THE DYNAMICS OF HOUSE PRICE CAPITALIZATION AND LOCATIONAL SORTING:

EVIDENCE FROM AIR QUALITY CHANGES

by

Corey Lang *
University of Rhode Island

CES 12-22 September, 2012

Abstract

Despite extensive use of housing data to reveal valuation of non-market goods, the process of house price capitalization remains vague. Using the restricted access American Housing Survey, a high-frequency panel of prices, turnover, and occupant characteristics, this paper examines the time path of capitalization and preference-based sorting in response to air quality changes caused by differential regulatory pressure from the 1990 Clean Air Act Amendments. The results demonstrate that owner-occupied units capitalize changes immediately, whereas rent capitalization lags. The delayed but sharp rent capitalization temporally coincides with evidence of sorting, suggesting a strong link between location choices and price dynamics.

JEL codes: R23, R31, Q51, Q53
Keywords: hedonic valuation, capitalization, sorting, air quality

*Thanks to Ian Schmutte for assistance with the RDC proposal, to seminar participants at several universities, the EPA, and CES, and to Kevin Roth and Andrew Waxman for valuable comments. Special thanks to Antonio Bento and Matthew Freedman for guidance and suggestions throughout this project. Any opinions and conclusions expressed herein are those of the author and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed.

1 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. Rosen's model assumes costless relocation and thus immediate and full price capitalization, and this assumption is implicit in most empirical applications. However, if this assumption is false, estimates of the amenity value could be biased as a result of too short of a study time span. Further, too long of a time span may also produce biased estimates if important determinants of house prices, which change on the time span of a decade but not one or two years, are unobserved.

This paper addresses these dynamic extensions of Tiebout's ideas in the context of large improvements in air quality that occurred during the 1990s. Specifically, I examine the path of prices for both owners and renters that capitalize a change in air quality – going beyond *if prices change* to *how prices change*. Further, I analyze preference-based sorting in response to air quality changes and seek to understand the links between sorting behavior and the dynamics of capitalization. For example, after air quality has improved a new type of resident with more money or different preferences may move in, driving up prices and displacing the old type of resident.

To address these questions, I utilize the American Housing Survey (AHS), which 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 is essential for examination of the dynamics of capitalization and sorting.[1] I match

[1] 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

housing units from the AHS to particulate matter (PM_{10}) concentrations measured from nearby air quality monitors. I exploit the structure of the 1990 Clean Air Act Amendments (CAAA) to identify quasi-exogenous variation in PM_{10}. Similar to the seminal work of Chay and Greenstone (2005), I employ an Instrumental Variables (IV) strategy that relies on non-attainment designations of the air quality standards to address the endogenous relationship between air quality and housing prices.

Importantly, I gained access to the confidential 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 confidential 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 IV identification strategy can exploit localized air quality regulation intensity (which would be masked at the MSA level) stemming from within-county differential regulatory pressure. Auffhammer et al. (2009) show that within-county regulation is the principal factor determining reductions in PM_{10}, and Bento et al. (2012) use these insights to construct instruments similar to those I use in this paper.

The results suggest that while both owners and renters value air quality, the path of capitalization markedly differs. Owner-occupied housing units capitalize changes in air quality immediately, and capitalization rates and marginal willingness-to-pay estimates (MWTP) stay fairly constant across time, lending credence to the hedonic method. On the other hand, renter-occupied housing units show statistically insignificant and economically small capitalization rates shortly after air quality changes, but the estimated valuation sharply increases at a lag of six years and continues to increase after that. Ten years after air quality began to change, estimated capitalization rates and MWTP are comparable to the owner-occupied units. This suite of results holds when just looking at movers or non-movers separately, excluding rent-controlled units, and several other robustness specifications. The results support the idea that the owner and renter market are fairly distinct, likely due to dramatic differences in the occupants and housing stock, which is documented by Glaeser and Gyourko (2007) and in this paper using the AHS. The

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.

different patterns of capitalization lead to arbitrage opportunities between the owner and renter market purely based on air quality. However, at a maximum the disparity in annual housing costs is only $124, which is unlikely to be enough for households to delay home purchase and certainly not enough to cover the financial costs of selling a home in order to be a renter.

The striking dynamics observed in the rental market offer an excellent opportunity to examine the interplay between valuation and preference-based sorting. I analyze changes in turnover and demographic variables related to age, race, education, and income in response to changes in air quality. The results suggest that neighborhoods that experience improvements in air quality see an increase in the turnover frequency and the likelihood of families with children moving in relative to other neighborhoods, but only at a lag of six or more years. In sum, the results indicate a temporal correspondence between capitalization and preference-based sorting and offer a strong empirical confirmation of Tiebout's ideas.

