the county non-attainment group of monitors has lower concentrations is Figure 2.5b compared to 5a (in 1990, 34.0 in 5b versus 36.0 in 5a), reflecting the fact that the dirtiest monitors have been removed and put in their own group. It is this group, the out-of-attainment monitor group, that experiences by far the largest declines in concentrations over the decade, 15.4 $\mu g/m^3$. This pattern highlights the substantial within-county variation in pollution driven by the CAAA and its attainment designations. As we would expect, and as Aufhammer et al. (2009) show, PM_{10} reductions are localized and center around monitors responsible for inciting regulatory action.[32] That the EPA's nonattainment designations were generally effective in reducing pollution levels is consistent with past work documenting the enforcement of air quality standards under the CAAA (Henderson 1996, Nadeau 1997, Becker and Henderson 2000).

Our IV strategy uses both county attainment status and monitor attainment status to best capture the full and heterogeneous effects of the policy. Our county (monitor) instrument is the ratio of years that the county (monitor) is out of attainment to the number of years for which

[31] Monitors out of attainment that are located in counties that are in attainment are included among monitors out of attainment for the purposes of the figure. All but one monitor out of attainment are located in non-attainment counties.

[32] Even for counties and monitors out of attainment, average PM_{10} concentrations often fall below 50 on average. With respect to the non-attainment monitors, 14 of 27 monitors are designated out of attainment purely because they did not meet the secondary standard (i.e., PM_{10} level exceeded 150 $\mu g/m^3$ for 24 hours).

there is a record during the time span 1992-1997.[33] With the county instrument, the denominator is always 6 years; for the monitor instrument, due to some monitors not having valid data for all years, the denominator can vary from 1 to 6 years. We opt for a ratio instrument over binary to differentiate areas of persistent air quality violations, like southern California, from areas that infrequently violated the standards.

To the extent that there are unobservable differences in housing and demographic characteristics between areas that are correlated with the instrument, or that the CAAA itself caused changes in housing and demographic characteristics, our estimates of the house price-air quality gradient could be biased. Table 2.3 explores these differences in 1990 levels and decadal changes for many of the key variables used in the analysis at both the county level and tract level. In levels, there are clearly large differences; monitor non-attainment areas have the lowest house prices, the lowest median incomes, the lowest share of residents that are white, the highest unemployment rate, and the lowest share of houses with three or more bedrooms. These differences suggest that there are likely other unmeasured characteristics of locations that might bias cross-sectional estimates of the relationship between house prices and air quality. However, to the extent that these characteristics are time-invariant, a differencing approach will sweep out these effects.[34]

---

[33] Chay and Greenstone use only county attainment status in 1975 and 1976 as an instrument for changes in TSPs between 1970 and 1980. Because 62% of the observed monitor exceedances in our sample occurs before 1994 (78% outside of California), we expand the years of attainment status data to construct our instruments. Nonetheless, our instruments are still mid-decade, and as such, allow some time for local mid-decade shocks to dissipate on either side of the decade. They also still allow sufficient time for pollution to respond to nonattainment status, which based on Figures 2.5a and 2.5b, appears to take 2-3 years. Further details on the instrument construction as well as comparisons with alternative specifications appear in the Appendix.

[34] With regard to the changes between 1990 and 2000 reported in Table 2.3, changes in PM_{10} conform to our expectations, but surprisingly changes in median house values do not. While aggregate data at the county level show that house prices appreciated more in non-attainment counties, house prices at the tract level do not appreciate in line with changes in PM_{10}. This lack of correlation points to the importance of including other covariates that affect house values in our specification.

Table 2.3: Summary statistics split by attainment status for counties and tracts

		County in-attainment		County non-attainment			
		County	Tract	County	Tract		
					all monitors	monitor in-attainment	monitor non-attainment
		(1)	(2)	(3)	(4)	(5)	(6)
Sample size	Counties	173		57			
	Tracts		234		118	91	27
PM10 concentration	1990	29	29.3	36.7	38	34.3	50.7
	2000-1990	-5.8	-5.6	-9.5	-9.1	-7.3	-14.9
Median house value	1990	130,017	110,023	142,354	110,045	111,668	104,574
	2000-1990	-4,553	1,323	8,262	1,626	1,780	1,108
Median rent	1990	510	746	559	819	834	770
	2000-1990	-23	-215	-48	-275	-277	-268
Median family income	1990	45,804	36,953	45,259	38,842	40,189	34,303
	2000-1990	3,824	3,840	3,659	2,030	2,608	81
Share white	1990	79.30%	72.20%	79.60%	69.80%	71.30%	64.70%
	2000-1990	-5.00%	-6.30%	-6.00%	-9.20%	-8.90%	-10.30%
Share same house as 5 years ago	1990	52.60%	49.70%	50.20%	48.40%	49.90%	43.60%
	2000-1990	1.50%	-0.10%	1.40%	1.10%	1.00%	1.30%
Share unemployed	1990	6.60%	9.20%	7.50%	9.90%	9.50%	11.30%
	2000-1990	-0.50%	-0.40%	-0.80%	-0.50%	-1.10%	1.60%
Population density (per sq. mile)	1990	978	2,951	2,374	3,453	3,674	2,706
	2000-1990	65	3	78	161	188	67
Total housing units	1990	137,780	1,645	258,918	1,615	1,704	1,317
	2000-1990	14,846	100	26,600	104	109	87
Share owner-occupied housing units	1990	64.70%	51.40%	63.80%	53.70%	55.50%	47.80%
	2000-1990	1.50%	0.40%	2.40%	0.30%	0.40%	0.00%
Share housing units built last 10 years	1990	18.80%	11.20%	20.20%	12.40%	12.50%	12.30%
	2000-1990	-1.30%	-1.60%	-0.80%	-3.20%	-3.70%	-1.40%
Share housing units with 3 or more bedrooms	1990	71.90%	61.60%	69.90%	62.00%	64.70%	53.00%
	2000-1990	1.60%	0.50%	1.30%	-0.30%	-1.10%	2.40%

Notes: Housing and demographic data come from Neighborhood Change Database and USA Counties. 1990 housing prices are adjusted using the CMHPI. PM10 data are from Air Quality Standards Database.

There are two potential problems with our research design, and Table 2.3 helps shed light on both issues. First, our capitalization estimates may fail to capture willingness-to-pay because there is ample time for both supply response and residential sorting before we observe house prices a second time in 2000. Second, the CAAA may affect house prices through other channels than PM_{10}. With respect to the first point, Table 2.3 shows no statistically significant differences between in-attainment and non-attainment areas in the number of new units and the share of units with three or more bedrooms, and thus there is no evidence of a supply response. In terms of sorting, changes in population density and turnover rates in attainment vs. non-attainment counties as well as in tracts with attainment vs. non-attainment monitors were similarly not significantly different. However, the share of whites dropped substantially more in monitor non-attainment areas than other tracts.

With regard to the second point, we may be concerned that the regulation would hamper economic activity and employment or wages may suffer.[35] In line with this argument, median household income and the unemployment rate were statistically significantly different across some of the non-attainment group comparisons. Instead of a result of the policy, we contend that this pattern is due to regional trends, specifically economic trends in Los Angeles, which plays a prominent role in our monitor non-attainment treatment group. Given that Southern California experienced a recession in the early 1990s, these trends are not surprising. When California is removed from the sample, the economic differences across non-attainment areas no longer persist. Further, the change in share white is no longer significantly different, which again points to Los Angeles as the driving factor, with its vast numbers of immigrants during the 1990s. Thus,

[35] This effect of environmental regulation on the economy has been explored widely in the literature – e.g., Becker and Henderson (2000), Greenstone (2002), Hanna (2010), and Kahn and Mansur (2010).

changes in economic activity and population characteristics are not caused by the policy, only correlated with it and its geographic distribution.[36]

We argue based on these results that nonattainment status affects house prices only through its affect on air quality, and thus that our instruments satisfy the exclusion restriction. Further, these results suggest that ex-post sorting does not compromise the interpretation of our IV estimates of the relationship between pollution and house prices as the marginal willingness-to-pay.[37]

Econometric Approach

In this section, we outline our econometric approach to estimating the implicit value of air quality derived from housing market data. We rely on the hedonic model developed by Rosen (1974), which characterizes a market for heterogeneous goods and allows one to assign prices to the attributes of those goods. We estimate regressions at various levels of geographic aggregation in order to examine the extent to which air quality improvements are experienced differentially and how willingness-to-pay varies across certain segments of the population.

We begin with a cross-sectional analysis of the relationship between housing prices and PM_{10} concentrations. Our basic specification is

$$p_i = \theta PM_i + \mathbf{X}_i\boldsymbol{\beta} + \varepsilon_i \tag{2.1}$$

[36] We further explored the potential exclusion restriction and sorting problems by running a series of county- and tract-level univariate regressions of first differences in each variable in Table 2.3 on county and monitor nonattainment instruments. For tracts, in our specification with both instruments, neither county or monitor non-attainment status is significantly related to changes in housing or population characteristics, except a decline in the proportion of white people.

[37] Using microdata from the American Housing Survey and an approach that takes into account regional trends in demographics, Lang (2011) also finds little evidence of sorting among households with different characteristics and potentially different MWTP for improvements in air quality in response to changes in pollution prompted by the CAAA in the 1990s.

where p_i is the natural log of either median owner-occupied housing value or median rent in area i, PM_i is the concentration of PM_{10} in area i, and $\mathbf{X}_i$ is a vector of area i's housing and neighborhood characteristics, which are listed in Table 2.2.[38] In principle, the coefficient θ measures the gradient of housing prices with respect to air quality at a given moment in time. However, estimates of θ from a cross section may suffer from omitted variable bias, in that variables in $\mathbf{X}$ might fail to control for characteristics of areas that affect local housing prices and are also correlated with PM_{10} concentrations. To mitigate this bias, we also estimate a simple first-difference specification that controls for time-invariant features of locations:

$$\Delta p_i = \theta(\Delta PM_i) + (\Delta \mathbf{X}_i)\boldsymbol{\beta} + \Delta\varepsilon_i \tag{2.2}$$

where in our application, housing prices, PM_{10} concentrations, and the vector of controls are differenced between 2000 and 1990. This approach controls for both observable and unobservable time invariant characteristics of areas that might be correlated with house prices and air quality, such as climate and topographical features, transportation infrastructure, and population density. In the time differenced specification, θ is no longer measuring a cross-sectional gradient, but instead capitalization. If we assume that the importance of air quality in the valuation of housing did not change in the 1990s, then θ can be interpreted as the marginal willingness-to-pay for improvements in air quality.

Though more credible than the cross-section regressions, the results of the first-difference regressions can still be biased if there are unmeasured changes over time in local characteristics that affect both pollution and housing prices. To address the simultaneity problem, we exploit the

[38] As documented in Cropper et al. (1988), such a functional form performs well in terms of reducing errors in measuring marginal attribute prices.

structure of the 1990 CAAAs to construct instruments for PM_{10} reductions between 1990 and 2000. In addition to a county-level instrument much like that used by Chay and Greenstone (2005), we introduce monitor nonattainment status as an additional instrument and estimate effects at several levels of geographic aggregation. In particular, we consider tracts in which monitors are located as well as rings of 0-1, 1-3, 3-5, 5-10, and 10-20 miles around each monitor. The first stage and reduced form equations of the IV analysis are

$$\Delta PM_i = \varphi N_i + (\Delta \mathbf{X}_i)\boldsymbol{\beta} + \Delta\varepsilon_i \tag{2.3}$$

and

$$\Delta p_i = \gamma N_i + (\Delta \mathbf{X}_i)\boldsymbol{\beta} + \Delta v_i \tag{2.4}$$

where N_i is the instrument equal to the ratio of non-attainment years during the time span 1992 to 1997. When we use only county or monitor nonattainment as an instrument, the model is just identified, and the parameter φ captures the first-stage effect of nonattainment status on changed in PM_{10} concentrations, controlling for the covariates in **X** and any time-invariant features of area i. The parameter γ captures the reduced-form effect of nonattainment status on house prices, again controlling for the covariates in **X** and fixed characteristics of areas. The IV estimator in the just-identified model is simply $\theta_{IV} = \gamma/\varphi$.

Figures 2.5a and 2.5b suggest that while the effects of the CAAA regulations are most pronounced for non-attainment monitors, some additional reductions occur in non-attainment counties as well. Hence, in addition to models that include either the county-level or the monitor-level instrument, we also consider overidentified models that include both instruments in our tract and buffer analysis and may best capture the heterogeneous effects of the policy as well as spillovers that may occur from non-attainment monitors.

If the hedonic approach is founded on an individual-level model and we aggregate our data to tracts, rings, or counties, our estimates may suffer from aggregation bias. Non-linearities in individual relationships between pollution and house prices may not be captured using our approach. However, our focus on relatively fine geographic units helps to mitigate any aggregation bias. Further, our estimates presented below are similar to those of Lang (2011), which are based on individual housing units, suggesting that any aggregation bias is not severe.

In sum, cross-sectional regressions are likely to produce biased estimates of marginal willingness-to-pay for improvements in air quality. While mitigating omitted variable bias, first differenced regressions still potentially suffer from simultaneity. An IV approach using mid-decade nonattainment status holds out the promise of producing robust estimates of the hedonic price schedule gradient. However, we would expect estimates of rates of capitalization to vary across specifications that use different levels of geographic aggregation and that use different instruments. Estimates from county-level regressions capture average changes in capitalization owing to changes in pollution levels for counties. However, as previously discussed, there is substantial variation within counties in both changes in pollution and in demographic and economic characteristics. Results based on county-level regressions will fail to fully capture heterogeneity in air quality improvements and may mask heterogeneity in rates of capitalization across segments of the population. Indeed, as Chay and Greenstone (2005) point out, to the extent that taste heterogeneity and sorting within counties is greater than that between counties, county-level estimates will tend to understate individual-level dispersion in marginal willingness-to-pay.

Also, each IV estimate is a local average treatment effect (LATE), capturing the effects of nonattainment status on those counties, tracts, or rings whose air quality levels can be changed by each instrument. Given that the instrument is defined differently at different levels of geography, LATEs may vary across specifications. Each LATE measures the effect of treatment on those areas who were induced by the regulation to reduce pollution levels. The estimates do not apply to those who would have had reductions in pollution regardless of the law, nor do they apply to places that would not have had reductions regardless of the law. To the extent that the subpopulation affected by the instrument differs across specifications, we may expect estimated valuations to vary. Generally speaking, the monitor instrument identifies the effect based on relatively few, but disproportionately dirty places, whereas the county instrument identifies the effect based on relatively more, but generally cleaner areas. If populations have sorted based on preferences for air quality, then estimates of valuation would differ purely because the underlying populations are different. Our results in Section 7 bear on the issue of preference based sorting.

We do not view the potential variance in our estimates as a negative. On the contrary, we are uncovering heterogeneity and see value in each of our estimates. Further, our analysis enables us to address distributional concerns that have often been overlooked by previous studies on pollution regulations. By focusing on a traditionally disadvantaged population (poor, minority), we can address issues of environmental justice and whether these groups benefited more or less than others.

Results

We first present results from estimating Equations (2.1) and (2.2) for homeowners in Table 2.4. For the 1990 cross section (presented in the first panel), the coefficient estimates on PM_{10} suggest that the correlation between PM_{10} levels and house values is statistically and economically significant. The county-level estimates in the first column suggest that, controlling for other housing and population characteristics, PM_{10} levels one unit lower are associated with approximately 0.46% higher housing values among homeowners, which corresponds to an elasticity of -0.17. Turning to estimates based on more disaggregated information, the tract level estimates suggest housing prices are 0.32% higher for each unit reduction in PM10, implying an elasticity of -0.12. The analysis based on rings finds larger effects, especially for the 0-1 mile and 1-3 mile rings, which have estimated elasticities of -0.23 and -0.22, respectively. For the 2000 cross section (in the second panel), the estimate at the county level is fairly consistent with 1990 estimates coming in at 0.36% house price premium per unit of PM_{10}. However, estimates at the tract and ring levels are smaller and generally not significant.

In the first-differences model (shown in the third panel), which controls for time invariant features that might otherwise bias cross sectional estimates of the relationship between PM_{10} and home values, the range of estimated coefficients is of a similar magnitude. However, the estimated relationship between pollution and home values is now nearly three times larger at the tract level compared to the county level. Similarly, the elasticity at the county level is -0.07 compared to -0.23 at the tract level. Small radius rings also have larger coefficients than the

Table 2.4: Cross section and first difference regression results for homeowners

	County	Tract	0-1 miles	1-3 miles	3-5 miles	5-10 miles	10-20 miles
1990 Cross Section							
PM_{10} (1/100)	-0.46	-0.32	-0.60	-0.57	-0.44	-0.33	-0.44
	(0.15)***	(0.22)	(0.18)***	(0.14)***	(0.13)***	(0.11)***	(0.13)***
R-Squared	0.88	0.75	0.81	0.84	0.84	0.85	0.86
2000 Cross Section							
PM_{10} (1/100)	-0.36	-0.28	-0.26	-0.08	-0.15	-0.17	-0.33
	(0.17)**	(0.23)	(0.25)	(0.16)	(0.16)	(0.15)	(0.16)**
R-Squared	0.91	0.80	0.85	0.89	0.89	0.89	0.89
2000-1990 First Difference							
PM_{10} (1/100)	-0.20	-0.61	-0.44	-0.41	-0.33	-0.09	0.04
	(0.16)	(0.22)***	(0.18)**	(0.15)***	(0.14)**	(0.12)	(0.13)
R-Squared	0.47	0.29	0.43	0.50	0.44	0.50	0.46
Sample Size	230	352	375	375	373	370	334

Notes: All regressions include the full set of controls listed in the Table 2.2. Standard errors are shown in parentheses and are estimated using the Eicker-White formula to correct for heteroskedasticity and are clustered at the county level. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

county, but this effect dissipates as the ring gets larger. Variation in the magnitude of the association between PM_{10} and home values suggest that fresh insights may be gleaned from a closer analysis of the particular populations within counties that may be affected by targeted reductions in pollution. Specifically, this provides early evidence that the areas closest to monitors could be driving the aggregated, county level estimates.