There are three main contributions of this paper. The first is to shed new light on how housing prices respond to a change in amenity levels. Despite extensive use of hedonic valuation, only one paper to date has addressed the dynamic details of capitalization. Cellini, Ferreira and Rothstein (2010) examine the effect of school bond spending on house price sales.[2] Their comprehensive sales data allow them to observe prices as much as 15 years after a bond referendum, which is typically 5-7 years after bond-related spending ceases. Cellini et al. use several models to estimate the WTP for a dollar of school bond money. While the dynamics of capitalization are slightly different for each model, valuation tends to increase for several years following the referendum and then stabilizes, likely reflecting the trend that bond spending ramps up for three to four years following the referendum and then declines. I complement what Cellini et al. have done by examining a different amenity, which could lead to a different capitalization pattern given different preferences for and information about air quality *versus* school spending.

Second, I go beyond looking at the owner market, as Cellini et al. have done, and examine the dynamics of rental market capitalization as well. This aspect complements recent work by Grainger (2012), who assesses the distributional impacts of the 1990 CAAA in terms of gains for owners versus renters, and Davis (2011), who examines capitalization of new power

[2] Pope (2008) is an additional paper that examines the responsiveness of house prices. In his setting it is not amenity levels that change but information about the amenity, and he finds prices are quick to respond to the new information.

plants. Both papers find that rental prices are responsive to amenity changes, but less so than their owner-occupied counterparts. However, both papers use decennial census data, and thus only address differences in levels of capitalization, whereas I employ the high frequency AHS to investigate differences in patterns of capitalization.

Third, this paper complements prior work that examined preference-based sorting (see Cameron and McConnaha (2006) and Greenstone and Gallagher (2008) in response to Superfund cleanups; Bayer et al. (2007) and Cellini et al. (2010) for school quality; Card et al. (2008) for racial preferences; Banzhaf and Walsh (2008) for toxic emissions; Davis (2011) for power plants). However, with the exception of Cellini et al., who happen to find no evidence of sorting, these papers use decennial census data, which do not permit inquiry into how sorting and capitalization may temporally relate. Again by using the AHS, I contribute to this literature by finding empirical evidence suggesting that capitalization and sorting are dynamically linked, thus reinforcing the idea that households "vote with their feet."

The paper proceeds as follows. The next section discusses the setting in which I examine capitalization patterns, namely the CAAA and particulate matter. Section 3 discusses data. I explain the use of the CAAA as a quasi-experiment in Section 4, and discuss the empirical models in Section 5. Sections 6 and 7 present the capitalization results and the sorting results, respectively. Section 8 discusses the results and concludes.

2 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).[3]

Responding to calls to action and mounting scientific evidence, the United States Congress passed the 1970 CAA, which was the first federal legislation establishing air quality control.[4] The 1970 CAA created the 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 CAAA, the second major update of the CAA, replaced TSPs with PM_{10} to parallel current scientific understanding of pollution's effects.[5] 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 $\mu g/m^3$.[6] In order to achieve the NAAQS, the EPA held counties and states accountable for meeting those standards. Importantly, 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.

In terms of the effectiveness of these regulations, Greenstone (2004) surprisingly finds that county attainment status has no bearing on improvements in SO_2. Auffhammer et al. (2009)

[3] 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 damages crops and buildings and reduces visibility (EPA 2010).

[4] 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 (EPA 2010).

[5] 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.

[6] 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.

bolster Greenstone's finding for the case of PM_{10}, but additionally show that concentrations are responsive to whether an individual air quality monitor exceeds the standards. It is this finding that motivates my focus on monitor exceedences as an instrument, rather than county attainment status. Beyond reducing pollution concentrations, strong effects of the Clean Air Act and its amendments have been well documented with respect to industrial activity (e.g., Henderson 1996, Becker and Henderson 2002, Hanna 2010) and health outcomes (Chay and Greenstone 2003, Currie and Neidell 2005).

3 DATA

This section discusses the source and relevant features of the air quality data, regulatory data, housing data (including how local house price trends are managed), neighborhood data, and county economic activity data. The Data Appendix provides a complete listing of all variables used in the regression analysis.

3.1 Air quality data

Individual air quality monitor records were obtained from the Air Quality Standards (AQS) database (EPA 2009). Monitors are placed throughout the United States, but are primarily located in urban areas. 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 monitor-year observations to those that are sufficiently reliable.[7] For the key measure of air quality, I use the annual mean PM_{10} concentration.

3.2 Attainment status

I construct an attainment status for each monitor-year directly from the AQS data using the same threshold rules as the EPA's county designation. If in a given year 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

[7] 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, I exclude monitor-year observations that are affected by "extreme natural events" beyond human influence.

two days or more, then that monitor-year is designated non-attainment. Monitor attainment status serves as the main instrument in the IV model. In addition, I obtained the county attainment designations from the annual CFR.

3.3 Housing unit and occupant characteristics

Housing data for the years 1989-2001 were obtained from the restricted access American Housing Survey National Sample (AHS). The AHS is a panel of housing units that are surveyed every two years, usually between August and November. The AHS collects information about self-reported house value (if owner occupied), rent (if renter occupied), dwelling characteristics (e.g., number of bedrooms, number of bathrooms, whether the unit is rent controlled), occupant characteristics (e.g., race, education, income), and when the current occupants moved in. Importantly, the AHS follows units, not occupants, yielding a high frequency panel of prices.

Unlike the public use version, the restricted access AHS records the census tract where each unit is located. Using GIS, I determined the distance between each tract's geographic center, or centroid, and surrounding air quality monitors. I was 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 three miles apart. As units get further away from monitors, more measurement error is introduced into the key air quality variable. Additionally, Bento et al. (2012) find that valuation declines as distance from monitors increases. Thus, a three mile cutoff is used to balance measurement error and valuation concerns with sample size concerns. Additional details about the matching process will follow in Section 5, which describes the specifics of the empirical approach.

The first two columns of Table 1 provide summary statistics examining the differences between included versus excluded housing units and their occupants for owners and renters separately.[8] Primarily reflecting the propensity of air quality monitors to be placed in urban areas, the two samples are substantially different. 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

---

[8] In addition to the sample restrictions based on distance to a monitor, I exclude unit-year observations when prices are interpolated/"hot decked" or error coded. Further, prices are edited if they are obviously miscoded by omission or insertion of a digit. For instance, if a unit's price sequence is $100,000, $10,000, $100,000, then the middle price would be changed to $100,000 as it appears a zero was omitted. After editing prices, I exclude units whose prices change by a factor of four in a two-year interval. Lastly, I omit unit-intervals that do not have a continuous sequence of valid observations for all years in that interval.

included units are more likely to be elderly, more likely to be Black or Hispanic, and less likely to be a high school graduate. Given these disparities, results may not extend to the entire population.

A major theme of this paper is disparities between the owner and renter market. Column 3 of Table 1 compares owners to renters for the included sample and shows that the units themselves as well as the occupants differ dramatically between the two groups. One of the starkest differences is in the turnover rate, where renter units were 32% more likely (40% compared to 8%) to turn over between 1987 and 1989. Intuitively, the fluidity of the renter market may cause rental prices to respond faster to changes in amenities than owner prices, but this is of course not the only difference between the two markets. Owner-occupied units are on average nearly 1000 square feet larger, have 1.2 more bedrooms, have 0.5 more bathrooms, and are 30% more likely to have a dishwasher and 20% more likely to have an air conditioner than renter-occupied units. Owner households make on average $17,750 more in income and are 20% less likely to be a minority than renter households. These substantial differences are especially remarkable given that all included units are located in primarily urban areas and that, in general, much of the difference between the owner and renter market stems from the fact that most rental units are in the urban core and owner-occupied units are often in the suburbs. These differences reinforce the motivation for estimating all models separately for owner-occupied and renter-occupied units.

3.4 Local housing market trends

Because the scope of the study is national and thus compares appreciation rates across many cities and regions, it is necessary to control for local housing market trends. If regional house price trends are correlated with patterns of air quality improvements, the valuation estimates could reflect those trends instead of responses to air quality changes. I institute a novel method that controls for local housing market trends by using external data. Freddie Mac publishes the Conventional Mortgage Home Price Index (CMHPI), which gives quarterly estimates of home price levels (Freddie Mac 2010). 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 state specific indices for every state. If a housing unit is located within one of the 11 MSAs, then it is matched to that index, otherwise it is matched to the state index. Using each of these indices, all housing prices were brought to 2001 levels using the third quarter (due to the sampling schedule of the AHS) index for each year 1989-2001. Given the adjustments made to house prices using the CMHPI, the dependent variable in the valuation regressions is, in essence, appreciation relative to the local housing market trend. Thus, while the scope of the study is national, I am able to compare housing price changes from one part of the country to another because all price changes are relative to a smaller market.[9]

3.5 Neighborhood characteristics

The socioeconomic characteristics of neighbors are, without question, an important piece of a housing unit's value or rent. 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.[10] Of course, these data are only available in 1990 and 2000, and I assume a linear trend to impute values for all years 1989-2001.