Based on these results, we can calculate a MWTP and compare our estimates with those in the literature based on similar models. In doing so, it is important to keep in mind that much of the past work was based on pollution changes in the 1970s, often considered pollutants other than PM_{10}, and/or relied on data from single cities as opposed to data for the U.S. as a whole. To calculate an annual MWTP for a one-unit drop in PM_{10}, we assume an 8% interest rate and a 30-year mortgage. Based on our 1990 cross-sectional results in Table 2.4, MWTP estimates are $53 for the county level, $31 for the tract level, and range from $42 to $62 for the rings, with the largest value from the 1-3 mile ring. Based on our results from the first-differences model in Table 2.4, MWTP estimates are $23 for the county level, $59 for the tract level, and range from -$5 to $44 for the rings, with the largest value from the tract-level regression. These estimates are on the high side of those presented in the meta-analysis of Smith and Huang (1995), who find an interquartile range of $0-$10 for annual MWTP. However, the results are comparable with two locational equilibrium estimates of MWTP of ozone: $44 from Sieg et al. (2004) and $50 from Tra (2010).

As previously discussed, our cross-section and first-difference estimates, as well as those from past studies, are potentially biased due to unmeasured changes over time in local characteristics that might affect both pollution and housing prices. Our IV approach, which we turn to in the next section, addresses this possible source of bias.

Our main IV results appear in Tables 2.5 and 2.6. Table 2.5 presents IV estimates at the county level, which more closely approximate Chay and Greenstone's (2005) original work on the capitalization of pollution changes in the 1970s. Each specification includes the full set of controls listed in Table 2.2. The first column of Table 2.5 presents results using all monitors in the main sample and the 1995-1996 binary instrument. Column (2) uses a more expansive instrument that takes on a value of 1 if the county was out of attainment any year between 1992 and 1997. And lastly, column (3) uses our preferred ratio instrument, which equals the number of non-attainment years divided six (the number of years of mid-decade observation).

Comparing columns 1 and 2, the results are identical whether we use a binary instrument defined for only 1995-1996 or for 1992-1997. This is due to the fact that county designation does not change much over time; if a county is designated non-attainment in 1995 or 1996, then it is

Table 2.5: IV regression results at the county level for homeowners

	(1)	(2)	(3)
First Stage			
County Non-attainment	-3.39	-3.39	-3.77
	(0.97)***	(0.97)***	(1)***
F-stat	12.25	12.25	14.13
R-Squared	0.34	0.34	0.35
Second Stage			
ΔPM_{10} (1/100)	-1.05	-1.05	-0.74
	(0.63)*	(0.63)*	(0.54)
R-Squared	0.38	0.38	0.44
Sample Size	230	230	230

Notes: Column 1 uses all monitors in the main sample and the 95/96 binary instrument. Column 2 uses all monitors in the main sample and the 92-97 binary instrument. Column 3 uses the 92-97 ratio instrument. All regressions include the full set of controls listed in the Table 2.2. Standard errors are shown in parentheses and are estimated using the Eicker-White formula to correct for heteroskedasticity. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

likely to have been out of attainment sometime 1992-1994. As foreshadowed by the descriptive statistics and Figure 2.5a, the first-stage results in Table 2.5 imply that county non-attainment designation has a negative effect on PM_{10} changes over the course of the decade. Controlling for changes in other observable characteristics and exploiting a first-difference specification that controls for time invariant unobservable characteristics, the first-stage results in columns (1) and (2) imply that non-attainment counties experienced drops in PM_{10} 3.4 $\mu g/m^3$ larger than counties in attainment on average. The first-stage coefficient in column (3) suggest that counties that were always designated non-attainment experienced drops in PM_{10} 3.8 $\mu g/m^3$ larger than counties always in attainment on average. If a county was designated non-attainment half of years 1992 to 1997, then the drop in PM_{10} would be on average half of that, or 1.9 $\mu g/m^3$. The estimated effects of regulation are significant and the F-statistics are strong at 12.25 and 14.13, the latter being for the ratio instrument. The larger coefficient and F-statistic in column (3) indicates that the ratio instrument is capturing additional heterogeneity that separates counties that are always out of attainment versus those that are only sometimes out of attainment.

The second-stage results measuring the valuation of changes in PM_{10} are in the range of -0.0105 to -0.0074, which is four to five times larger than the county-level first difference results in Table 2.4. The corresponding elasticities are -0.38 and -0.27, for columns (2) and (3) respectively, which are similar to the range estimated by Chay and Greenstone (-0.20 to -0.35). Based on the same assumptions regarding interest rates and mortgage term, the corresponding MWTP estimates from the county-level IV model are in the range of \$92 to \$131. These estimates are larger than the cross sectional and first difference estimates from Table 2.4. They are also larger than the estimates derived from the locational equilibrium models of Sieg et al. (2004) and Tra (2010), which suggests that identification of exogenous changes in air quality affects valuation estimates. Thus, while structural models may be used to calculate general equilibrium estimates of the WTP that incorporate household relocation in response to price and air quality changes, their estimates may suffer from endogeniety.

The tract- and ring-level IV results appear in Table 2.6. At the tract level and for each ring, we present first- and second-stage estimates from three regressions. The first two columns in each set are just-identified models that use the county and monitor instruments, respectively. The third column is an overidentified model that uses both instruments, thus integrating the heterogeneous treatment effects of the CAAA in a single model. In each case, the instrument is the ratio of non-attainment years to the total number of observation years.

The first-stage results at the tract- and ring-levels using only the county instrument are similar to those from the county-level model in Table 2.5. For instance, tracts with monitors that are located in counties that are always out of attainment experience drops in PM_{10} that are about 3.5 $\mu g/m^3$ larger than tracts with monitors that are in counties always in attainment. Estimates for rings of up to 20 miles around these tracts are very similar, ranging from 3.0 to 4.1 $\mu g/m^3$. All the estimates are highly significant, and with F-statistics ranging from 13.5 to 24.2, the instruments appear to be highly relevant.[39]

As expected, the monitor instrument is an even stronger predictor of changes in PM_{10} than the county instrument. Relative to tracts with monitors always in attainment, tracts with monitors always out of attainment experience a 13.0 $\mu g/m^3$ decline in PM_{10}. However, most monitors in the treatment group have an instrument value less than one, and the average instrument value is 0.4. Thus, for the average monitor in the monitor non-attainment treatment group, the expected PM10 decline is 5.1 $\mu g/m^3$. The first-stage estimates using the monitor

[39] Staiger and Stock (1997) suggest that when the F-statistic exceeds 10, as it does in all the specifications in Table 2.6, weak instrument bias is small.

Table 2.6: Instrumental variables regression results for tract and ring level analysis for homeowners

	Tract			0-1 mile			1-3 mile		
First Stage	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
County Non-attainment	-3.36		-2.53	-3.37		-2.75	-3.24		-2.62
	(0.85)***		(0.85)***	(0.78)***		(0.78)***	(0.82)***		(0.82)***
Monitor Non-attainment		-12.99	-10.76		-12.15	-10.01		-12.44	-10.49
		(2.53)***	(2.4)***		(2.91)***	(2.7)***		(2.52)***	(2.41)***
F-stat	15.60	26.32	18.49	18.78	17.39	15.24	15.54	24.44	18.17
R-Squared	0.23	0.24	0.27	0.27	0.25	0.25	0.29	0.30	0.32
Second Stage									
ΔPM_{10} (1/100)	-0.86	-0.95	-0.92	-1.71	-1.08	-1.39	-1.54	-0.86	-1.16
	(0.9)	(0.48)**	(0.53)*	(0.84)**	(0.5)**	(0.5)***	(0.83)*	(0.3)***	(0.41)***
R-Squared	0.29	0.29	0.29	0.33	0.41	0.38	0.40	0.49	0.46
Sample Size	352	352	352	375	375	375	375	375	375

	3-5 mile			5-10 mile			10-20 mile		
First Stage	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
County Non-attainment	-2.97		-2.25	-4.06		-3.02	-3.90		-2.68
	(0.81)***		(0.76)***	(0.83)***		(0.7)***	(0.83)***		(0.68)***
Monitor Non-attainment		-13.74	-12.28		-15.49	-12.79		-15.81	-13.17
		(2.63)***	(2.44)***		(2.61)***	(2.39)***		(2.44)***	(2.33)***
F-stat	13.53	27.21	22.52	24.16	35.19	30.68	22.07	42.12	28.45
R-Squared	0.32	0.28	0.32	0.25	0.27	0.31	0.29	0.33	0.36
Second Stage									
ΔPM_{10} (1/100)	-1.88	-0.66	-1.03	-0.96	-0.48	-0.67	-0.28	-0.04	-0.12
	(0.88)**	(0.31)**	(0.35)***	(0.57)*	(0.24)**	(0.28)**	(0.56)	(0.23)	(0.28)
R-Squared	0.17	0.42	0.38	0.40	0.48	0.46	0.45	0.46	0.46
Sample Size	373	373	373	370	370	370	334	334	334

Notes: All regressions include the full set of controls listed in the Table 2.2 and use the ratio instruments constructed from years 1992-97. Standard errors are shown in parentheses and are estimated using the Eicker-White formula to correct for heteroskedasticity and are clustered at the county level. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

instrument are similar for different rings. The F-statistics for these models are even larger than the county instrument, ranging from 17.4 to 42.1, again suggesting a strong first-stage.

Finally, we include both instruments in overidentified models that appear in columns (3), (6), (9), (12), (15), and (18).[40] Both county nonattainment and monitor nonattainment are associated with declines in PM_{10}. Echoing the descriptive statistics, the results imply that the largest drops in PM_{10} occur near offending monitors that are located in counties out of attainment, while smaller drops occur near in attainment monitors that are located in counties out of attainment. Both instruments decline in magnitude from the just-identified models to the overidentified models suggesting the coefficients in the just-identified models are picking up some of the variation caused by the other instrument.

Turning to the IV estimates on the effects of changes in PM_{10} on capitalization, the second-stage coefficient on PM_{10} when using only the county instrument is -0.0086 at the tract level and ranges from -0.0028 to -0.0188 for the rings. All estimates are significant at the 10% level or better, except at the tract level and the 10-20 mile ring. Corresponding elasticities range from -0.11 to -0.71. Consistent with the first difference results, magnitudes tend to be larger for tight rings around monitors (up to 5 miles), but decline by over one-half going from the 3-5 mile ring to the 5-10 mile ring and by another two-thirds to the 10-20 mile ring.

Using only the monitor instrument, the second-stage estimate of the effects of PM_{10} reductions on home values is similar at the tract level, -0.0095, but for the rings tends to be smaller than when using the county instrument, ranging from close to zero to -0.0108. All estimates are significant at the 5% level, except the 10-20 mile ring, and elasticities range from -0.02 to -0.55. A similar pattern of declining magnitudes as rings get larger is observed.

[40] The 3-5 mile ring regression in column (12) fails the Sargan test for overidentification. We do not put too much stock in the results at these tests, as our data are spatial related and the Sargan test cannot account for that.

Finally, using both instruments, the estimated effect of PM_{10} reductions on home values is -0.0092 for the tract level and ranges from -0.0012 to -0.0139 for the rings. These coefficients, which are all significant at least at the 10% level except for the 10-20 mile buffer, are close to the average between the estimates using only the county instrument or only the monitor instrument. This highlights the local nature of our IV estimates. The IV estimator in our model that includes both the county and the monitor instruments represents a weighted average of the causal effects for the instrument-specific affected populations. The affected populations are different for each instrument, in one case comprising tracts in counties out of attainment (whether their monitors were out of attainment or not) and in the other comprising tracts with monitors out of attainment. The estimate we obtain with multiple instruments is simply a linear combination of the instrument-specific LATEs from using the instruments one at a time (Imbens and Angrist 1994).

It might come as a surprise that estimates of the effect of pollution reductions on home values that rely on the monitor instrument alone are smaller than those that rely on the county instrument alone. This again underscores the point that each estimate is a LATE. To the extent that within counties, those living very close to nonattainment monitors have different characteristics than those living further away, each estimate will capture a different MWTP. As previously discussed, those who live close to monitors tend to have lower income relative to those who live further away, and as such, our results are consistent with a large body of literature that suggests that the MWTP for reductions in pollution increases with income.

Estimates of MWTP tend to differ depending on the model and vary substantially by distance. The three models estimated at the tract level in Table 2.6 actually provide remarkably similar estimates of MWTP, ranging from \$84 to \$90. Based on the models in Table 2.6 that use only the county instrument, estimates of the MWTP for a one-unit decrease in PM_{10} range from

$37 to $181. Based on the models that use only the monitor instrument, estimates of the MWTP for a one-unit decrease in PM_{10} range from $5 to $123.[41] In our model with multiple instruments, we can calculate MWTP for those in neighborhoods around a monitor in attainment but who were located in a non-attainment county as well as one for those in neighborhoods around a monitor out of attainment (and in a county out of attainment). The MWTP for a one-unit decrease in PM_{10} for the first group ranges from $16 to $152, and the MWTP for the second group is a very similar $16 to $130. These estimates are two to three times larger than the corresponding tract and ring level estimates from the first differenced model in Table 2.4 and on either side of the preferred county level estimate from Table 2.5.[42]

Our previous results include California, which could be considered an outlier given the harsh recession and sharp pullback in house prices in the early 1990s, as well as some of the dirtiest air in the country. In order to gauge the influence of California on our main results, Table 2.7 presents the tract- and ring-level IV regression results for the sample excluding California. The first-stage estimates of the effects of non-attainment at the county level, the monitor level, or both levels are similar to those based on the sample excluding California however, the standard errors are larger and F-statistics are lower. The second-stage IV estimates tend to be similar in magnitude and exhibit the same decline as the radius increases, but again less significant compared to the sample that includes California.

[41] These estimates are consistent with Lang (2011), who used unit-level housing data from the American Housing Survey.
[42] We conducted several other tests to verify the robustness of the results. First, we explored the effect of relaxing the reliability requirements on monitors on our main tract and ring level results. Without the reliability requirements, an additional 216 monitors can be included in the sample, a 58% increase over the main sample. While the first stage results are weaker as a result of relaxing the reliability requirements on monitors, the second stage coefficients on PM_{10} are consistent with those from Table 2.6, which show diminishing magnitude as distance grows. We also investigated the robustness of results to changes in covariates. First stage results are very robust and change little across specifications. However, the second-stage results are sensitive to inclusion of population characteristics, which is not surprising given the relatively weak unconditional correlation between attainment status and house values discussed in Section 4.

Table 2.7: Instrumental variables regression results for tract and ring level analysis for homeowners, excluding California

	Tract			0-1 mile			1-3 mile		
First Stage	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
County Non-attainment	-3.32		-2.71	-3.36		-2.86	-3.24		-2.89
	(0.89)***		(0.92)***	(0.83)***		(0.85)***	(0.93)***		(0.95)***
Monitor Non-attainment		-17.81	-13.63		-16.11	-11.79		-13.48	-9.68
		(4.38)***	(4.29)***		(5.19)***	(4.96)**		(5.26)**	(5.13)*
F-stat	13.91	16.57	12.44	16.55	9.64	10.68	12.15	6.56	8.18
R-Squared	0.22	0.21	0.25	0.25	0.23	0.21	0.28	0.25	0.29
Second Stage									
ΔPM_{10} (1/100)	-1.03	-0.97	-1.00	-1.71	-0.86	-1.41	-1.39	-0.81	-1.23
	(0.99)	(0.86)	(0.76)	(0.91)*	(0.56)	(0.67)**	(0.89)	(0.53)	(0.68)*
R-Squared	0.31	0.31	0.31	0.38	0.46	0.42	0.48	0.55	0.51
Sample Size	301	301	301	323	323	323	323	323	323

	3-5 mile			5-10 mile			10-20 mile		
First Stage	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
County Non-attainment	-2.65		-2.08	-3.35		-2.72	-3.18		-2.53
	(0.82)***		(0.82)**	(0.77)***		(0.73)***	(0.74)***		(0.72)***
Monitor Non-attainment		-17.55	-14.99		-18.05	-14.31		-17.81	-14.26
		(3.97)***	(3.81)***		(5.09)***	(4.62)***		(4.12)***	(3.74)***
F-stat	10.39	19.60	13.19	18.78	12.58	15.39	18.40	18.68	17.71
R-Squared	0.29	0.27	0.28	0.26	0.25	0.29	0.36	0.36	0.38
Second Stage									
ΔPM_{10} (1/100)	-1.90	0.03	-0.76	-0.63	0.58	-0.06	0.01	0.46	0.24
	(1.06)*	(0.44)	(0.55)	(0.72)	(0.31)*	(0.43)	(0.72)	(0.38)	(0.43)
R-Squared	0.21	0.48	0.46	0.54	0.54	0.58	0.54	0.52	0.53
Sample Size	321	321	321	318	318	318	283	283	283

Notes: All regressions include the full set of controls listed in the Table 2.2. Standard errors are shown in parentheses and are estimated using the Eicker-White formula to correct for heteroskedasticity and are clustered at the county level. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

In all, the results are consistent with the main results in Table 2.6, though the results suggest a weaker relationship between changes in pollution and home values. This is not surprising given that much of the treatment group, especially the monitor non-attainment treatment group, is discarded when California is removed. Notably, though, the broader patterns hold whether we include California or not; the CAAA and the behavioral response of officials to the regulations affected some areas more than others, and treated locations generally experienced increases in home values. In addition, the consistency suggests that our approach of using the CMHPI to detrend house price changes from across the country is effective.