3.6 County economic characteristics

The correlation between air pollution and economic activity can confound the analysis of how households value air quality. While households value clean air, they also value jobs and income (and probably much more so). The IV strategy discussed in the next section is designed to eliminate the bias that arises from this correlation. But additionally, I control for county economic characteristics to mitigate this bias. From the Bureau of Economic Analysis (BEA), I use annual average county income, and I include employment data from the County Business Patterns (CBP) database. CBP gives yearly employment counts broken down by Standard Industrial Classification/North American Industry Classification System codes for each county.[11]

---

[9] One limitation of this method is that the CMHPI is an index of owner-occupied prices only. Glaeser and Gyourko (2007) argue that the relationship between rental prices and owner prices is weaker than under conditions of perfect arbitrage, but empirically the two markets are still strongly related. In the Online Appendix, I test alternative means of controlling for regional rent trends and find similar results as when using the CMHPI.

[10] The available years of AHS data only give census tract codes for 1980 and 1990, with 1990 codes only appearing for observations in 1997 and after. I obtained the historical tract boundaries from NHGIS (Minnesota Population Center 2004). Using GIS, I overlaid historical boundary files with the 2000 boundaries to determine weights such that the neighborhood variables from Geolytics could be reconstructed for 1980 or 1990 boundaries.

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

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.

4 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. (2012). 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 non-attainment 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, in addition to county attainment status, use individual air quality monitor attainment designations. 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 (Auffhammer et al. 2009).

Consistent with Bento et al., I use individual air quality monitor readings instead of county averages, however I focus primarily on the monitor attainment designation to identify exogenous changes in air quality. The monitor instrument I construct is the ratio of years that the monitor is out of attainment to the number of years for which there is a record. I 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. This construction is clarified with respect to the various panels of differing lags in the following section where I discuss the empirical framework.

For my monitor non-attainment instrument to be valid, it must be correlated with PM_{10} changes and only affect housing prices through its impact on air quality. First, I examine the time series of PM_{10} with respect to attainment designation. Figure 1 shows PM_{10} trends for the years 1989-2001 for three mutually exclusive groups: 1) monitors that exceeded the standards at some point during 1989-2001, 2) monitors in counties that were designated non-attainment at some point during 1989-2001, and 3) monitors never designated non-attainment nor located in a non-attainment county. Figure 1 shows that, broadly, PM_{10} levels converged over the 1990s, which is exactly the intent of the NAAQS. The largest reductions in PM_{10} clearly occurred for non-attainment monitors, which declined by a total of 18.0 $\mu g/m^3$ over the years 1989-2001, 10.1 more than in-attainment monitors in non-attainment counties and 10.5 more than monitors in in-attainment counties. While county non-attainment monitors do experience additional declines in PM_{10} compared to county in-attainment monitors, the differential is small and the trends are graphically very similar. Figure 1 suggests that the non-attainment status for monitors is strongly correlated with PM_{10} reductions, but that non-attainment status for counties only weakly so. Also, Figure 1 demonstrates that the majority of air quality improvements occurred early in the decade; 80% of total PM_{10} reductions observed for non-attainment monitors had occurred by 1992.

Next consider the relationship between attainment designation and housing unit and occupant characteristics. Similar to Cellini et al. (2010), Columns 4 and 5 of Table 1 examine the pre-treatment conditions of the housing units, comparing unit and occupant characteristics in 1989 for housing units matched to non-attainment monitors versus in-attainment monitors, again comparing owners and renters separately. Owner-occupied units in monitor non-attainment areas are worth less, are smaller, and have fewer appliances. 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 or other locational amenities across different areas that happen to have different air quality levels. Rental rates in monitor non-attainment areas are actually higher, which underscores the difficulty with cross sectional estimation. While many demographic variables are not significantly different between the groups, non-attainment areas have fewer high school graduates, but more college graduates. One of the most important pre-treatment characteristics to consider is prior appreciation of housing prices because an existing price trend could bias valuation estimates. Table 1 shows no statistically significant difference in prior appreciation; a robustness check in Section 6 confirms inclusion of past price changes does not affect results. In addition, there is not a statistically significant difference in either changes in household income or the rate of turnover between 1987 and 1989. While the groups are not perfectly balanced pre-treatment, the statistics offer no reason to be concerned about confounding effects of the identification strategy.

My IV approach relies on the assumption that, conditional on other observable housing, neighborhood, and county 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 indirectly affects home prices. In fact, a substantial body of research has shown that air quality regulation has a significant effect deterring new firms (Becker and Henderson 2000), off-shoring production (Hanna 2010), and on employment levels (Kahn and Mansur 2010, Walker 2011). However, these findings focus on the economic decisions and outcomes of polluting firms only, which represent a relatively small portion of the total economic activity of an area.

To get a sense of the impact of regulation on the *overall* economic robustness of an area, I analyzed the effect of individual monitor exceedences, as well as the EPA county level attainment designation, on annual measures of county average income, population, and total employment. The results, detailed in the Online Appendix, show that both non-attainment measures have an insignificant effect on the three economic measures considered, which is consistent with the ambiguous total impact of non-attainment status found by Kahn and Mansur (see their Table 5). While this analysis is limited, it gives no reason to think that the IV exclusion restriction is violated.

5 EMPIRICAL FRAMEWORK

In this section, I outline the econometric approach I use to estimate the capitalization of air quality changes into housing prices and how households may adjust their location preferences in response to the air quality changes. My primary estimates come from IV first difference models with lags between observations ranging from two to ten years. The range of lags enables examination of how the rate of capitalization and sorting behavior may change over time. Further, this section details how housing units are matched with air quality monitors and the exact construction of the instrument.

The first and second stage equations of the first difference IV analysis are

$$\Delta PM_{ijt} = \gamma N_{jt} + (\Delta \boldsymbol{X_{ijt}})\boldsymbol{\delta} + \Delta v_{ijt} \tag{1}$$

and

$$\Delta p_{ijt} = \theta\left(\Delta \widehat{PM}_{ijt}\right) + (\Delta \boldsymbol{X_{ijt}})\boldsymbol{\beta} + \Delta \varepsilon_{ijt} \tag{2}$$

where p_{ijt} is the natural log of price (either house value or annual rent) of housing unit i matched to monitor j in time period t, PM_{ijt} is the concentration of PM_{10} for monitor j in time period t, and $\boldsymbol{X_{ijt}}$ is a vector of unit and location covariates (a complete list of which appears in the Data Appendix). The instrument in Equation (1), N_{jt}, is based on monitor non-attainment designation. If in Equations (1) and (2), the first difference is taken between years t_1 and t_2, then the monitor instrument equals the ratio of non-attainment years to the total number of years of observation during the years [t_1, t_2-1]. θ is the coefficient of interest and measures the change in log of house price due to a change in PM_{10}, or the capitalization rate. Implicit in the first difference model is a unit specific fixed effect that absorbs time invariant characteristics of areas that might be correlated with house prices and air quality, such as climate and topographical features, proximity to open space, and transportation infrastructure.

In order to examine the dynamics of capitalization, I construct multiple 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 earlier intervals, as most of the PM_{10} reductions occurred early in the decade. 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 three miles in distance and include only the match of least distance in the event of more than one monitor-interval matching to a unit-interval. This method ensures 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 intervals. This strategy balances the competing goals of minimizing measurement error and maximizing sample size.[12]

To explore preference-based sorting in response to exogenous changes in air quality, I estimate a variant of the IV first difference model above, except the second stage dependent variable measures turnover and changes in demographic characteristics of occupants. The second stage equation becomes

$$y_{ijt} = \theta\left(\Delta\widehat{PM}_{ijt}\right) + (\Delta\boldsymbol{X_{ijt}})\boldsymbol{\beta} + \Delta\varepsilon_{ijt} \qquad (3).$$

First, I define y_{ijt} to be binary and equal to one if unit i has turned over during the interval ending at t and zero otherwise. In this case, Equation (3) is estimated using a logistic model. Second, I estimate models of sorting on observable characteristics that may be correlated with preferences for air quality: household head over the age of 65, the presence of children under 18, household head is either Black or Hispanic, educational attainment of the household head, and household income. For all characteristics except household income, y_{ijt} is binary and equals one if the demographic characteristic of choice has "increased" (i.e., the household head of the prior occupants was under 65 and the household head of the new occupants is over 65) and equals zero if the demographic characteristic of choice has "decreased".[13] All unit-intervals that either do not experience turnover or do not have a change in the given demographic group are excluded. This restriction allows me to estimate (3) for this class of sorting models using a logistic model. For household income, y_{ijt} equals the differenced log household income of the entering household from the exiting household, and the sample is still restricted to those units that turnover. Since y_{ijt} is continuous in this case, Equation (3) is estimated using two stage least squares.

Much like housing prices, there are regional trends in turnover and demographics that must be taken into account in order to accurately measure preference-based sorting at a national scale. Unfortunately, no analogous CMHPI exists for demographic changes. As a second best, I include changes in turnover and demographic characteristics at the MSA level (aggregated from

[12] I choose not to interpolate PM10 values between monitors, so that the attainment designations can be clearly assigned at the monitor level.

[13] The dependent variable for the educational attainment specification equals one if educational attainment has increased with the new occupants (i.e., a high school dropout moves out and either a high school graduate or university graduate moves in) and zero if educational attainment has declined.

decennial census data) as covariates in the estimation of (3).[14] With this strategy, the locational choice of, for example, households over 65 in response to changes in PM_{10} is conditional on the MSA level changes in the over 65 population.