For comparison purposes, we present the tract- and ring-level IV regression results for renters in Table 2.8. As in the homeowner first-stage regressions, both county and monitor instruments are strong predictors of changes in PM_{10}, even though the F-statistics are generally slightly smaller than the corresponding value for the homeowner analysis.[43] The second-stage IV coefficients flip sign and are never significant at even the 10% level, suggesting the relationship between rental prices and air quality is substantially weaker than it is for owner occupied units. These results are consistent with a large body of evidence that demonstrates that changes in amenities are capitalized to a much smaller degree among renters than among homeowners (Greenstone and Gallagher 2008, Grainger 2010, Davis 2011).

Welfare analysis

We now take the preferred estimates from both the county and monitor-level IV models presented above to examine the change in housing prices, overall benefits, and the distribution of benefits that accrue to different groups of the population.

[43] Note that the set of covariates included the regressions for renters differ slightly from those included the regressions for homeowners.

Table 2.8: Instrumental variables regression results for tract and ring level analysis for renters

	Tract			0-1 mile			1-3 mile		
First Stage	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
County Non-attainment	-3.22		-2.30	-2.90		-2.18	-3.00		-2.31
	(0.88)***		(0.87)***	(0.81)***		(0.8)***	(0.84)***		(0.84)***
Monitor Non-attainment		-13.87	-11.89		-11.80	-10.05		-11.82	-9.99
		(2.5)***	(2.47)***		(2.42)***	(2.37)***		(2.51)***	(2.41)***
F-stat	13.51	30.75	21.11	12.95	23.89	15.63	12.65	22.19	14.61
R-Squared	0.18	0.22	0.24	0.24	0.22	0.24	0.24	0.25	0.27
Second Stage									
ΔPM_{10} (1/100)	1.62	-0.01	0.53	1.31	0.26	0.66	-0.23	-0.40	-0.33
	(1.04)	(0.49)	(0.56)	(1)	(0.61)	(0.62)	(0.83)	(0.39)	(0.46)
R-Squared	0.21	0.34	0.33	0.16	0.33	0.29	0.37	0.37	0.37
Sample Size	352	352	352	375	375	375	375	375	375

	3-5 mile			5-10 mile			10-20 mile		
First Stage	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
County Non-attainment	-2.74		-1.91	-3.89		-2.92	-3.40		-2.08
	(0.89)***		(0.83)**	(0.78)***		(0.68)***	(0.86)***		(0.73)***
Monitor Non-attainment		-13.30	-11.82		-14.89	-12.33		-15.22	-12.90
		(2.68)***	(2.53)***		(2.66)***	(2.35)***		(2.4)***	(2.37)***
F-stat	9.44	24.65	15.93	25.17	31.34	27.21	15.56	40.19	24.33
R-Squared	0.27	0.23	0.27	0.24	0.26	0.30	0.29	0.34	0.36
Second Stage									
ΔPM_{10} (1/100)	-0.05	-0.19	-0.15	0.18	-0.21	-0.05	0.18	-0.35	-0.20
	(0.9)	(0.4)	(0.39)	(0.57)	(0.38)	(0.35)	(0.74)	(0.33)	(0.36)
R-Squared	0.35	0.35	0.35	0.38	0.39	0.39	0.33	0.34	0.35
Sample Size	373	373	373	370	370	370	334	334	334

Notes: All regressions include the full set of controls listed in the Table 2.2, with the appropriate substitutions made for the renter-specific variables. Standard errors are shown in parentheses and are estimated using the Eicker-White formula to correct for heteroskedasticity and are clustered at the county level. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

Table 2.9 displays the breakdown of the benefits by treatment group (monitor non-attainment areas and county non-attainment areas) and by ring. Columns (1) and (2) present the average 1990 median income and 1990 median house value. The patterns shown here echo the patterns shown in Figure 2.4 and Table 2.3 with incomes and housing values being larger in the county non-attainment areas, as well as increasing with radius for both groups. Column (3) gives the total value of owner-occupied housing. Total value is substantially larger for county non-attainment areas compared to monitor non-attainment areas which just reflects the fact that county non-attainment areas are more abundant, and thus as a group have more total housing units, than monitor non-attainment areas. Of particular interest are columns (4), (5) , (6), and (7), which display the estimated appreciation per house due the CAAA and the estimated total appreciation for owner occupied housing due to CAAA based on the results of the county model of Table 2.5 and the monitor model of Table 2.6. To compute the appreciation per house due to the CAAA, we multiply the capitalization coefficient from the second stage of the IV model by estimated reduction in PM_{10} caused by the CAAA (the PM_{10} coefficient from the first stage of the IV times the instrument) and the 1990 median house value. To calculate total appreciation of owner occupied housing, we multiply appreciation per house by the number of housing units.

Our estimates of the overall benefits of the CAAA are found by vertically summing columns (6) and (7), the result of which appears in the last row of those columns. The two estimates of overall benefits are remarkably close, $36.0 billion for the county model and $36.8 billion for the monitor model.[44] We view this parity as a validation of our empirical design. By

[44] This estimate is smaller than the $45 billion of total benefits that Chay and Greenstone (2005) report for the 1970s. Our estimate would grow if we extended it to counties not in the sample, as they did. However, we feel this is not consistent with the idea of LATE.

Table 2.9: Breakdown of benefits for treated groups by instrument status and ring radius

Ring radius and instrument status	1990 median income	1990 median house value	1990 total value of owner-occupied housing (millions)	Appreciation per house due to CAAA		Total appreciation of owner occupied housing due to CAAA (millions)	
				County model	Monitor model	County model	Monitor model
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
0-1 mile							
County non-att.	42,089	124,632	35,600	3,315	4,560	917	1,260
Monitor non-att.	35,576	107,098	4,860	2,702	9,982	112	419
1-3 mile							
County non-att.	45,079	129,690	161,000	3,454	3,801	4,240	4,660
Monitor non-att.	40,629	120,344	36,400	3,029	9,541	890	2,850
3-5 mile							
County non-att.	49,043	141,629	201,000	3,791	3,162	5,480	4,570
Monitor non-att.	42,840	130,233	60,500	3,273	9,684	1,580	4,640
5-10 mile							
County non-att.	49,686	142,595	404,000	3,843	2,814	11,100	8,150
Monitor non-att.	47,162	144,202	150,000	3,625	7,974	4,060	8,040
10-20 mile							
County non-att.	51,341	155,476	248,000	4,203	488	6,770	786
Monitor non-att.	47,634	150,591	62,000	3,748	1,496	1,700	624
Total Benefits						35,999	36,849

Notes: All dollar amounts are adjusted to 2000 levels. Columns 4 and 6 rely on the estimates from the county model results shown in Column 3 of Table 2.5. Columns 5 and 7 rely on the estimates from the monitor model results shown in Columns 3, 6, ..., 18 of Table 2.6.

constructing rings around monitors and allowing valuation to vary by ring, we are decomposing the total county level benefits into various sub-populations.

There are non-trivial changes in the value of owner-occupied housing, as shown in columns (4) and (5). The county model estimates of appreciation per house, shown in column (4) are relatively constant across the radii and attainment groups, ranging from \$2,700 to \$4,200. The appreciation is about \$500 greater for the county non-attainment group compared to the monitor non-attainment group for most rings, and appreciation increases by about \$1,000 as the radius increases from 0-1 miles to 10-20 miles. Because the estimated PM_{10} reduction and the estimated valuation are nearly identical across radii and groups for the model, these differences are driven entirely by house values, which increase with distance from monitors and are smaller in monitor non-attainment areas (shown in column (2), but also in Figure 2.4 and Table 2.3).

The appreciation per house estimates for the monitor model, shown in column (5), contrast with the uniform estimates of column (4). On average, the appreciation per house for non-attainment counties is \$4,560, \$3,801, \$3,162, \$2,814, and \$488 for houses located in rings 0-1 miles, 1-3 miles, 3-5 miles, 5-10 miles and 10-20 miles, respectively. For houses located in monitor non-attainment areas, the average appreciation per house is \$9,982, \$9,541, \$9,684, \$7,974, \$1,496 for houses located within 0-1 miles, 1-3 miles, 3-5 miles, 5-10 miles and 10-20 miles, respectively. Both estimates show near monotonic declines as the radius increases, dropping off precipitously in the 10-20 mile ring. This pattern is a result of the decline in valuation coefficients, shown in Table 2.6, as the radius grows. Most striking is that appreciation is two to three times larger in monitor non-attainment versus county non-attainment areas. This disparity is due to the substantially larger declines in PM_{10} in areas surrounding non-attainment monitors, and the ability of the monitor level model to account for that heterogeneity.

Despite the similarity between the county and monitor models in estimated total benefits, comparing columns (4) and (5) highlights that the distribution of the benefits across space critically depends on the model. For the 0-1 mile ring, the monitor model implies that appreciation per house was 38% larger for county non-attainment areas and a whopping 369% larger for monitor non-attainment areas compared to the county model. For the 10-20 mile ring, the monitor model implies that appreciation per house was 88% smaller for county non-attainment areas and 60% smaller for monitor non-attainment areas compared to the county model. In general, the county model tends to underestimate appreciation close to monitors and severely underestimate appreciation in monitor non-attainment areas (up to a radius of 10 miles). These differences are striking and could lead to incorrect conclusions about the distributional impacts of the policy.

Columns (6) and (7) provide estimates of the total benefits accruing to each group in each ring as a result of CAAA. These columns equal columns (4) and (5) multiplied by the number of housing units in each area. Since the appreciation estimates of the county model and monitor model are being multiplied by the same number of housing units, the disparities present in columns (6) and (7) are exactly the same as in (4) and (5).

We do not provide a direct test of the incidence of the CAAA. However, the combination of Figure 2.4, Table 2.3, and the calculations displayed in Table 2.9 based on the monitor model suggest that the policy may actually have been progressive. As previously documented, the poor and minorities disproportionately live close to monitors. These are the populations that appear to be reaping most of the benefits of the program. A major advantage of the approach taken here is that it allows us to learn how the benefits of the program are distributed across space, shedding light on the potential incidence to specific populations of interest.

In Table 2.10, we display statistics and estimates related to the welfare implications of the CAAA for census tracts in either the monitor non-attainment or county non-attainment treatment group. The first three columns present median income, median house price, and average PM_{10} concentration, respectively. Column (4) presents estimates of MWTP for a 1 $\mu g/m^3$ reduction in PM_{10}. Column (5) gives estimates of WTP, which is what our estimates imply a household is willing to pay in annual housing expenditure to live in an area that experienced the declines associated with the CAAA. Finally, column (6) gives estimates of proportional WTP, which equals WTP divided by median income. This calculation is done at the tract level and then summarized, which is why column (6) does not necessarily equal column (5) divided by column (1).

The first row presents overall means for all tracts in the treated groups. The results suggest that, on average, the 1990 CAAA provided significant benefits to households. We find a MWTP of $88 and a WTP of $302 in affected tracts. The latter corresponds to 0.82% of the median income.

The next set of rows in Table 2.10 show the same statistics and estimates for the top, middle, and bottom of the income distribution. The bottom 10% is comprised of tracts with an average 1990 median income of $16,501, whereas the top 10% includes tracts with an average 1990 median income of $72,534. Column (3) confirms our intuition that the poorest people live in areas with the worst air quality on average. Consistent with prior literature (e.g., Sieg et al. 2004), the MWTP for air quality improvements displayed in column (4) increases with income. We find that the MWTP for air quality improvements were $46, $80, and $178 for households in the bottom 10%, middle 10%, and top 10% of income, respectively. The magnitude of increases from MWTP to WTP (which are entirely a function of the magnitude of changes in PM_{10}

Table 2.10: Welfare calculations for treated tracts, by 1990 income and pollution levels

	Median Income	Median house price	Average PM10 concentration	MWTP	WTP	Proportional WTP
	(1)	(2)	(3)	(4)	(5)	(6)
Overall mean	38,842	110,045	38	88	302	0.82%
Sorted by Income						
Bottom 10%	16,501	56,633	42	46	169	1.03%
Middle 10%	35,542	99,639	31	80	268	0.76%
Top 10%	72,534	221,198	32	178	502	0.72%
Sorted by PM10						
Bottom 10%	29,568	99,662	67	80	660	2.15%
Middle 10%	30,319	91,160	37	73	171	0.64%
Top 10%	48,660	149,639	19	120	323	0.71%

Notes: Calculation of MWTP and WTP rely on estimates from the tract-level specification using both monitor and county instruments from Table 2.6. Proportional WTP equals estimated WTP divided by median income.

resulting from the CAAA) are about the same across income groups, although they diminish slightly as income increases. This suggests that the monitor non-attainment treatment group is slightly more represented in the low income group, though still spread throughout the income distribution. Column (6) shows that, despite WTP increasing with income, proportional WTP is actually largest in the low income group and smallest in the high income group, which suggests that lower income households are willing to pay a greater percentage of their income for clean air. This result implies that the lowest income group benefits disproportionally relative to the middle and high income groups.

The final set of rows in Table 2.10 examine the distribution of benefits for treated tracts sorted by 1990 PM_{10} levels. Since 1990 PM_{10} levels and 1990 median income are correlated, column (1) shows that the most polluted areas had an average median income less than the other groups, though the disparity is much smaller compared to when tracts are sorted on income. However, the average PM_{10} concentration is 48 $\mu g/m^3$ higher in the dirtiest 10% compared to the cleanest 10%, compared to only 10 $\mu g/m^3$ higher when sorted on income. The estimates of MWTP are \$80, \$73, and \$120 for the dirtiest, middle, and cleanest groups, respectively. The house prices in the dirtiest group are actually larger than those in the middle group, which is driving the inversion of the relationship between MWTP and income. The magnitude of the increase from MWTP to WTP is substantially greater for the dirtiest areas relative to the middle and cleanest areas (by a factor of 8.25 for the dirtiest compared to 2.34 for the middle and 2.69 for the cleanest). This is a result of the monitor non-attainment treatment group being disproportionately in the dirtiest group, which is bound to be the case when the average 1990 concentration is well above the annual standard of 50 $\mu g/m^3$. Estimates of proportional WTP again show that the residents of the dirtiest and poorest areas were willing to pay a greater share of their income for air quality improvements, and that the benefits resulting from the CAAA disproportionately accrued to those residents.

Prior literature has suggested that the design of the CAA and its amendments may be inefficient by seeking to improve the air quality in the dirtiest places. If households have heterogeneous MWTP and have already sorted into areas consistent with their air quality preferences, then total benefits to an air quality policy may be maximized by improving air quality in the areas that are already clean. The results of Tables 2.9 and 2.10 dispute this idea and strongly suggest that the CAAA were a progressive policy. The very households living in the areas with the worst air quality are those that benefited most from the improvements. Thus, our results defend the design and implementation of the CAAA. By placing standards on criteria pollutants, the dirtiest areas will be cleaned up as local regulators respond to incentives, and the disadvantaged residents of the dirtiest areas will benefit.

REFERENCES

Auffhammer, M.; A. M. Bento and S. E. Lowe. 2009. "Measuring the Effects of the Clean Air Act Amendments on Ambient Pm10 Concentrations: The Critical Importance of a Spatially Disaggregated Analysis." *Journal of Environmental Economics and Management*, 58(1), 15-26.

Banzhaf, S. and R. P. Walsh. 2008. "Do People Vote with Their Feet? An Empirical Test of Tiebout's Mechanism." *American Economic Review*, 98(3), 843-63.

Becker, R. and J. V. Henderson. 2000. "Effects of Air Quality Regulations on Polluting Industries." *Journal of Political Economy*, 108(2), 379-421.

Cameron, T. A. and I. T. McConnaha. 2006. "Evidence of Environmental Migration." *Land Economics*, 82(2), 273-90.

Chay, K. Y. and M. Greenstone. 2005. "Does Air Quality Matter? Evidence from the Housing Market." *Journal of Political Economy*, 113(2), 376-424.

Cropper, M., L. Deck, and K. McConnell. 1988. "On the Choice of Functional Forms for Hedonic Price Functions." *Review of Economics and Statistics* 70(4), 668-675.

Currie, J. and M. Neidell. 2005. "Air Pollution and Infant Health: What Can We Learn from California's Recent Experience?" *Quarterly Journal of Economics*, 120(3), 1003-30.

Davis, L. 2010. "The Effect of Power Plants on Local Housing Values and Rents." *Review of Economics and Statistics*, Forthcoming.

Gamper-Rabindran, S., R. Mastromonaco, and C. Timmins. 2011. "Valuing the Benefits of Superfund Site Remediation: Three Approaches to Measuring Localized Benefits." NBER Working Paper No. 16655.

Grainger, Corbett. 2009. "The Distributional Effects of Pollution Regulations: Rental Housing and Air Quality Improvements." University of California-Santa Barbara Working Paper.

Greenstone, M. and J. Gallagher. 2008. "Does Hazardous Waste Matter? Evidence from the Housing Market and the Superfund Program." *Quarterly Journal of Economics*, 123(3), 951-1003.

Hall, J. V.; A. M. Winer; M. T. Kleinman; F. W. Lurmann; V. Brajer and S. D. Colome. 1992. "Valuing the Health Benefits of Clean-Air." *Science*, 255(5046), 812-17.

Harrison, D. and D. L. Rubinfeld. 1978. "Hedonic Housing Prices and Demand for Clean-Air." *Journal of Environmental Economics and Management*, 5(1), 81-102.

Henderson, J. V. 1996. "Effects of Air Quality Regulation." *American Economic Review*, 86(4), 789-813.

Imbens, G. and J. Angrist. 1994. “Identification and Estimation of Local Average Treatment Effects.” *Econometrica* 62(2), 467-475.

Lang, C. “The Dynamics of Air Quality Capitalization and Locational Sorting in the Housing Market.” Cornell University Working Paper.

Nadeau, L. W. 1997. "EPA Effectiveness at Reducing the Duration of Plant-Level Noncompliance." *Journal of Environmental Economics and Management*, 34(1), 54-78.

National Archives and Records Administration. 1987. Federal Register title 40 public health, Chapter 50, United States Code of Federal Regulations.