6 CAPITALIZATION RESULTS

Tables 2 and 3 present the results for the IV first difference model, presented in Equations (1) and (2), for interval lengths of two, four, six, eight, and ten years for owner-occupied and renter-occupied units, respectively.[15] Consistent with expectations, the monitor non-attainment instrument performs well in predicting drops in particulate matter. Just focusing on Table 2, the coefficient on monitor non-attainment is -5.75 for the two-year lag, then increases to -12.10 for the four-year lag and -19.21 for the six-year lag, and then remains roughly at that level: -22.03 and -16.22 for the eight- and ten-year lags, respectively. F-statistics range from 21.36 to 75.61, suggesting a strong instrument and, indeed, that regulation has a strong effect on air quality. The first stages are approximately the same for owners and renters, but differ slightly due to different values for covariates and a slight difference in the set of monitors matched to the sample. The first stage estimates parallel the pattern observed in Figure 1 that air quality improved quickly after the onset of the 1990 CAAA and then stayed relatively constant across areas.

Turning to the second stage, Table 2 convincingly shows that declines in PM_{10} cause owner-occupied housing prices to appreciate. Coefficient estimates range from -0.0059 to -0.0139, with all estimates statistically significant at the 5% level or more. Corresponding elasticities range from -0.23 to -0.63 and MWTP estimates range from \$83 to \$212, with the six-year and ten-year interval consistently being the minimum and maximum, respectively.[16] These estimates are consistent with estimates from other data sets and identification techniques (e.g.,

[14] Effective MSAs are constructed by grouping units in the same MSA, and then grouping units not located in a MSA by state.

[15] For the sake of comparison, the Online Appendix additionally reports results from cross-section and first difference (without instrumenting) models. The first difference results are qualitatively similar to those using an instrument, but the magnitudes of capitalization are much smaller reinforcing the importance of the IV strategy.

[16] MWTP is defined as the amount a household would be willing to pay annually for a one unit reduction in PM_{10}. For owners, this calculation requires an estimate of the annual cost of owning a home. I abstract from all costs except mortgage payments and assume an 8% interest rate and a 30-year mortgage. (8% is roughly in the middle of the range of mortgage rates observed during 1989-2001, according to historical rates downloaded from www.hsh.com/mtghst.html in May 2011.) For renters, the calculation is straightforward – just the estimated coefficient times the annual rent.

Bayer et al. (2009) find a MWTP of $149 ($1982-84) and the relevant MWTP estimate from Bento et al. (2012) is $150 ($2000)). The between-interval variance in estimates is likely due to the volatility of prices in the owner market, and the "true" MWTP probably lies in the middle of the range. While there is variance in the estimates across intervals, there is no pattern of increasing or decreasing capitalization. Thus, the results suggest that owner-occupied units capitalize changes in air quality quickly.[17]

Switching to the second stage renter results, Table 3 demonstrates that, like owner-occupied units, rental units capitalize changes in air quality, but the pattern of capitalization is markedly different. For intervals two and four, coefficient estimates are small (-0.0021 and -0.0025) and statistically insignificant, but then sharply jump to a statistically significant -0.0076 for the six-year interval and continue all the way to -0.0203 at a lag of ten years. At a lag of two years, the elasticity is -0.06 and the MWTP is $13. At a lag of ten years, the elasticity and MWTP are ten times larger at levels of -0.63 and $161. Despite being initially low, elasticities for rental price responses reached parity with owners by the ten year interval.

Comparing the owner and renter results leads to counterintuitive insights. First, conventional wisdom may suggest that the renter market is better suited for valuation studies than the owner market because the high turnover rates and low financial costs of moving should cause prices to adjust quickly. However, the results indicate that the fluidity of a market does not correspond to the speed of capitalization. Second, the results suggest that the rental market and owner market are not as closely linked as conventional wisdom holds. The valuation sharply and immediately diverges, but there is no evidence that this divergence is arbitraged either through financial instruments or people delaying home purchase to rent. However, the maximum disparity in annual costs resulting from air quality improvements between homeowners and renters is $124, which may not be enough to induce arbitrage given the substantial economic and psychological costs of moving.

[17] Another possibility is that households are forward-looking and, as a result, the two-year and perhaps the four-year capitalization coefficients are capturing households' expectations about future air quality improvements. Cellini et al. (2010) develop a forward looking estimator that delivers a larger estimated valuation for years shortly following the bond referendum compared to the more myopic estimators. Expectations are certainly possible in this setting, but given the IV framework it is infeasible to test for it.

6.1 Robustness

Tables 4 and 5 offer a series of robustness checks that test whether the main results hold with various sample restrictions.[18] 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 similar to that of Tables 2 and 3.

Panel A in both tables examines whether the spatial distribution of owners and renters could affect the capitalization patterns. If the spatial distribution is uneven and different areas experience different appreciation rates, then comparisons between owners and renters could be flawed. For these estimates, the sample is restricted to air quality monitors that are matched to both owner- and renter-occupied units guaranteeing that the two groups have geographic overlap. The results support the finding of immediate and constant capitalization for owners and delayed capitalization for renters.

Panel B in both tables tests for the possibility that pre-treatment trends could be affecting capitalization estimates. Each regression now includes the change in unit-specific log housing prices 1987-1989 as a control variable. Intuition would suggest that the short interval estimates would be more affected by pre-treatment trends than long interval coefficients. The two- and four-year interval estimates for owners are consistent with those in Table 2, statistically significant at the 1% level, and remarkably stable. The two- and four-year interval estimates for renters are also consistent with those in Table 3, showing small and insignificant capitalization. Capitalization patterns for other intervals are consistent as well for both owners and renters.

One concern with the main results of Tables 2 and 3 is that two- and four-year intervals that occur late decade may only experience small fluctuations in air quality not accompanied by any change in housing prices. Similar to classical measurement error, this could attenuate the estimated average capitalization rate for the shorter interval panels. Panel C in both tables includes only intervals that begin in 1989 or 1991. For example, this criterion implies that for the two-year dataset only the intervals 1989-1991 and 1991-1993 will be included. This specification ensures that each unit-interval observation contains the pivotal time of rapid PM_{10} change during the early 1990s and of all specifications in this paper is most akin to an event study. The results are similar to the patterns observed in Tables 2 and 3 and suggest that including the late-decade

[18] Additional specifications testing the robustness of the results to changes in covariates and CMHPI de-trending are presented in the Online Appendix. Results are generally consistent with those presented in Tables 2 and 3.

observations do not bias the short interval estimates. This finding is consistent with idea that the estimated local average treatment effect is most accurate for the treated sample, which in this case is the intervals experiencing the rapid PM_{10} change (Imbens and Angrist 1994).

Panel D of Table 4 examines capitalization for owner-occupied houses that do not turn over during an interval.[19] The purpose here is to check if information about sales prices percolates to households not taking part in that market. This is additionally important as a check on how self-reported prices follow those of units that do transact and can proxy for market prices.[20] The results again suggest immediate and full capitalization, as in Table 2. In addition to validating self-reported prices and ubiquitous information, these results also suggest that households living in cleaned up areas that do not move still reap financial gains from the policy.

Panels D and E of Table 5 test whether price rigidities in the rental market could be causing the delay in capitalization. The concern is that landlords tend to keep rental prices nominally unchanged from year to year when tenants renew a lease, which has been shown to be the case by Genesove (2003). Further, tenants in rent-controlled apartments tend to move less, and typically landlords can only modestly (if at all) adjust prices unless there is turnover.[21] Thus, only when tenants leave do landlords raise rents to market levels that would reflect a change in amenities. To address this possibility, the Panel D specification includes only unit-intervals with turnover and Panel E excludes all units under rent control. The results of Panel D show a remarkably similar pattern of delayed capitalization to Table 3, suggesting that delayed capitalization is common to all rental units, regardless of occupant tenure. In Panel E, the two- and four-year capitalization rates double compared to those seen in Table 3, and the four-year coefficient is significant at the 5% level. However, the ten-year capitalization estimate is still five times larger than the two-year capitalization estimate. This suggests that rent control does impede some short term capitalization, but a lag in full capitalization still persists.

The last two panels of each table examine the robustness of results when the radius of match changes from three miles to two and four miles. The results of the two mile match model are qualitatively similar to the main results in Tables 2 and 3, though there is less statistical

[19] Though it would be a valuable test similar to using sales data, examining only units that sold is not feasible due to small sample size.

[20] Kiel and Zabel (1999) examine bias stemming from self-reported prices and find that self-reported values tend to be inflated over market prices, but consistently so. Thus, as long as self-reported values are being compared to other self-reported values, the bias should not affect results.

[21] This is especially important since rent control is practiced in Los Angeles (Los Angeles Housing Department 2012), which represents a substantial portion of the treatment group.

significance reflecting the decrease in sample size. The four mile match model is consistent with the main results, though the magnitude of the coefficients declines.

7 SORTING RESULTS

The delayed but sharply rising capitalization rates observed in the rental market offer an excellent opportunity to examine the relationship between valuation and preference-based sorting.[22] The capitalization results suggest that something fundamentally changes going from the two- and four-year interval to the six-, eight-, and ten-year intervals. If there is a relationship between sorting and valuation, a similar shift in the turnover rate or demographics should occur at the six-year interval.