National Archives and Records Administration. 2005. Federal Register title 42 public health, chapter 85, subchapter 1, part D, subpart 1, par. 7509, United States Code of Federal Regulations.

Palmquist, R. 1984. “Estimating the Demand for the Characteristics of Housing.” *Review of Economics and Statistics* 66(3), 394-404.

Rosen, S. 1974. "Hedonic Prices and Implicit Markets - Product Differentiation in Pure Competition." *Journal of Political Economy*, 82(1), 34-55.

Sieg, H., V. K. Smith, H. S. Banzhaf, and R. Walsh. 2004. “Estimating the General Equilibrium Benefits of Large Changes in Spatially Delineated Public Goods.” *International Economic Review* 45(4), 1047-1077.

Smith, V. K., and J. Huang. 1995. “Can Markets Value Air Quality? A Meta-Analysis of Hedonic Property Value Models. *Journal of Political Economy* 103(1), 209-227.

Staiger, D. and J. Stock. 1997. “Instrumental Variables Regression with Weak Instruments.” *Econometrica*, 65(3), 557-586.

Tra, C. 2010. “A Discrete Choice Equilibrium Approach to Valuing Large Environmental Changes.” *Journal of Public Economics*, 94(2), 183-196.

United States Environmental Protection Agency. 2005. National Ambient Air Quality Standards (NAAQS). http://www.epa.gov/air/criteria.html. July.

Walker, W. R. 2011. “Environmental Regulation and Labor Reallocation: Evidence from the Clean Air Act.” *American Economic Review: Papers and Proceedings*, forthcoming.

CHAPTER 3

The Dynamics of Air Quality Capitalization and Locational Sorting in the Housing Market

Introduction

Charles Tiebout (1956) argued that households should "vote with their feet" and choose residential locations with the optimal bundle of amenities and price. Since that time, and especially after Rosen's (1974) formal development of hedonic theory, economists have exploited housing market data to uncover people's preferences and values for a wide range of spatially delineated non-market goods including school quality, crime, and air pollution. When amenity levels are constant, the compensating differential in housing prices across locations should accurately reflect the value of amenity differences, such that the marginal mover is indifferent between locations. However, our understanding of how prices and households dynamically respond to a change in amenity levels is limited. What is the evolution of housing price capitalization of the changed amenity levels? Does the housing market efficiently aggregate new information or is there a delay in capitalization? Do households spatially sort based on preferences for the new amenity levels? Is sorting necessary for capitalization?

In this paper, I address these questions in the context of large improvements in air quality that occurred during the 1990s. I utilize a high frequency panel of house prices, turnover, and occupant characteristics to determine the dynamic path of capitalization and assess what role preference based sorting plays in driving those dynamics.

An estimate of people's willingness to pay for clean air is necessary in order to evaluate a proposed policy in a benefit-cost framework.[45] Dating back to Ridker and Henning (1967) and Harrison and Rubinfeld (1978), a voluminous literature seeks to identify the housing price-air

[45] Another prominent approach for measuring the benefits of air quality is to directly examine health impacts (e.g., Chay and Greenstone 2003, Currie and Neidell 2005, and Currie and Walker 2010).

quality gradient. The bulk of this literature relies on housing data for a single year and a single MSA.[46] In addition to estimates being biased from omitted relevant variables, these estimates lack any sort of context by which the housing price-air quality gradient came to be or any indication that the market is even in equilibrium.

Chay and Greenstone (2005) significantly improve on the prior literature by utilizing panel data and instrumental variables to alleviate prior estimation problems. They use the 1970 Clean Air Act to instrument for changes in air quality between 1970 and 1980 and the corresponding county level decennial census data to measure housing price changes. A similar approach is taken by Bento, Freedman, and Lang (2010), who use the 1990 Clean Air Act Amendments (CAAA) and tract level decennial census data from 1990 and 2000 to examine heterogeneous valuation of air quality and the distributional impacts of the 1990 CAAA. While these studies have mitigated past concerns of bias, they are still unable to contextualize the housing price changes due to the ten year difference between observations. Price evolution is necessarily ignored due to data limitations, even though the path of capitalization could cause the decadal estimate to be an over- or underestimation of the capitalization rate that reflects what households are truly willing to pay for air quality improvements.

Even looking beyond air quality to other amenities households value, surprisingly little is known about the microeconomic dynamics of capitalization.[47] Cellini, Ferreira, and Rothstein (2010) examine the effect of school bond passage on house price sales. While they are not explicitly interested in the dynamics, their results show that the price premium for living in a

---

[46] Smith and Huang (1995) offer a meta-analysis of published and unpublished estimates of the housing price-air quality gradient and show that the estimates vary widely, few are statistically significant, and 25% have a perverse sign.

[47] On the more macroeconomic side, dynamics are more often and more easily studied. Case and Shiller (1989) examine quarterly price trends for repeat home sales in three MSAs and find evidence of serial correlation indicating market inefficiency. Glaeser and Gyourko (2007) support these findings with evidence from quarterly price indices from 115 MSAs and additionally show that prices demonstrate mean reversion in the median run.

district that passed a bond as much as triples from one year after passage to ten years after passage and continues to increase even after bond-related spending has ceased. Case et al. (2006) measure the price discount for home sales that arises with the revelation of groundwater contamination in a localized area. The discount reaches a maximum ten years after the revelation and then begins to abate. While both articles find interesting dynamics, neither attempts to explain the mechanism behind the results.

One potential driver of housing price dynamics is preference based sorting of households. In general, examining sorting patterns serves as a second revealed preference approach, other than price differentials, to understand valuation of location specific amenities. In the current context, there are two reasons to be interested in the locational sorting patterns of households. First, if households are re-locating in response to a change in amenities, then this serves as evidence of heterogeneous preferences and a validation of Tiebout's ideas. (However, even with homogeneous preferences, price differentials can exist.) Second, household sorting may preclude the estimated capitalization rate from a welfare interpretation. For instance, if wealthy households value air quality more and move into initially poor neighborhoods with recently improved air quality, the housing price increase no longer reflects the initial occupants' willingness-to-pay. Further, there could be an additional indirect appreciation in house prices if households like to have wealthy neighbors and move to the recently cleaned up neighborhood for that reason rather than the air quality.[48]

Due to the fundamental importance of understanding sorting behavior, the empirical literature has grown substantially in recent years, though limitations still exist. Cameron and McConnaha (2006) and Greenstone and Gallagher (2008) examine changes in the socioeconomic

[48] This process is described and confirmed by Bayer, Ferreira, and McMillan (2004) for the case of school quality.

characteristics of housing occupants surrounding Superfund sites and find little to no sorting. In contrast, Banzhaf and Walsh (2008) and Davis (2010) find strong evidence of sorting in response to toxic emissions and new power plants, respectively. All of these examples have used decennial census data to measure sorting. While the implications of results are unlikely to change if data with more frequent observations were used, the findings offer no context for how the sorting occurred. Further, if evidence of capitalization and sorting was found, there is no way to determine the relationship between the two.

One reason for the lack of focus on price evolution and household location dynamics in the housing market is the tremendous data requirement. However, the American Housing Survey (AHS), which I use in this study, is ideal for this purpose. The AHS collects information from a nationally representative panel of housing units and their occupants every two years, including self-reported home value or rent. The high frequency and regularity of observations allows inference into the dynamics of capitalization and sorting.

Additionally, the structure of the AHS obviates standard concerns when estimating a hedonic model. First, the omission of unobserved unit or location characteristics commonly biases hedonic estimates. The AHS offers multiple observations for each housing unit and thus time-invariant omitted variables do not pose a problem. When using sales data, researchers often rely on repeat sales to purge these time-invariant confounders. However, a repeat sales model can exclude as much as 97% of observations (Case and Quigley 1991). Further, transacting properties are not random; Case, Pollakowski, and Wachter (1997) show that properties that transact more tend to appreciate more, as well as have different structural characteristics. Appreciation estimates from the AHS will not have this same bias since all units report price changes, not just those that sell, and the units are randomly sampled.

Importantly, I gained access to the restricted access version of the AHS through a Census Restricted Data Center. Unlike the public use AHS, which only identifies the geographic location of a housing unit at the MSA level, the restricted access version identifies the census tract where each unit is located. This fine scale enables two critical aspects of the present research. First, the air quality that a given household faces can be measured with far greater precision. Second, the identification strategy can exploit localized air quality regulation intensity that would be masked at the MSA level.

I match housing units from the AHS to particulate matter (PM_{10}) concentrations measured from nearby air quality monitors. PM_{10} is chosen as the air quality measure of choice due to its impact on human health, its visibility, its prior use in the literature, and, importantly, the fact that it began to be regulated in 1990.

My empirical approach exploits the structure of the 1990 CAAA and the timing of air quality improvements to estimate the impact of changes in PM_{10} on housing prices, turnover frequency, and demographic characteristics. Due to the endogeneity of air quality, I construct instrumental variables (IV) to estimate the exogenous changes in PM_{10} concentrations. In order to encourage all areas of the United States to have pollution levels below thresholds set for human health and welfare, the Environmental Protection Agency (EPA) classifies all counties as in- or out-of-attainment of the thresholds and then uses a variety of tools to encourage abatement in the out-of-attainment counties. However, Auffhammer, Bento, and Lowe (2009) find that the regulation acts through the channel of individual air quality monitors rather than county level averages. Motivated by their findings and those of Bento et al. (2010), I develop an instrumental variables strategy that relies on both a monitor attainment designation and the more standard county level designation. First and foremost, these dual instruments allow for estimation of exogenous changes in PM_{10} concentrations, but they also allow for heterogeneous impacts of the CAAA, consistent with Auffhammer et al.

Within this IV framework, I examine the differential impact PM_{10} has on housing prices, turnover frequency, and demographic characteristics for various lags between observations ranging from two to ten years. PM_{10} regulation began in 1990 and the most polluted areas rapidly reduced their PM_{10} concentrations in order to comply with the standards and avoid penalties. However, the relative distribution of PM_{10} changed little after 1993. Given the timing of PM_{10} reductions, I am able to test whether air quality changes are capitalized immediately into housing prices or if late-decade housing price changes can be attributed to early-decade air quality improvements.

The results suggest that while housing prices immediately appreciate in response to air quality changes, the bulk of the capitalization is delayed by as much as eight years. The implied marginal willingness to pay (MWTP) for a one unit reduction in PM_{10} increases from $45 to $183 when comparing a lag time between observations of two years to ten years. There is no evidence that households sort based on preferences for clean air, and further sorting cannot explain the observed pattern of capitalization dynamics.

My analysis indicates that dynamics should be understood in order to correctly value a rapidly changing amenity. Specifically, in the case of the valuation of the significant air quality improvements in the 1990s, if the time frame of study was too short, then the estimates could be less than half of the true valuation. In addition, the lag in capitalization can cause cross sectional estimates of marginal willingness to pay to be biased as changes in the cumulative distribution function (cdf) of air quality are not associated with in-kind changes in the cdf of housing prices.

Further, given that past changes in PM_{10} can predict future changes in housing prices, the results add microeconomic evidence of inefficiency in the housing market.

Particulate matter and the 1990 Clean Air Act Amendments

Particulate matter is a class of solid and liquid air pollutants that consists of nitrates, organic chemicals, metals, soot, smoke, and dust. Particulate matter enters the atmosphere either directly from a source, such as construction sites, unpaved roads, fields, smokestacks or fires, or indirectly as the result of reactions from sulfur dioxides and nitrogen oxides that are emitted from power plants, industrial facilities, or motor vehicles (Unites States Environmental Protection Agency [EPA] 2010). In general, the contribution of indirect sources is substantially larger to overall particulate matter concentrations than the contribution of direct sources.

Particulate matter is classified by the measurement of its diameter, with the diameter being inversely related to the potential for human health damage. PM_{10}, the pollutant of interest in this paper, is particulate matter that is less than 10 micrometers in diameter. At this diameter, particulate matter can penetrate deeply into the human respiratory system and cause numerous health problems, including aggravated asthma, chronic bronchitis, and even premature death for those with pre-existing lung and heart problems (EPA 2010).[49] Fine particulates, with a diameter of 2.5 or less, are most harmful.

Responding to calls to action and mounting scientific evidence, the United States Congress passed the 1970 Clean Air Act (CAA), which was the first federal legislation

[49] For a concise analysis of the health effects from exposure to PM10, see Hall et al (1992), Dominici et al (2002), and Daniels et al (2000). In addition to human health effects, particulate matter can damage crops and buildings and reduce visibility.

establishing air quality control.[50] The 1970 CAA created the Environmental Protection Agency (EPA) and authorized it to enforce National Ambient Air Quality Standards (NAAQS) for six common air pollutants, the so-called criteria pollutants. Particulate matter was included in this group in the form of total suspended particulates, or TSPs, which is particulate matter of diameter 100 micrometers or less. The 1990 Clean Air Act Amendments (CAAA), the second major update of the CAA, replaced TSPs with PM_{10} to parallel current scientific understanding of pollution's effects.[51] In 1997, the EPA further refined the NAAQS to target $PM_{2.5}$, again reflecting current understanding.

The objective of the NAAQS was to lower concentrations of the criteria pollutants below harmful levels everywhere in the United States. For PM_{10}, the EPA set an annual arithmetic mean daily readings concentration threshold of 50 $\mu g/m^3$ and a 24-hour arithmetic mean concentration threshold of 150 $\mu g/m^3$.[52] In order to achieve the NAAQS, the EPA held counties and states accountable for meeting those standards. If even a *single monitor* within a county exceeds the annual threshold or the 24-hour threshold for more than one day, then the *entire county* is considered in violation of the standard. The EPA can then move to designate that county as out of attainment, which then requires the county and state, in cooperation with the EPA, to develop an official plan to reduce pollution and attain the standards set forth by the NAAQS. As a means to encourage compliance, non-attainment counties can be subject to scrutiny over industrial activities, including the opening of new plants, and can even have federal highway funds withheld. The effectiveness of these regulations has been well documented with

[50] Prior federal legislation in the 1950s and 1960s merely provided funds for monitoring air quality and for research on the impacts of pollution on health and agriculture.
[51] In addition, the 1990 CAAA expanded the scope of federal regulation by adding control over the release of 189 toxic chemicals and by initiating the Acid Rain Program.
[52] The EPA sets primary and secondary standards for all criteria pollutants, where primary standards address human health, especially of vulnerable populations, and secondary standards address overall human welfare. For PM_{10}, the primary and secondary standards are identical.

respect to overall concentrations (e.g., Chay and Greenstone 2005, Auffhammer et al. 2009), industrial activity (e.g., Nadeau 1997, Henderson 1996), and health outcomes (Currie and Neidell 2005, Chay and Greenstone 2003).

Data

This section discusses the source and relevant features of the air quality data, the regulatory data, the housing data, neighborhood data, and county employment data as well as sample restrictions. Table 3.1 provides a complete listing of all variables used in the analysis.

Individual air quality monitor records were obtained from the Air Quality Standards (AQS) database, which is maintained by the Environmental Protection Agency (EPA).[53] Monitors are placed throughout the United States, but are primarily located in urban areas.[54] For each monitor, the database includes the annual mean PM_{10} concentration, the number of days the PM_{10} concentration was above the 24-hour threshold, the geospatial coordinates of the monitor, and several reliability measures.

For the purposes of my analysis, I restrict monitors to those that are sufficiently reliable and have readings in necessary years. Title 40 Part 58.12 and Title 40 Part 50 Appendix K of the Code of Federal Regulations (CFR) prescribe the monitoring frequencies for PM_{10} monitors, as well as the criteria for establishing whether a monitor is "representative" and therefore should be included in the analysis.[55]

---

[53] The data are available online at www.epa.gov/aqspubl1/
[54] The exact spatial distribution of included monitors is withheld to minimize disclosure risk.
[55] In the AQS data, a criteria flag is set based on data completeness criteria so that if it is set to "Y", then the assumption can be made that the data represent the sampling period of the year. These summary criteria are based on 75% or greater data capture and data reported for all four calendar quarters in each year. Additionally, I exclude monitor-year observations that are affected by "extreme natural events" beyond human influence.

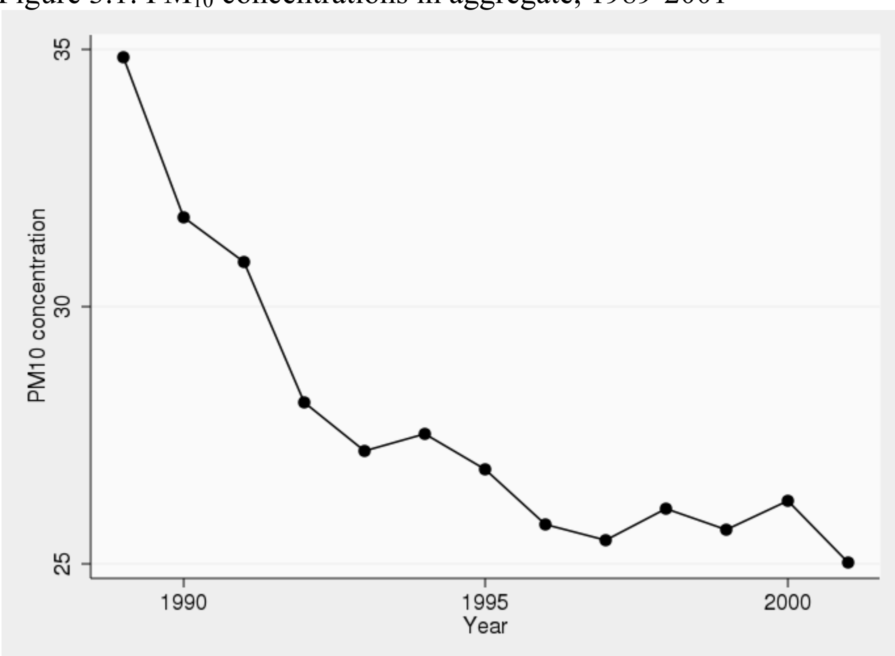
Table 3.1: List of variables used in regressions

Unit characteristics	
	Dishwasher (1=yes)
	Washing machine (1=yes)
	Dryer (1=yes)
	Stove (1=yes)
	Cracks present on exterior or interior (1=yes)
	Holes present on exterior or interior (1=yes)
	New roof (1=yes)*
	Kitchen remodeled (1=yes) *
	New or remodeled bathroom (1=yes) *
	Other addition (1=yes) *
Neighborhood (census tract) characteristics	
	Population density
	Share Black
	Share Hispanic
	Share over 60
	Share under 5
	Share foreign born
	Share high school graduate
	Share college graduate
	Share unemployed
	Share below poverty line
	Share receiving welfare benefits
	Share living in the same residence for 5 years
	Median family income
	Median rent+
	Median house value+
	Total housing units
	Share of units occupied
	Share of occupied units that are owner occupied
County Employment (CBP)	
	Number of jobs in construction
	Number of jobs in manufacturing
	Number of jobs in mining
	Number of jobs in agriculture or forestry
	Number of jobs in all other sectors

Notes: * Only included in owner-occupied specifications. + Both Median rent and Median house value are included in specifications that contain both owners and renters. Segregated specifications only contain the corresponding median.

For the key measure of air quality, I use the weighted annual mean PM_{10} concentration. Figure 3.1 shows the time series trend of annual PM_{10} concentrations for the years 1989-2001; in aggregate, concentrations declined by 28%, which is consistent with the findings of Auffhammer et al. (2009). For the purposes of the current analysis, it's important to note that 75% of the reductions occurred by 1993.

Figure 3.1: PM_{10} concentrations in aggregate, 1989-2001



Notes: Sample includes air quality monitors that are included in the main analysis and have at least 10 reliable observations over the timeframe 1989-1991.

As outlined in Section 3, the EPA classifies all counties in the United States as in attainment or non-attainment of the standards set forth by the 1990 CAAA. I obtained the county attainment designations from the annual CFR.

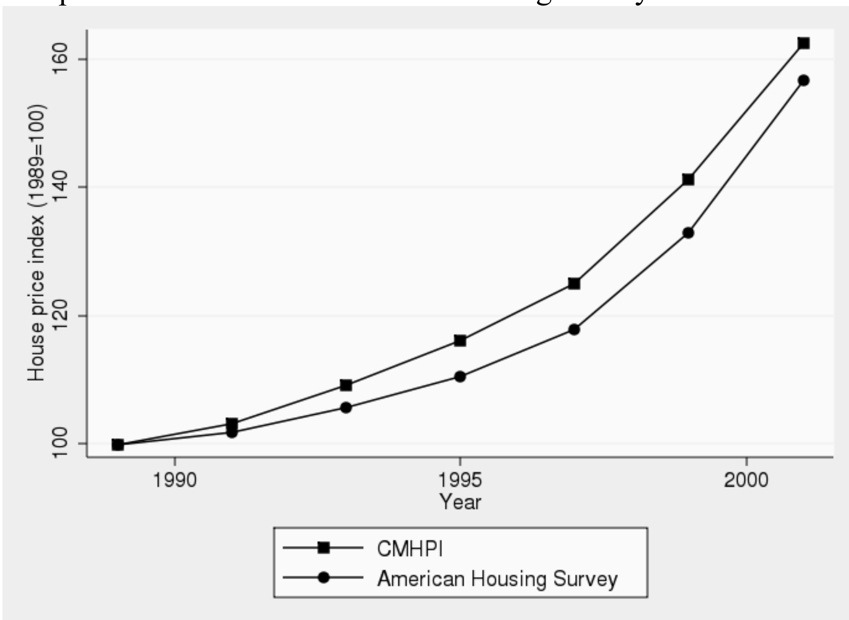
In addition, I construct an attainment status for each monitor directly from the AQS data using the same threshold rules as the EPA's county designation. If in year *t* a monitor's annual PM_{10} concentration is greater than 50 $\mu g/m^3$ or its 24 hour concentration exceeds 150 $\mu g/m^3$ for two days or more, then that monitor is designated non-attainment in year *t+1*. This logic follows closely Chay and Greenstone (2005), who assigned county attainment status due to EPA records being unavailable.

Housing data for the years 1989-2001 were obtained from the restricted access American Housing Survey National Sample (AHS).[56] The AHS is a panel of housing units that are surveyed every two years, usually between August and November. Importantly, the AHS follows units, not occupants. Units generally enter and exit the sample based on regional population growth and, in the cross section, the sample is representative of the United States' housing stock.

The AHS collects information about house value (if owner occupied), rent (if renter occupied), dwelling characteristics (e.g., number of bedrooms, number of bathrooms), occupant characteristics (e.g., race, education, income), and when the current occupants moved in. The price data suffer from input and measurement error owing to the fact that they are self-reported (though this should not be an issue with renter occupied units). In order to assess the significance of this issue, I construct an index of owner occupied housing values from the AHS (1989=100) and compare it to Freddie Mac's Conventional Mortgage Home Price Index (CMHPI) identically scaled. Figure 3.2 graphs the comparison and suggests that the appreciation trends are similar. Thus, at least in aggregate, the self-reported nature of the key price data does not prevent the results from being generalized to the housing market. For all subsequent analyses, prices are adjusted for inflation to 2001 levels using the total shelter component of the CPI.

[56] There is also a Metropolitan Sample of the AHS that surveys housing units in select MSAs on a rotating basis. This sample is not included in the present analysis because the time interval between observations is generally six or more years.

Figure 3.2: Comparing the Conventional Mortgage Home Price Index to an index of owner-occupied units from the American Housing Survey



Notes: Owner-occupied housing units must have at least two observations to be included in the sample used to construct the index.

Unlike the public use version, the restricted access AHS records the census tract where each unit is located. Using GIS, I determine the distance between each tract's geographic center, or centroid, and surrounding air quality monitors. I am then able to match housing units to air quality monitors on the basis of least distance, while excluding all unit-monitor matches that are greater than five miles apart. As units get further away from monitors, more measurement error is introduced into the key air quality variable. Additionally, Bento et al. (2010) find that valuation declines as distance from monitors increases, especially beyond five miles. Thus, a five mile cutoff is used to balance measurement error and valuation concerns with sample size concerns. Robustness checks are run with different cutoffs. Additional details about the matching process will follow in Section 6, which describes the specifics of the empirical approach.

The first two columns of Table 3.2 provide summary statistics for the included versus excluded housing units. The two samples are significantly different, primarily reflecting the propensity of air quality monitors to be placed in urban areas. Included units are worth less, rent for less, are more often renter occupied, are smaller, and are less likely to have appliances. The occupants of included units are more likely to be elderly, more likely to be Black or Hispanic, and less likely to be a high school graduate.

Without question, the socioeconomic characteristics of neighbors are an important piece of a housing unit's value. While the AHS offers many benefits, the observations are nowhere near spatially dense enough to measure important neighborhood variables. To alleviate this restriction, I use the census tract identifier in the AHS to include tract level decennial census data from GeoLytics Neighborhood Change Database.[57] Of course, these data are only available in 1990 and 2000. I assume a linear trend to impute values for all years 1989-2001.

As discussed above, when studying how households value air quality, one issue that can confound the analysis is the correlation between air pollution and economic activity. While households value clean air, they also value jobs (and probably more so). In an attempt to mitigate the bias that arises from this correlation, I include employment data from the County Business Patterns (CBP) database maintained by the US Census. CBP gives yearly employment counts broken down by Standard Industrial Classification/North American Industry Classification System codes for each county.[58] I aggregate the number of jobs into five major categories that are intended to be most relevant to air quality: construction, manufacturing, mining, agriculture/forestry, and a catch all for the remaining codes.

[57] More information is available at www.geolytics.com/

[58] The classification switched from SIC to NAICS in 1997. The broad classifications I use are unaffected by the change.

Table 3.2: Summary statistics for housing units and occupant characteristics

Variable	out of sample (1)	in sample (2)	p value 1 v 2 (3)	county in attainment (4)	county non-attainment (5)	monitor non-attainment (6)	p value 4 v 5 (7)	p value 4 v 6 (8)	p value 5 v 6 (9)
% of units owner occupied	64.9	56.5	0.00	56.7	60.2	49.4	0.06	0.00	0.00
House value	159,919	138,557	0.00	145,531	128,621	118,318	0.00	0.00	0.09
Annual Rent	7,459	6,872	0.00	6,753	6,947	7,278	0.26	0.05	0.25
House value appreciation (percent), 1987-1989	3.4	1.1	0.00	1.3	2.9	-3.0	0.15	0.00	0.00
Annual rent increase (percent), 1987-1989	-1.1	0.3	0.00	-0.1	2.3	-0.4	0.02	0.83	0.06
% of owner occupied units that turned over, 1987-1989	9.9	9.4	0.41	8.6	10.7	11.3	0.16	0.18	0.81
% of renter occupied units that turned over, 1987-1989	49.3	44.6	0.00	44.4	49.3	39.3	0.10	0.11	0.01
Number of bedrooms	2.6	2.4	0.00	2.5	2.4	2.3	0.17	0.00	0.02
Number of bathrooms	1.6	1.4	0.00	1.4	1.4	1.4	0.87	0.48	0.64
Square feet	1,842	1,759	0.02	1,830	1,726	1,479	0.09	0.00	0.00
Number of kitchens	0.998	0.990	0.00	0.993	0.994	0.971	0.94	0.01	0.02
New roof in last two years	8.5	7.5	0.03	8.2	5.7	6.9	0.01	0.27	0.34
New kitchen in last two years	4.3	4.4	0.95	4.0	5.9	3.5	0.03	0.56	0.03
New bathroom in last two years	5.8	5.5	0.44	5.4	6.1	4.7	0.44	0.46	0.23
Other addition in last two years	1.2	0.7	0.00	-	-	-	0.69	0.18	0.18
Stove (1=yes)	98.3	98.4	0.53	98.5	98.2	98.3	0.48	0.72	0.84
Dishwasher (1=yes)	54.8	43.4	0.00	43.8	42.5	42.5	0.49	0.55	0.99
Washing machine (1=yes)	76.6	68.2	0.00	69.6	67.8	62.2	0.30	0.00	0.03
Dryer (1=yes)	71.2	61.0	0.00	61.6	62.2	56.2	0.74	0.02	0.02
Central AC (1=yes)	43.8	35.1	0.00	37.1	31.1	31.7	0.00	0.01	0.82
Household income	40,308	35,197	0.01	33,650	33,565	45,624	0.97	0.06	0.06
% of units with kids under 6	17.7	17.2	0.41	17.1	17.1	17.7	0.98	0.73	0.75
% of units with household head over 60	26.0	28.3	0.00	27.8	31.0	26.1	0.07	0.39	0.04
% of units with Black or Hispanic household head	14.8	20.0	0.00	21.1	16.8	19.2	0.00	0.29	0.22
% of units with household head a high school graduate	57.2	54.5	0.00	54.6	56.7	50.4	0.26	0.06	0.02
% of units with household head a university graduate	29.1	29.8	0.34	28.8	29.4	35.6	0.72	0.00	0.01

Notes: All values are for the year 1989 unless specified otherwise. All monetary variables are in $2001. Units are included in sample statistics if they are included in the main IV analysis and have valid price observations for 1987 and 1989. "-" signifies that the mean was withheld due to disclosure risk. Sample sizes for each of the columns are 19,271 (1), 4,429 (2), 2,905 (4), 931 (5), and 593 (6).

I omit units that record two-year price changes greater than 100% or less than -50%. Units that switch from owner occupied to renter occupied and back (or vice versa) are excluded. Unit-year observations are excluded when there is no survey respondent, prices are interpolated/"hot decked", or no reliable PM_{10} monitor has a reliable reading for that year.

The state of California is excluded from the analysis. California experienced a severe recession in the early 1990s that coincided with dramatic improvements in air quality. Due to the national scale of the present analysis and California's large presence in the treated group, the inclusion of California attenuates the capitalization estimates due to macroeconomic factors affecting only that state.

The 1990 Clean Air Act Amendments as a Quasi-Experiment

Empirical estimates of the true relationship between air quality and housing prices are obscured by a suite of unobserved factors that simultaneously influence both air quality and housing prices. For example, a recently built highway may worsen local air quality but increase local house prices, or the shutdown of a manufacturing plant may improve local air quality but decrease local house prices. Such correlations, which we generally expect to bias the estimated valuation of air quality toward zero, are at the root of the endogeneity problem that calls for an instrumental variable strategy. My approach treats the 1990 CAAA as a quasi-experiment to address the unobserved factors that affect both housing prices and pollution.

My identification strategy stems from that of Chay and Greenstone (2005) and follows closely that of Bento et al. (2010). Chay and Greenstone exploit the structure of the 1970 CAA to instrument for changes in TSPs at the county level between 1970 and 1980 using county nonattainment status in the mid-1970s. They demonstrate that the attainment designations in

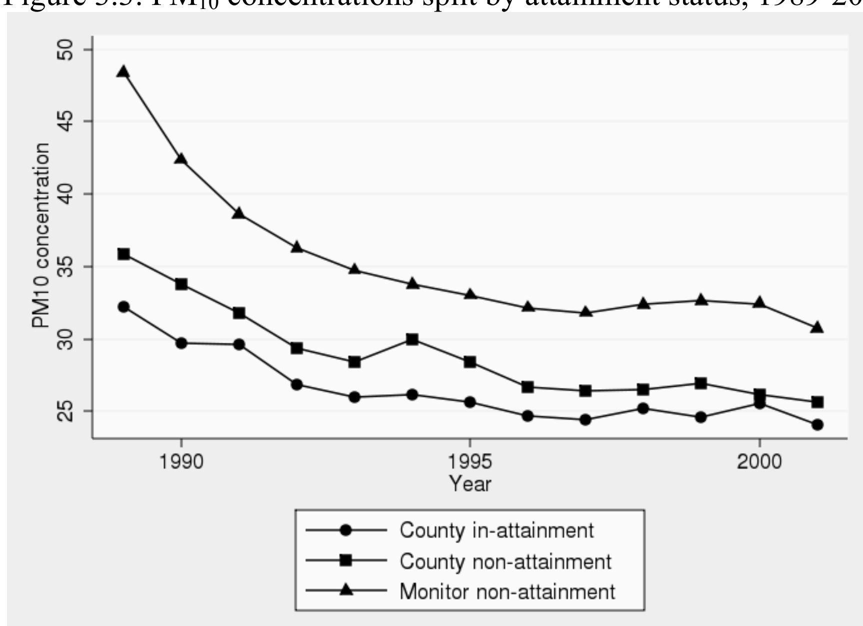
1975 and 1976 are strongly correlated with decadal changes in TSPs and housing prices, but not other county characteristics that may affect home prices (e.g., average income, population), and thus their instrument satisfies the exclusion restriction.

Bento et al. again utilize federal regulation, this time the 1990 CAAA, to instrument for changes in PM_{10} between 1990 and 2000. In contrast to Chay and Greenstone, they expand the window of mid-decade attainment designations to 1992-97 and, instead of county averages, use individual air quality monitor readings. Both changes are made to better capture the impact of the CAAA. Due to the fact that an entire county is designated non-attainment when just one monitor exceeds either threshold, optimizing local officials will exclusively target clean-up efforts in the areas around non-attainment monitors. Bento et al. construct a monitor level instrument from de facto monitor attainment designations and combine it with the more standard county designation instrument. Their spatially disaggregated model better identifies changes in PM_{10} and the resulting beneficiaries than the aggregated county-level model.

Consistent with Bento et al., I use both a monitor attainment designation and county attainment designation to identify exogenous changes in air quality at the monitor level. I construct mutually exclusive instruments based on the pair of attainment designations between observations. If a monitor was ever designated non-attainment between housing observations, then the monitor non-attainment instrument equals one and is zero otherwise. If a monitor is never designated non-attainment, but is contained within a county that is designated non-attainment at some point between housing observations, then the county non-attainment instrument equals one and is zero otherwise. This construction is clarified with respect to the various panels of differing intervals in the following section where I discuss the empirical framework.

For my instruments to be valid, they must be correlated with PM10 changes and only affect housing prices through their impact on air quality. First, I examine the time series of PM_{10} with respect to attainment designation. Figure 3.3 decomposes the aggregate PM_{10} trends shown in Figure 3.1 by three groups corresponding to the instruments: 1) monitors that are ever designated non-attainment, 2) monitors never designated non-attainment but located within a county that is designated non-attainment, and 3) monitors never designated non-attainment nor located in a non-attainment county. Figure 3.3 shows that, broadly, PM_{10} levels converged over the 1990s, which is exactly the intent of the NAAQS.[59] The largest reductions in PM_{10} clearly occurred for non-attainment monitors, which declined by a total of 17 $\mu g/m^3$, 7 more than county

Figure 3.3: PM_{10} concentrations split by attainment status, 1989-2001



Notes: The monitors from Figure 3.1 are split into attainment categories by the following rule. If a monitor ever exceeds the NAAQS PM10 thresholds during 1989-2001, then that monitor is put in the 'Monitor non-attainment' group. If a monitor never exceeds the thresholds, but is located within a county designated non-attainment by the EPA at some point during 1989-2001, then that monitor is put in the 'County non-attainment' group. All other monitors are put in the 'County in-attainment' group.

[59] For counties and monitors out of attainment, average PM_{10} concentrations are often below 50 $\mu g/m^3$, which is the non-attainment threshold for annual concentration. Frequently, a monitor or county will trigger the non-attainment designation by exceeding the 24 hour standard, even though average concentrations do not exceed annual threshold.

non-attainment monitors and 8.7 more than county in-attainment monitors. While county non-attainment monitors have a similar trend to county in-attainment monitors, they do experience an additional 1.7 $\mu g/m^3$ decline in PM_{10}. Figure 3.3 clearly shows that the instruments are correlated with PM_{10} reductions. Further evidence is given in the results sections with the first stage of the IV model. Also, like Figure 3.1, Figure 3.3 demonstrates that the majority or air quality improvements occurred early in the decade; 80% of total reductions observed for monitor non-attainment and county non-attainment monitors had occurred by 1993. Thus, attainment status appears to have a strong effect on air quality, consistent with Chay and Greenstone (2005), Auffhammer et al. (2009), and Bento et al. (2010).

Next consider the relationship between attainment designation and housing unit and occupant characteristics. In a process fully described in Section 5, housing units are matched to air quality monitors. For the purposes of the Table 3.2 and Figures 3.4a-3.5b, I classify units into the same three categories (county in-attainment, county non-attainment, and monitor non-attainment) based on monitor(s) with which the unit is matched.[60]

In an effort to understand the pre-treatment conditions of the housing units, Columns 4-6 of Table 3.2 compare housing and occupant characteristics in 1989 for the three classifications. Units in monitor non-attainment areas are worth less, are smaller, have fewer appliances, and appreciated less since 1987. The differences in housing values are suggestive of a compensating differential for air quality differences, but could also be mostly or entirely due to differences in housing stock across different areas with different air quality levels. The rental rates exhibit the

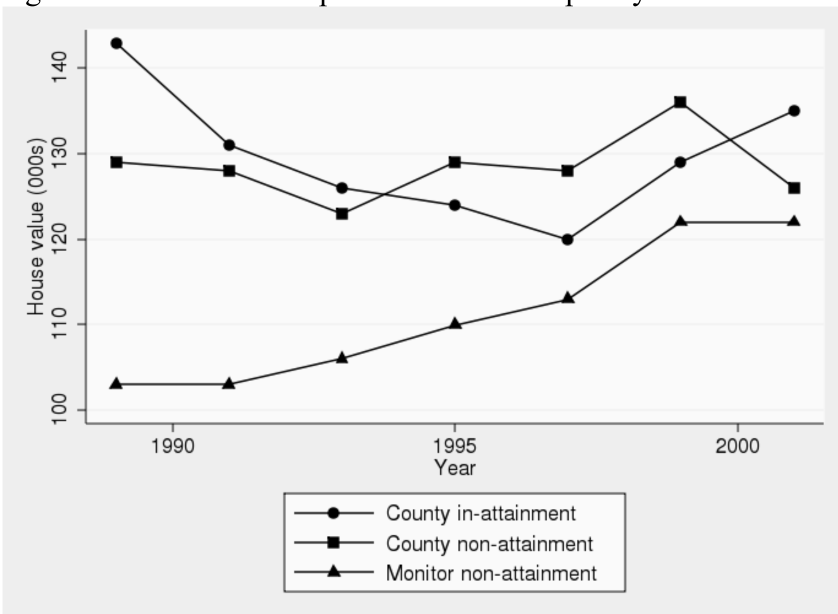
[60] In the main panel analysis, units can be matched to different monitors for different time periods. If at any point in time, a unit is matched to a monitor that is designated non-attainment, then that unit (for all time periods) is put into the monitor non-attainment bin. If a unit is within a county designated non-attainment, but is never matched to a non-attainment monitor, then that unit (for all time periods) is put into the county non-attainment bin. The remaining units are put into the county in-attainment bin.

opposite trend and underscore the problems with cross sectional estimation. Most demographic variables are not significantly different between the groups. While the groups are not perfectly balanced pre-treatment, the statistics offer no reason to be concerned about confounding effects of the identification strategy.

Next, Figures 3.4a and 3.4b examine the time series trends 1989-2001 of housing values and rents, respectively, of units in the three classifications. Both figures demonstrate a similar pattern: units near non-attainment monitors are initially worth less, but appreciate at a greater rate than the other units over the decade. Housing units near county non-attainment monitors also appreciate at a faster rate than county in-attainment units for the years 1989-1999, but then steeply depreciate 1999-2001. These patterns suggest that attainment designation is strongly correlated with changes in housing values, but, for reasons previously discussed, the correlation likely exists only because attainment status affects changes in air quality.

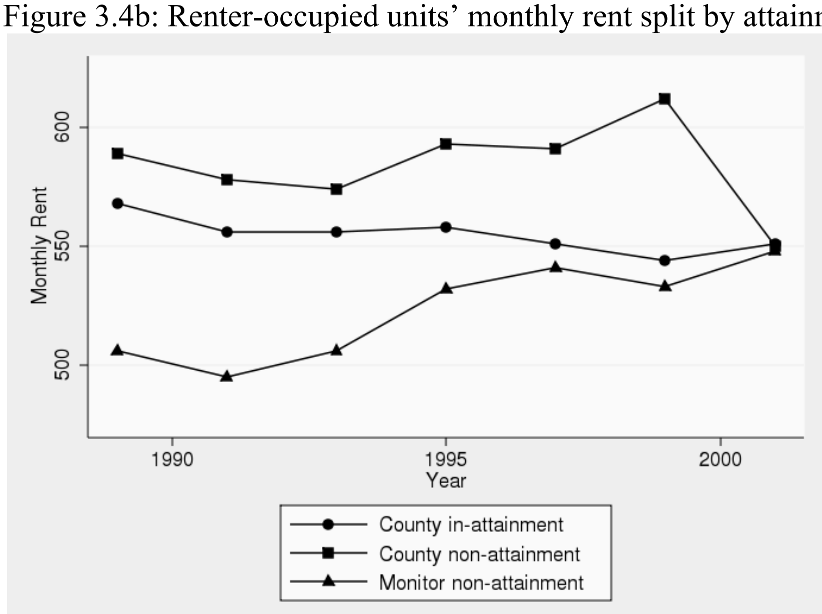
The previous figures strongly suggest that the proposed IV strategy will effectively identify the valuation of air quality in the housing market, but the focus of this paper goes beyond that effect to examine the evolution of valuation. From the data underlying Figures 3.3, 4a, and 4b, I construct time series of PM_{10} concentrations and housing prices in monitor non-attainment areas compared to county in-attainment areas. This ratio represents the deviation of the dirtiest areas from the "normal" (i.e., control or untreated) areas. Figures 3.5a and 3.5b illustrate these ratios for housing values and rental rates, respectively, and plot them against the ratio for PM_{10} concentrations. The PM_{10} ratio declines rapidly between 1989 and 1991 and then is basically flat until 2001. If air quality was immediately capitalized into housing prices, the time series of housing price ratios would be a mirror image of the PM_{10} time series, large

Figure 3.4a: Owner-occupied units' values split by attainment status, 1989-2001



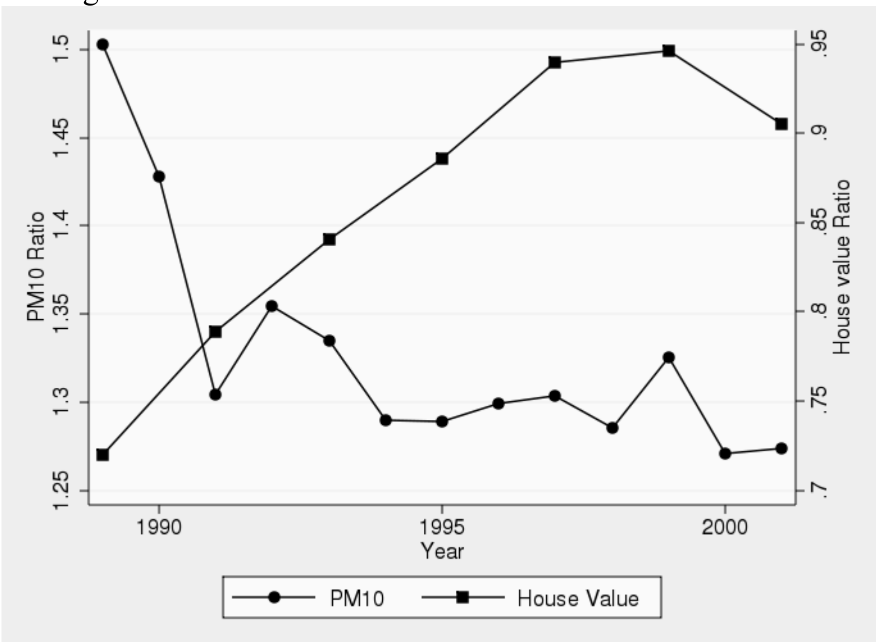
Notes: Owner-occupied housing units are split into attainment categories by the following rule. If a unit is ever matched to a monitor that exceeds the NAAQS PM_{10} thresholds while matched to that unit, then that unit is put in the 'Monitor non-attainment' group. If a unit is never matched to a monitor that exceeds the thresholds, but is matched to a monitor located within a county designated non-attainment by the EPA while matched to that unit, then that unit is put in the 'County non-attainment' group. All other units are put in the 'County in-attainment' group. In order to be included, a unit must have at least five observations during 1989-2001.

Figure 3.4b: Renter-occupied units' monthly rent split by attainment status, 1989-2001



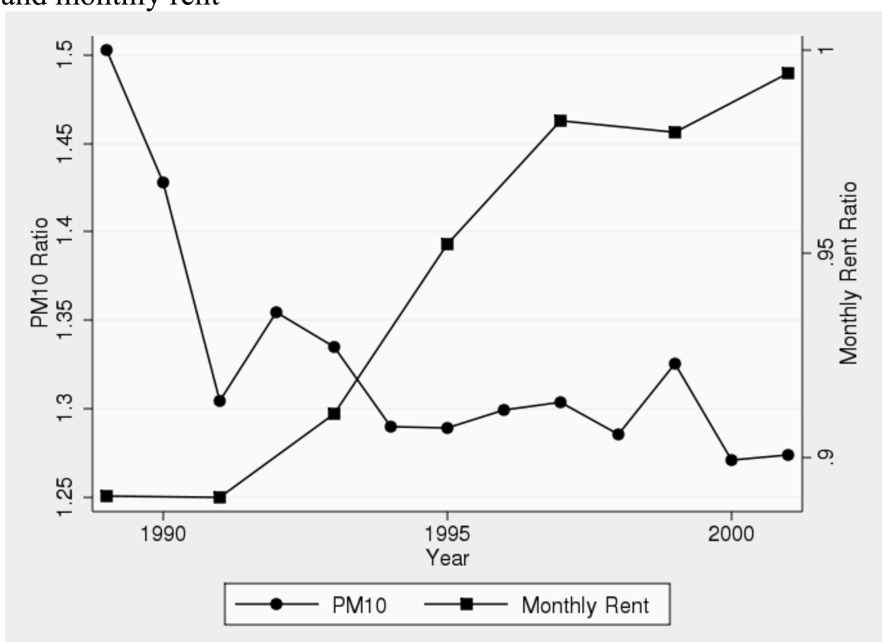
Notes: See notes to Figure 3.4a.

Figure 3.5a: Comparing monitor non-attainment areas to county in-attainment areas for PM_{10} and housing values



Notes: See notes to Figures 3.3 and 3.4a.

Figure 3.5b: Comparing monitor non-attainment areas to county in-attainment areas for PM_{10} and monthly rent



Notes: See notes to Figures 3.3 and 3.4b.

appreciation gains between 1989 and 1991 and then flat until 2001. In contrast, however, both sets of housing prices rise steadily for the better part of the timeframe suggesting that capitalization is not immediate, but instead that prices continue to capitalize changes in air quality that happened several years in the past. While these time series do not control for covariates, they are strongly suggestive that capitalization is delayed and provide the motivation for the analysis and the intuition for the results.

My IV approach relies on the assumption that, conditional on other observable housing, neighborhood, and employment characteristics, nonattainment status only affects house prices through its impact on local pollution levels. One concern with this assumption is that the CAAA regulation affects the local economy, and thus directly affects home prices. To address this concern, I analyze the effect of the EPA county level attainment designation and individual monitor exceedences on annual measures of county average income and population from the Bureau of Economic Analysis, as well as county employment from CBP.[61] The results show that both non-attainment measures have an insignificant effect on changes in income, population, and employment.[62] [63]

Another concern over the instruments is that households may spatially sort in response to the regulation and thus bias the valuation estimate. In Section 6, I examine sorting on observable

---

[61] There is not a perfect correspondence between the EPA county designation and individual monitor exceedences. First, a monitor needs to exceed the threshold for three years in order to for the county to be designated non-attainment. Second, the findings of Greenstone (2004) and Auffhammer et al (2009) suggest that county designation may be somewhat unresponsive to changes in air quality. However, both may theoretically spur a county to action and affect the local economy, and thus, I examine the effects of both.

[62] I use the universe of county data for years 1985-2003 and include county time trends in the regression of attainment status on changes in income, population, and employment (all regressed separately). For the models using individual monitor exceedences, I include only monitor-year observations that are reliable and exclude all counties without monitors. While no counties were designated non-attainment prior to 1990, including data from 1985-1989 allows for a pre-regulation trend to be estimated.

[63] At first blush, this finding is contrary to the results of Becker and Henderson (2000) and Hanna (2010), who each demonstrate that economic activity is displaced as a result of the Clean Air Acts. However, the results are not inconsistent because both articles focus on the economic decisions of affected polluting firms only, which represent a relatively small portion of the total economic activity of the county.

household characteristics in an identical IV framework as the main valuation analysis and find that changes in demographic composition are not related to the instruments. While it is impossible to completely validate the exclusion restriction, all evidence suggests that attainment designations only affect housing prices via the air quality channel.

Empirical Framework

In this section, I outline the econometric approach I use to estimate the implicit value of air quality derived from housing market data. I rely on the hedonic framework and estimate IV first difference models with lags between observations ranging from two to ten years in order to examine the extent to which air quality improvements are capitalized into housing prices and how the rate of capitalization may change over time. Further, this section details how housing units are matched with air quality monitors and the exact construction of the instruments.

The standard cross-sectional hedonic specification of the relationship between housing prices and PM_{10} concentrations is:

$$p_i = \theta PM_i + \mathbf{X}_i'\boldsymbol{\beta} + \varepsilon_i \quad , \tag{3.1}$$

where p_i is the natural log of price of housing unit i, PM_i is the concentration of PM_{10} for housing unit i, and $\mathbf{X}_i$ is a vector of unit and location covariates.

In principle, the coefficient θ measures the gradient of housing prices with respect to air quality at a given moment in time. However, there are two reasons why estimates of θ from a cross section might be biased. The first reason, which is well recognized in the literature, is that a cross sectional specification may suffer from omitted variable bias. The variables in $\boldsymbol{X}$ may fail to control for characteristics that affect local housing prices and are also correlated with PM_{10} concentrations. The second reason, which I am introducing in this paper, is that price changes may lag behind improvements in air quality, which would cause cross sectional estimates of θ to be biased downwards.[64] The intuition of this bias is as follows. From Equation (3.1), we get

$\theta = \frac{\partial p_i}{\partial PM_i}$, which we expect to be negative. If the distribution of PM is compressed, but the distribution of prices remains the same, then θ would increase in magnitude.

To mitigate the bias resulting from omitted variables and explore the lag in capitalization, I exploit the panel nature of the AHS and estimate a series of first-difference models. Incorporating multiple observations from each housing unit controls for both observable and unobservable time invariant characteristics of areas that might be correlated with house prices and air quality, such as climate and topographical features, transportation infrastructure, and population density.

In order to examine the lag in capitalization, I construct first-difference datasets with varying time intervals between observations. Since the AHS surveys housing units every other year, I create datasets with two, four, six, eight, and ten years between observations. For example, if a housing unit was surveyed in 1989, 1991, 1993, and 1995, this unit would enter the two-year difference panel with years 1989-1991, 1991-1993, and 1993-1995, the four-year difference panel with years 1989-1993, and the six-year panel with years 1989-1995. The differenced intervals are non-overlapping and priority is given to the earlier of the intervals. In the example just given, only one of 1989-1993 and 1991-1995 can be included in the four-year panel due to the overlap and 1989-1993 is chosen because it occurs earlier in the decade.

In a similar manner as done with the units, I construct monitor-interval pairs that are then matched with the unit-interval pairs. I exclude all matches greater than five miles in distance and

[64] Valuation is biased upwards when air quality improves, but would be biased towards zero if air quality were to get worse.

include only the match of least distance in the event of more than one monitor-interval matching to a unit-interval. Differencing observations allows me to ensure that units are matched to the same air quality monitors at the beginning and end of an interval, while still allowing the monitor-unit match to change across time periods. This strategy balances the competing goals of minimizing measurement error and maximizing sample size.

The first difference model is

$$\Delta_t p_i = \theta(\Delta_t PM_i) + \Delta_t \mathbf{X}_i' \boldsymbol{\beta} + \Delta_t \varepsilon_i \quad , \tag{3.2}$$

where PM_i, $\mathbf{X_i}$, and p_i are differenced by t = 2, 4, 6, 8, or 10 years depending on the model.

However, as discussed in Section 4, additional bias is likely to arise due to time-varying unobserved factors correlated with both air quality and housing prices. To address this problem, I utilize the structure of the 1990 CAAAs to construct instruments for changes in PM_{10}. The first and second stage equations of the first difference IV analysis are

$$\Delta_t PM_i = \mathbf{N}_i' \boldsymbol{\gamma} + \Delta_t \mathbf{X}_i' \boldsymbol{\beta}_1 + \Delta_t \nu_i \tag{3.3}$$

and

$$\Delta_t p_i = \theta(\Delta_t \hat{P}M_i) + \Delta_t \mathbf{X}_i' \boldsymbol{\beta}_2 + \Delta_t \varepsilon_i \quad , \tag{3.4}$$

where $\mathbf{N_i}$ is a vector of two binary, mutually exclusive instruments, one based on monitor designation and one based on county designation. If in Equations (3.3) and (3.4), the first difference is taken between years t_1 and t_2, then the monitor instrument equals one if the monitor is designated out of attainment for some year in the interval [t_1+1, t_2]. The county instrument is equal to one if the county is designated out of attainment for some year in the interval [t_1+1, t_2] and the monitor instrument equals zero. In addition to the overidentified model that captures heterogeneous regulation, I also estimate and present results for a just identified model that uses only the monitor non-attainment instrument.

The main IV first difference specification I estimate combines owner occupied units and renter occupied units in order to boost sample sizes. Since the model is log-linear, the enormous difference between annual rent and home values is absent because the focus is on percent change. In an additional set of regressions, I estimate the two samples separately.

To explore preference based sorting in response to exogenous changes in air quality, I estimate a variant of the IV first difference model above with a different dependent variable:

$$y_i = \theta(\Delta_t \hat{P}M_i) + \Delta_t \mathbf{X}_i' \boldsymbol{\beta}_2 + \Delta_t \varepsilon_i \qquad (3.5).$$

In this framework, I estimate models of sorting on observables for binary classifications of household head over the age of 60, the presence of children under 6, the presence of children under 18, household head is either Black or Hispanic, and educational attainment of the household head (high school dropout, high school graduate, college graduate). The dependent variable y_i is binary and equals one if the unit has turned over and the demographic characteristic of choice has increased (i.e., the household head of the prior occupants was under 60 and the household head of the new occupants is over 60) and equals zero if the unit has turned over and the demographic characteristic of choice has decreased. All units that either do not turn over or do not have a change in the given demographic group are excluded.

I additionally test for changes in the frequency of turnover as a proxy for sorting on unobservables. Still using the IV first difference framework, I define y_i in Equation (3.5) to first be binary and equal to one if unit i has turned over at some point during the interval and zero otherwise. In a second specification, I define y_i to be the number of moves that occurred to unit i during the interval.

Results

Table 3.3 presents least squares estimates of Equation (3.2), for interval lengths of two, four, six, eight, and ten years. Since the estimated model is log-linear, the coefficient on PM_{10} is interpreted as the percent change in housing prices in response to a 1 unit change in PM_{10}. A negative coefficient implies that a decrease in PM_{10} is associated with an increase in housing prices, which is the predicted relationship. The coefficients are negative only for intervals 4, 8, and 10, and none are significant. Further, there is no pattern present in the coefficients that would suggest that dynamics are important. The ten-year capitalization rate is -0.08%, which indicates an elasticity -0.03. Converting owner-occupied house value to yearly expenditure using a simplified 7.5% factor (as in Bajari and Kahn 2005), I can back out the implicit MWTP for a one unit reduction of PM_{10} for renter and owners combined. For the ten year lag, MWTP is $7.

Table 3.3: Least Squares Results

	Interval				
	2	4	6	8	10
PM_{10} (1/100)	0.05	-0.04	0.06	-0.03	-0.08
	(0.04)	(0.09)	(0.11)	(0.14)	(0.17)
R^2	0.02	0.04	0.07	0.09	0.17
N	40381	20584	12082	7088	4038

Notes: Each specification includes all covariates listed in the Table 3.1. Standard errors are shown in parentheses and are estimated using the Eicker-White formula to correct for heteroskedasticity and are clustered at the monitor level. *, ** and *** indicate significance at 10%, 5% and 1%, respectively.

Table 3.4 presents the main results for the IV first difference model, presented in Equations (3.3) and (3.4). Consistent with expectations, the first stage instruments perform well in predicting drops in particulate matter. The magnitude of the monitor non-attainment

Table 3.4: IV results for owner occupied units

	Interval				
	2	4	6	8	10
First Stage					
Monitor non-attainment	-4.42	-7.32	-7.79	-6.79	-7.25
	(1.24)***	(1.58)***	(1.37)***	(1.42)***	(1.83)***
County non-attainment	-0.66	-0.84	-0.86	-1.14	-2.58
	(0.24)***	(0.38)**	(0.58)	(0.90)	(0.97)***
F-stat	10.9	13.3	16.94	11.64	9.83
Second Stage					
PM_{10} (1/100)	-0.51	-0.63	-1.00	-2.03	-2.19
	(0.3)*	(0.32)*	(0.4)**	(0.81)**	(0.81)**
R^2	0.01	0.03	0.03	0.00	0.05
N	40381	20584	12082	7088	4038

Notes: Each specification includes all covariates listed in Table 3.1. Standard errors are shown in parentheses and are estimated using the Eicker-White formula to correct for heteroskedasticity and are clustered at the monitor level. *, ** and *** indicate significance at 10%, 5% and 1%, respectively.

instrument nearly doubles between two and four years, but then remains constant suggesting that air quality improves as much as it is going to in four years and then stays constant. The effect of county non-attainment is substantially smaller in magnitude than the monitor non-attainment effect, consistent with the findings of Auffhammer et al. (2009). Together the instruments are robust across specifications and generally have a joint F statistic above ten, which bolsters the argument that regulation has a strong effect on air quality, albeit in a heterogeneous manner.

The first stage estimates parallel the pattern demonstrated in Figures 3.1, 3.3, 3.5a, and 3.5b that air quality improved quickly after the onset of the 1990 CAAA and then stayed relatively constant. The coefficient on monitor non-attainment is -4.4 for the two-year lag, then increases to -7.3 for the four-year lag, and then remains at that level: -7.79, -6.79, and -7.25 for the six-, eight-, and ten-year lags, respectively. These results reinforce the idea that any increase in capitalization between four and ten years is due to a lag in capitalization, rather than continued improvements in air quality.

Turning to the second stage, the results robustly show that declines in PM_{10} cause housing prices to appreciate. All of the second stage coefficients are negative and statistically significant at the 10% level or better. Comparing these results to the least squares estimates of Table 3.3 reaffirm the importance of instrumental variables when measuring the valuation of air quality.

Most interestingly, the magnitude of the coefficient on PM_{10} monotonically increases as the lag between observations increases. For example, after two years, a one unit reduction in PM_{10} causes a treated housing unit to appreciate 0.51%, whereas after 10 years, that same unit in PM_{10} increases housing prices 2.19%. This trend is robust across specifications. The implied elasticities increase from -0.17 to -0.84 from two to ten years, and the associated MWTP increases from $45 to $183.

The ten year estimates are remarkably similar to estimates from other data sets and identification techniques. Bayer, Keohane, and Timmins (2009) use the IPUMS decennial census microdata and find a MWTP of $149 ($1982-84) for a unit reduction in PM_{10}. Bento et al. (2010) use decennial census data at the tract level and find MWTP estimates ranging from $75 to $263 ($2000) depending on the radius of included tracts.

Tables 3.5 and 3.6 present similar estimates as Table 3.4, but for owners and renters separated and only for the specification that includes all of the covariates and time fixed effects. Again, the first stage instruments are strong predictors of drops in PM_{10}, with the magnitude of effect leveling off after four years, and the second stage estimates show the same increase in the per unit rate of capitalization as the lag time increases.

Table 3.5: IV results for owner occupied units

	Interval				
	2	4	6	8	10
First Stage					
Monitor non-attainment	-4.80	-7.21	-8.22	-7.17	-7.32
	(1.18)***	(1.69)***	(1.57)***	(1.5)***	(1.72)***
County non-attainment	-0.58	-0.70	-0.68	-0.78	-2.38
	(0.23)**	(0.39)*	(0.62)	(1.01)	(1)**
F-stat	11.13	10.53	14.03	11.52	10.21
Second Stage					
PM_{10} (1/100)	-0.47	-0.74	-0.84	-1.67	-2.18
	(0.29)	(0.42)*	(0.46)*	(0.81)**	(0.99)**
R^2	0.03	0.08	0.10	0.13	0.15
N	22626	11997	7345	4504	2778

Notes: Each specification includes all covariates listed in Table 3.1. Standard errors are shown in parentheses and are estimated using the Eicker-White formula to correct for heteroskedasticity and are clustered at the monitor level. *, ** and *** indicate significance at 10%, 5% and 1%, respectively.

Table 3.6: IV results for renter occupied units

	Interval				
	2	4	6	8	10
First Stage					
Monitor non-attainment	-3.84	-7.10	-7.39	-6.39	-6.84
	(1.44)***	(1.65)***	(1.29)***	(1.54)***	(2.44)***
County non-attainment	-0.78	-1.03	-1.36	-2.04	-2.93
	(0.27)***	(0.43)**	(0.57)**	(0.87)**	(1.01)***
F-stat	9.31	12.18	19.32	10.86	6.99
Second Stage					
PM_{10} (1/100)	-0.27	-0.14	-0.70	-1.51	-1.65
	(0.32)	(0.3)	(0.34)**	(0.64)**	(0.63)***
R^2	0.00	0.02	0.02	-	0.04
N	17755	8587	4737	2584	1260

Notes: Each specification includes all covariates listed in Table 3.1. Standard errors are shown in parentheses and are estimated using the Eicker-White formula to correct for heteroskedasticity and are clustered at the monitor level. *, ** and *** indicate significance at 10%, 5% and 1%, respectively.

Generally, prior literature shows that rental prices tend to covary less than owner prices with a range of amenities, both in cross section and panel settings. Tables 3.5 and 3.6 present two pieces of evidence that are consistent with this general finding. First, the results suggest that rental rates are unresponsive to changes in air quality for up to four years (and then start to capitalize the changes), while the value of owner-occupied units begins to appreciate immediately. Second, the ten year capitalization rate is 25% smaller for renters than owners, suggesting that even in the long term rental units capitalize changes in air quality less.

Table 3.7 presents results for a just-identified model using only the monitor non-attainment instrument. When the set of instruments changes, the affected population changes, and the model will yield a different Local Average Treatment Effect (LATE). While this model forgoes identifying heterogeneous impacts of regulation, using only the monitor non-attainment instrument captures the effects on the population that underwent the most change. The first stage

Table 3.7: IV results for just identified model

	Interval				
	2	4	6	8	10
First Stage					
Monitor non-attainment	-4.25	-7.14	-7.56	-6.47	-6.43
	(1.24)***	(1.58)***	(1.37)***	(1.42)***	(1.79)***
F-stat	11.72	20.48	30.42	20.93	12.86
Second Stage					
PM_{10} (1/100)	-0.34	-0.47	-0.81	-1.53	-1.58
	(0.29)	(0.31)	(0.4)**	(0.8)*	(0.86)*
R^2	0.01	0.04	0.04	0.03	0.11
N	40381	20584	12082	7088	4038

Notes: Each specification includes all covariates listed in Table 3.1. Standard errors are shown in parentheses and are estimated using the Eicker-White formula to correct for heteroskedasticity and are clustered at the monitor level. *, ** and *** indicate significance at 10%, 5% and 1%, respectively.

results are strong with F statistics ranging between 11.7 and 30.4. Additionally, the same pattern holds of no further concentration reductions past a four-year lag. The capitalization estimates are less statistically significant: only the six-, eight-, and ten-year lag estimates are significant at the 10% level or better. The marginal capitalization rate increases in magnitude from -0.34% at a two-year lag to -1.58% at a ten-year lag, and thus, the results of a different LATE are consistent with the patterns of delayed capitalization observed in Table 3.4.

Table 3.8 offers a series of robustness checks that test whether the main result of delayed capitalization holds with various sample restrictions. As the sample changes with each test, the sample sizes and first stage results are withheld to minimize disclosure risk. However, in each case the first stage is very similar to that of Table 3.4.

Panel A tests for selection bias and excludes all units that turnover at some point during the interval.[65] Prior literature (e.g., Case et al. 1997) has shown that transacting units are not random, and frequently transacting units appreciate at a faster than normal rate. The results show that the capitalization rate increases from -0.51% to -2.22% from two to ten years. Thus, even units that do not transact appreciate in response to air quality changes and the specific pattern of delayed capitalization is not being driven by transacting properties only.

Panel B tests for attenuation bias and includes only intervals that begin in 1989 or 1991. For example, this criterion implies that for two year differences only intervals spanning 1989-1991 or 1991-1993 will be included. Since PM_{10} declines almost exclusively occurred in the early part of the decade, this specification more directly tests for delayed capitalization, instead

[65] Additionally, it would be useful to examine a specification in which only units that have recently transacted, and thus are most like sales, are included. This would to some extent test whether the results were purely an artifact of the self-reported nature of the data. Unfortunately, the parameter estimates are not identified given the small sample size. Concern over this issue is mitigated by the fact that rental prices experience the same delay in capitalization as owner-occupied home value and should not be affected by any bias in self-reporting. Further, both Cellini et al. (2010) and Case et al. (2006) use sales data and find delays in capitalization.

Table 3.8: Robustness checks

	Interval			
2	4	6	8	10
Panel A: Selection bias (includes only units that do not turnover)				
-0.51	-0.77	-0.79	-1.50	-2.22
(0.32)	(0.34)**	(0.39)**	(0.75)**	(0.83)***
Panel B: Attenuation bias (includes only intevals that start in 1989 or 1991)				
-0.57	-1.07	-1.23	-2.13	-2.19
(0.43)	(0.55)*	(0.51)**	(0.89)**	(0.81)***
Panel C: Attrition bias (includes only units present in ten year sample)				
-0.85	-0.91	-0.61	-2.07	-2.19
(0.36)**	(0.46)**	(0.37)	(0.83)**	(0.81)***
Panel D: All units within 3 miles of an air quality monitor				
-0.84	-0.87	-1.20	-2.27	-2.24
(0.37)**	(0.35)**	(0.47)**	(0.93)**	(0.98)**
Panel E: All units within 7 miles of an air quality monitor				
-0.53	-0.52	-0.85	-1.95	-1.95
(0.26)**	(0.29)*	(0.37)**	(0.82)**	(0.74)***
Panel F: All units within 10 miles of an air quality monitor				
-0.54	-0.52	-0.84	-1.99	-1.70
(0.25)**	(0.3)*	(0.38)**	(0.8)**	(0.65)***

Notes: Each coefficient is the coefficient on PM10 in a separate regression. Each specification includes all covariates listed in Table 3.1 and uses both monitor and county non-attainment instruments. Sample sizes are censored to reduce disclosure risk. Standard errors are shown in parentheses and are estimated using the Eicker-White formula to correct for heteroskedasticity and are clustered at the monitor level. *, ** and *** indicate significance at 10%, 5% and 1%, respectively.

of Table 3.4 relying indirectly on the fact that the effects of regulation are basically the same for lags four through ten. Including time intervals that do not include the beginning of the decade introduces the possibility that changes in air quality are in fact capitalized quickly, but small fluctuations in air quality (not accompanied by any change in housing prices) late in the decade attenuate the average capitalization rate for difference models with a short lag. The results show that the capitalization rate increases from -0.57% to -2.19% from two to ten years. Thus, including the late-decade observations do not bias the short lag estimates, which is consistent with idea that the estimated LATE is most accurate for the treated sample.

Panel C tests for attrition bias and includes only housing units that are present in the ten year interval sample. Because units enter and exit the sample and are not always interviewed even if they remain in the sample, the short lag panels are not balanced and contain units that do not appear in the long lag panels. If the probability for housing units to stay in the sample and be interviewed year after year is not random, then the short lag estimates may be biased by the inclusion of the non-random units. The results show that the capitalization rate increases from -0.85% to -2.19% from two to ten years. Thus, the factors related to whether or not a unit is interviewed do not appear to affect the valuation estimates.

Finally, Panels D-F examine the impact of the five mile matching radius on results by testing specifications that change the radius of match to three, seven, and ten miles.[66] In prior research, distance is primarily a means of measuring exposure, be it to Superfund sites (Greenstone and Gallagher, 2008; Cameron and McConnaha, 2006), toxic emissions (Banzhaf and Walsh 2008), or power plants (Davis, 2010). However, all of these examples consider very localized disamenities, whereas PM_{10}, while by no means a global pollutant, tends to be in similar concentrations for nearby areas. Thus, in expectation, there should not be a sharp discontinuity in valuation at any distance. The results show that, from two to ten years, the capitalization rate increases from -0.84% to -2.24% for the three mile cutoff, from -0.53% to -1.95% for the three seven cutoff, and from -0.54% to -1.70% for the ten mile cutoff. Thus, there

[66] At a distance of two miles or less, there are too few observations to conduct the analysis.

is no sharp discontinuity, and so the valuation estimates are not likely due to other factors. However, there is a modest decline in the magnitude of coefficients as distance increases, which is consistent with the findings of Bento et al. (2010).

Table 3.9 shows the results of the sorting on observables analysis giving the coefficient estimates of the effects of PM_{10} on demographic changes and the implied marginal effects. A negative coefficient implies that a reduction in PM_{10} increases the propensity of the given demographic group to move into rather than move out of that neighborhood. The results suggest that older households are more likely, though typically insignificantly so, to move into, rather than out of, a neighborhood in response to a decline in PM_{10}, which is consistent with the health risks that particulate matter poses to the elderly. The effect of PM_{10} on households with children under six years of age is positive for a two year lag and negative for lags four through ten, but always insignificant. Similarly for households with children under the age of eighteen, the effect of PM_{10} is positive for two and four year lags and negative for six, eight, and ten year lags, but again, always insignificant. The results suggest that Black or Hispanic headed households are more likely to move into a cleaned up neighborhood, though the coefficients are insignificant in all cases except the ten year lag. Lastly, the results indicate that households with less human capital are more likely to move in and households with more human capital are more likely to move out for all lags except ten, but all coefficients are insignificant.

Table 3.10 presents the results of the sorting on unobservables analysis that examines whether exogenous changes in PM_{10} cause turnover frequency to increase. A negative coefficient implies that a reduction in PM_{10} increases the rate if turnover, which would be suggestive of preference-based sorting (that is potentially uncorrelated with observable

Table 3.9: Demographic response to air quality changes

	Interval				
	2	4	6	8	10
Panel A: Over 60					
PM_{10} (1/100)	-10.03	-5.42	-10.31	-2.82	-4.01
	(4.93)**	(3.5)	(4.21)***	(3.74)	(4.27)
Marginal effect	-3.3%	-1.7%	-3.1%	-0.8%	-1.2%
Panel B: Children under 6					
PM_{10} (1/100)	2.97	-1.89	-1.91	-3.31	-3.44
	(3.93)	(2.95)	(2.3)	(3.91)	(4.61)
Marginal effect	1.2%	-0.7%	-0.7%	-1.2%	-1.2%
Panel C: Children under 18					
PM_{10} (1/100)	2.59	0.96	-1.15	-2.47	-0.10
	(3.83)	(2.62)	(3.12)	(3.44)	(3.89)
Marginal effect	1.0%	0.4%	-0.4%	-0.9%	0.0%
Panel D: Black or Hispanic					
PM_{10} (1/100)	-4.42	-3.31	-5.11	-8.96	-17.79
	(3.24)	(5.77)	(5.29)	(10.22)	(7.21)**
Marginal effect	-1.6%	-1.2%	-1.6%	-2.6%	-4.5%
Panel E: Human capital					
PM_{10} (1/100)	2.58	2.21	1.49	0.36	-0.23
	(3.14)	(2.1)	(1.98)	(3.33)	(3.41)
Marginal effect	1.0%	0.9%	0.6%	0.1%	-0.1%

Notes: Each coefficient represents a different regression. Each specification includes all covariates listed in Table 3.1 and is estimated using an IV probit specification using both monitor and county instruments. Samples include only units where the demographic characteristic of choice changed due to turnover during the interval. The dependent variable is binary and equals one if the unit gained in the given characteristic and equals zero if it lost. Maximum educational attainment of all household heads is classified into high school dropout, high school graduate, and college graduate. The variable 'human capital' takes the value one if there is turnover and the result of the turnover is an increase in human capital along the lines of the three classifications (i.e., high school dropout moves out and high school grad moves in) and takes the value zero if human capital declines. Sample sizes are censored to reduce disclosure risk. Standard errors are shown in parentheses and are estimated using the Eicker-White formula to correct for heteroskedasticity and are clustered at the monitor level. *, ** and *** indicate significance at 10%, 5% and 1%, respectively.

Table 3.10: Turnover in response to air quality changes

	Interval				
	2	4	6	8	10
Panel A: Turnover (probit)					
PM_{10} (1/100)	-1.49	1.33	0.33	0.43	-0.53
	(1.76)	(1.08)	(1.18)	(2.04)	(2.13)
Marginal effect	-1.5%	1.3%	0.3%	0.4%	-0.5%
Panel B: Number of moves (poisson)					
PM_{10} (1/100)	-1.34	1.31	-0.07	0.09	0.02
	(1.66)	(1.16)	(1.1)	(1.58)	(1.25)
Marginal effect	-0.3%	0.6%	0.0%	0.1%	0.0%

Notes: Each coefficient represents a different regression. Each specification includes all covariates listed in Table 3.1 and uses both monitor and county instruments. The sample is the same as in Table 3.3. For panel A, the dependent variable is binary and equal to one if new occupants moved in to the unit at some point during the interval and the model is estimated using a probit model in the second stage of the IV. For panel B, the dependent variable is the number of moves that occurred during the interval and the model is estimated using a Poisson model in the second stage of the IV. Standard errors are shown in parentheses and are estimated using the Eicker-White formula to correct for heteroskedasticity and are clustered at the monitor level. *, ** and *** indicate significance at 10%, 5% and 1%, respectively.

characteristics). In both the binary measure of turnover (Panel A) and the number of moves (Panel B), the coefficients are never significant and flip sign across the various intervals.

While a handful of the coefficient estimates in the sorting analyses are statistically significant, there is not enough support to suggest that preference based sorting is a factor in driving capitalization results. Further, no trends are apparent in the coefficient estimates across time that could explain the delay in capitalization.

While the sorting analyses offer no insights into the mechanism of delayed capitalization, they do offer more evidence that the exclusion restriction of the IV valuation estimates is valid. That is, conditional on other observable housing, neighborhood, and employment characteristics, nonattainment status does not affect household location decisions.

In addition, because there is no evidence of preference-based sorting, the capitalization estimates can be interpreted as willingness to pay and welfare conclusions can be made. In the context of the current study, the results imply that a substantial time delay is necessary in order to correctly estimate the welfare effects from the reductions in PM_{10} caused by the 1990 CAAA. In a larger context, this finding suggests that when the housing market is used for policy evaluation, substantial context may be necessary to accurately measure the welfare effects.

Conclusion

This paper uses unusually rich data and the large declines in pollution during the 1990s to examine the path of price discovery by which the housing market capitalized the changes in air quality. The results strongly suggest that households do indeed value air quality, but that a lag of as much as eight years exists before changes in particulate matter are fully capitalized into housing prices. These results inform policy by offering more evidence that households value clean air, as well as inform methods of policy evaluation. Price changes in the housing market are frequently used to assess damages and measure benefits, but if the timeframe of analysis is insufficient then the valuation could be severely underestimated.

While preference based sorting was unable to explain the dynamic path of air quality capitalization, other possibilities such as imperfect information about the housing market could be driving the results and future research should attempt to understand this mechanism. Further, the results could be very specific to air quality or even PM_{10} within the class of air pollutants. Other amenities and disamenities should be examined in this same framework to see if the salience of the amenity matters.

REFERENCES

Auffhammer, M.; A. M. Bento and S. E. Lowe. 2009. "Measuring the Effects of the Clean Air Act Amendments on Ambient Pm10 Concentrations: The Critical Importance of a Spatially Disaggregated Analysis." *Journal of Environmental Economics and Management*, 58(1), 15-26.

Banzhaf, S. and R. P. Walsh. 2008. "Do People Vote with Their Feet? An Empirical Test of Tiebout's Mechanism." *American Economic Review*, 98(3), 843-63.

Bayer, P.; F. Ferreira and R. McMillan. 2004. "Tiebout Sorting, Social Multipliers and the Demand for School Quality." NBER Working Paper 10871.

Bayer, P.; N. Keohane and C. Timmins. 2009. "Migration and Hedonic Valuation: The Case of Air Quality." *Journal of Environmental Economics and Management*, 58(1), 1-14.

Becker, R. and J. V. Henderson. 2000. "Effects of Air Quality Regulations on Polluting Industries." *Journal of Political Economy*, 108(2), 379-421.

Bento, A. M.; M. Freedman and C. Lang. 2010. "Spatial and Social Disparities in the Benefits from Air Quality Improvements." Cornell University Working Paper.

Cameron, T. A. and I. T. McConnaha. 2006. "Evidence of Environmental Migration." *Land Economics*, 82(2), 273-90.

Case, B.; P. F. Colwell; C. Leishman and C. Watkins. 2006. "The Impact of Environmental Contamination on Condo Prices: A Hybrid Repeat-Sale/Hedonic Approach." *Real Estate Economics*, 34(1), 77-107.

Case, B.; H. O. Pollakowski and S. M. Wachter. 1997. "Frequency of Transaction and House Price Modeling." *Journal of Real Estate Finance and Economics*, 14(1-2), 173-87.

Case, B. and J. M. Quigley. 1991. "The Dynamics of Real-Estate Prices." *Review of Economics and Statistics*, 73(1), 50-58.

Case, K. E. and R. J. Shiller. 1989. "The Efficiency of the Market for Single-Family Homes." *American Economic Review*, 79(1), 125-37.

Cellini, S. R.; F. Ferreira and J. Rothstein. 2010. "The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design." *Quarterly Journal of Economics*, 125(1), 215-61.

Chay, K. Y. and M. Greenstone. 2003. "Air Quality, Infant Mortality, and the Clean Air Act of 1970." NBER Working Paper 10053.

Chay, K. Y. and M. Greenstone. 2005. "Does Air Quality Matter? Evidence from the Housing Market." *Journal of Political Economy*, 113(2), 376-424.

Currie, J. and M. Neidell. 2005. "Air Pollution and Infant Health: What Can We Learn from California's Recent Experience?" *Quarterly Journal of Economics*, 120(3), 1003-30.

Daniels, M. J.; F. Dominici; J. M. Samet and S. L. Zeger. 2000. "Estimating Particulate Matter-Mortality Dose-Response Curves and Threshold Levels: An Analysis of Daily Time-Series for the 20 Largest Us Cities." *American Journal of Epidemiology*, 152(5), 397-406.

Davis, L. 2010. "The Effect of Power Plants on Local Housing Values and Rents." *Review of Economics and Statistics*, Forthcoming.

Dominici, F.; M. Daniels; S. L. Zeger and J. M. Samet. 2002. "Air Pollution and Mortality: Estimating Regional and National Dose-Response Relationships." *Journal of the American Statistical Association*, 97(457), 100-11.

Glaeser E. L. and J. Gyourko. 2007. "Housing Dynamics." HIER Working Paper #2137.

Greenstone, M. and J. Gallagher. 2008. "Does Hazardous Waste Matter? Evidence from the Housing Market and the Superfund Program." *Quarterly Journal of Economics*, 123(3), 951-1003.

Hall, J. V.; A. M. Winer; M. T. Kleinman; F. W. Lurmann; V. Brajer and S. D. Colome. 1992. "Valuing the Health Benefits of Clean-Air." *Science*, 255(5046), 812-17.

Hanna, R. 2010. "U.S. Environmental Regulation and FDI: Evidence from a Panel of U.S. Based Multinational Firms." *American Economic Journal: Applied Economics,* 2(3), 158-89.

Harrison, D. and D. L. Rubinfeld. 1978. "Hedonic Housing Prices and Demand for Clean-Air." *Journal of Environmental Economics and Management*, 5(1), 81-102.

Henderson, J. V. 1996. "Effects of Air Quality Regulation." *American Economic Review*, 86(4), 789-813.

Minnesota Population Center. *National Historical Geographic Information System: Pre-release Version 0.1.* Minneapolis, MN: University of Minnesota 2004. Available at http://www.nhgis.org

Nadeau, L. W. 1997. "EPA Effectiveness at Reducing the Duration of Plant-Level Noncompliance." *Journal of Environmental Economics and Management*, 34(1), 54-78.

Ridker, R. G. and J. A. Henning. 1967. "Determinants of Residential Property Values with Special Reference to Air Pollution." *Review of Economics and Statistics*, 49(2), 246-57.

Rosen, S. 1974. "Hedonic Prices and Implicit Markets - Product Differentiation in Pure Competition." *Journal of Political Economy*, 82(1), 34-55.

Tiebout, C. M. 1956. "A Pure Theory of Local Expenditures." *Journal of Political Economy*, 64(5), 416-24.

Unites States Environmental Protection Agency. 2010. www.epa.gov/

APPENDIX

In the AQS database, 3080 record observations for at least one year during 1989-2000. Of those, 594 monitors meet the time requirements necessary to be included in the analysis. Of those, only 378 meet the reliability requirements. Three of these monitors are excluded because they are in the same tract as another monitor; the multiple monitor tracts were located in Hampden County, Massachusetts and Sweetwater County, Wyoming. In these cases, the monitor with the most valid days of observation was selected to be included. So the final sample size is 375.

Table A1 compares included to excluded monitors for levels and changes in PM_{10}, as

Table A1: PM levels and changes for included and excluded monitors

		In sample	Out of sample	
			Reliable	Non-reliable
PM, 1989-1990	mean	32.9	28.7	28.0
	st. dev.	10.4	10.2	12.7
	N	375	579	461
PM, 1991-1996	mean	27.9	24.1	24.8
	st. dev.	7.5	8.4	16.7
	N	375	1382	341
PM, 1999-2000	mean	25.7	24.3	22.3
	st. dev.	8.1	11.4	10.0
	N	375	658	330
ΔPM, 2000-1990	mean	-7.2	-3.1	-7.8
	st. dev.	5.8	2.8	9.0
	N	375	18	239

Notes: The decadal difference row only requires monitors to have observations in 1989-1990 and 1999-2000. The 18 monitors in the Reliable, out of sample column are not included in the main analysis because they do not have a reliable mid-decade reading.

well as indicates the number of monitors meeting the reliability criteria. The levels for in-sample monitors tends to be higher than excluded monitors, but not significantly so. Decadal changes are more comparable between the groups. While we are only able to include a fraction of the monitors in the entire sample, we are confident our set of monitors is representative of the whole and that selection bias is not an issue.

Figures 2.5a and 2.5b clearly demonstrate two facts. First, monitor attainment status is a much better predictor of PM10 drops than county attainment status. Second, most of the reductions in PM10 occurred in the early part of the decade. With these two facts in mind, we constructed the instruments for our main analysis. Our county (monitor) instrument is the ratio of years that the county (monitor) is out of attainment to the number of years for which there is a record during the time span 1992-1997. In this case, our instruments are based on PM_{10} concentrations in the years 1991-1996.

Table A2 shows average reductions in PM_{10} for combinations of county and monitor nonattainment status (with the number of monitors represented in each cell in parentheses). The

Table A2: 1990-2000 PM_{10} reductions by county and monitor instrument

			Monitor instrument			
	Range		0	(0,0.5)	[0.5,1)	1
			-6 (348)	-14.9 (18)	-11.4 (5)	-18.8 (4)
County instrument	0	-5.6 (257)	-5.5 (255)	-11.4 (2)		
	(0,0.5]	-0.5 (1)		-0.5 (1)		
	(0.5,1)	-6.2 (10)	-5.2 (9)	-15.3 (1)		
	1	-9.3 (107)	-7.5 (84)	-16.5 (14)	-11.4 (5)	-18.8 (4)

Notes: The first number in each cell is the reduction in PM10 over the decade. The second number, in parentheses, is the number of monitors in each cell. The first column of values shows the distribution of reductions and monitors split only by the county instrument. The first row of values shows the distribution of reductions and monitors split only by the monitor instrument. The remaining cells show the distribution of reductions and monitors for the intersection of both instruments.

first column shows reductions in PM_{10} by county nonattainment instrument, regardless of monitor nonattainment status. The averages suggest that counties out of attainment for all years during 1992-97 experienced a 3.7 $\mu g/m^3$ larger decrease in PM_{10} than counties in attainment for the whole time period. Addidtionally, the number of monitors in each cell demonstrates the bi-modal nature of the county instrument. The first row, which shows reductions in PM_{10} by monitor nonattainment instrument regardless of county nonattainment status, reveals that monitors out of attainment in for all years 1992-97 saw a 12.8 $\mu g/m^3$ larger drop in PM_{10} than monitors always in attainment. The number of monitors in each cell tapers off, which is consistent with the implementation of the 1990 CAAA. The interior of Table A2 breaks down the average PM_{10} reductions for various combinations of the county and monitor instruments. The general patterns of the unconditional means hold. However, the county non-attainment differential drops to 2.0 $\mu g/m^3$ when nonattainment monitors are factored out. While merely suggestive, these results are consistent with a more concerted effort by local authorities to reduce pollution levels near offending monitors more so than near non-offending monitors in nonattainment counties.

Two different instrument constructions are shown in Tables A3 and A4 for illustrative purposes. Table A3 displays the PM10 reductions for binary county and monitor instruments that equal one if the county or monitor is ever designated non-attainment during 1992-1997. Similar reductions are seen for each instrument separately.[67] Table A4 also uses binary instruments, but restricts the years that can trigger the instrument to 1995 and 1996. This construction is most consistent with Chay and Greenstone (2005) whose instrument equals one if the county is

[67] The changes in PM_{10} concentrations in Figure 5b do not exactly match those in Table A3 because, as explained in Section 3, data for years 1989 and 1999 can be used to in place of data from 1990 and 2000, respectively, when no data exist for the latter years.

Table A3: 1990-2000 PM_{10} reductions by 1992-1997 binary instrument

			Monitor status	
	Range		0	1
			-6 (348)	-14.9 (27)
County status	0	-5.6 (257)	-5.5 (255)	-11.4 (2)
	1	-8.9 (118)	-7.2 (93)	-15.1 (25)

Notes: County (monitor) status equals one if that county (monitor) is designated non-attainment in any year 1992-1997, and zero otherwise. The first number in each cell is the average reduction in PM10 over the decade for the monitors in that cell. The second number, in parentheses, is the number of monitors in each cell. The first column of values shows the distribution of reductions and monitors split only by the county instrument. The first row of values shows the distribution of reductions and monitors split only by the monitor instrument. The remaining cells show the distribution of reductions and monitors for the intersection of both instruments.

Table A4: 1990-2000 PM_{10} reductions by 1995-1996 binary instrument

			Monitor status		
	Range		0	1	-
			-6.4 (352)	-16.9 (9)	-5 (14)
County status	0	-5.6 (257)	-5.6 (249)		-4.9 (8)
	1	-8.9 (118)	-8.4 (103)	-16.9 (9)	-5.2 (6)

Notes: County (monitor) status equals one if that county (monitor) is designated non-attainment in either 1995 or 1996, and zero otherwise. A monitor status of "-" implies that those monitors either had no recorded reading or readings that were not reliable in 1995 and 1996. The first number in each cell is the average reduction in PM10 over the decade for the monitors in that cell. The second number, in parentheses, is the number of monitors in each cell. The first column of values shows the distribution of reductions and monitors split only by the county instrument. The first row of values shows the distribution of reductions and monitors split only by the monitor instrument. The remaining cells show the distribution of reductions and monitors for the intersection of both instruments.

designated non-attainment in either 1975 or 1976. Differential county level reductions equal 3.3 $\mu g/m^3$, which is unchanged from Table A3 and similar to the 3.7 $\mu g/m^3$ reduction observed in Table A2. However, due to a lack of reliable monitor readings in 1995 and 1996, our monitor-level instrument suffers from sample size issues. Just one-third of the monitors designated as some way non-attainment in Tables A2 and A3 remain the non-attainment bin in Table A4 (The remaining two-thirds either get lumped in to the monitor in-attainment group or cannot be assigned in either the in-attainment or out-of attainment bin because there are no reliable observations for 1995 or 1996). This instrument construction is unusable for our disaggregated tract- and ring-level analysis, but the binary 95-96 instrument was examined in Table 2.5 and similar results were observed compared to the ratio instrument.