Table 6 presents the results from estimates of Equation (3), giving the second stage coefficients on PM_{10}, as well as the implied marginal effects.[23] Panel A examines whether changes in PM_{10} cause turnover to increase, and the model is estimated on the full sample appearing in Table 3. A negative coefficient implies that a reduction in PM_{10} increases the rate of turnover, which would be suggestive of preference-based sorting (with preferences potentially uncorrelated with observable characteristics). The estimated coefficients are positive and insignificant for intervals two and four. However, for intervals six and eight, the coefficient becomes negative and statistically significant. Further, the jump in the probability of turnover is economically significant; the total estimated increase in probability from four to six years given the total reductions in PM_{10} is 10.1%. For interval ten, the coefficient is positive and insignificant again.

Panels B through F of Table 6 show the results of the sorting on observables analysis giving the estimates of the effects of PM_{10} on demographic changes. 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 of Panel B suggest that

[22] The possibility that appreciation in the owner market spills over into the renter market, perhaps at a substantial lag, could also explain the disparate capitalization paths. This is explored in the Online Appendix, but little support is found for this idea. However, there is evidence that owner prices aggregate both past owner and renter price changes, furthering the implication of the capitalization results that the owner market more efficiently incorporates information into prices than the renter market. Additionally in the Online Appendix, the corresponding owner-occupied sorting analysis is presented in. Given the relative infrequency of moves among owner-occupiers, the data is inadequate to draw too many conclusions.

[23] Sample sizes and first stage results are omitted due to disclosure concerns. In all cases, the first stage estimates are similar in magnitude and significance with those shown in Table 3.

improvements in PM_{10} consistently increase the likelihood of households with children under 18 years of age of moving into an area. For intervals two and four the estimates are small and insignificant, but for intervals six, eight, and ten the estimates are larger and statistically significant. The six-year interval estimate implies a total increase in likelihood of 18.2%. Panel C suggests that older residents are more likely to be moving out of, rather than into, a neighborhood in response to a decline in PM_{10}. This result is surprising given that the elderly are vulnerable to adverse health effects from poor air quality. The results of Panel D suggest that the minority status of a household is unresponsive to changes in air quality, as coefficients flip signs multiple times and are mostly insignificant.[24] Panel E explores changes in the educational attainment of the household; as with minority status, the coefficients flip sign and are mostly insignificant. Lastly, the household income results in Panel F suggest that the income level of new residents was lower than old residents for interval two, but flips signs and is insignificant for the other intervals.

The results of the sorting analysis suggest that changes in turnover frequency and the location preferences of households with children are responsive to changes in air quality and correspond to the rental price capitalization patterns. Intuitively, households with children should value improvements in air quality more than households without children due to health concerns. However, we might also think that preferences may be correlated with income and education, but the results show no evidence of this. To be clear, I am unable to disentangle the causality between price movements and locational preferences. The results only suggest that they are strongly related.

8 DISCUSSION AND CONCLUSION

This paper uses rich housing market data and spatially disaggregated air quality data to examine changing housing prices and changing locational preferences in response to air quality changes brought on by differential regulatory pressure from the CAAA. This is the first paper to study the dynamic path of price capitalization in response to air quality changes and the first paper to examine the dynamics of capitalization in response to any changing amenity for owners and renters separately. The results strongly suggest a disconnect between air quality capitalization patterns in the owner and renter markets. While owner-occupied houses capitalize

[24] Examining Black households and Hispanic household separately does not qualitatively affect the results.

changes immediately and the capitalization rate stays fairly constant over time, renter-occupied units are slow to capitalize changes, and the capitalization rate increases substantially over time up to the point that the implied MWTP estimates are on par for the two groups.

There are several reasons why the disparities in dynamic capitalization paths between owners and renters could exist. First, owners may be more attentive of amenity levels given their anticipated tenure and financial stake in the neighborhood. This effect may be especially strong for an imperfectly observed good such as air quality. Second, the structure of the rental market may be such that it impedes rapid price changes. This claim is partially implicated in Table 5, which shows that short term capitalization rates increase when the sample excludes rental units that do not turn over or are subject to rent control. Third, if renters have heterogeneous preferences for air quality, changing locational preferences may lead to a lag of capitalization if households are slow to re-optimize their location (perhaps slow due to reasons one or two). Surprisingly, the fact that there is so much less turnover in the owner market does not seem to affect the capitalization patterns.

This paper examines preference-based sorting and finds evidence that the turnover of rental units and the propensity for households with children move into a unit both increase in recently cleaned up neighborhoods. Importantly, the timing of the sorting coincides with price capitalization, both at a six year lag, bolstering the long-standing intuition of Tiebout's "vote with your feet" ideas.

Given that we do see an increase in the turnover rate, there may be a concern that renters suffer a welfare loss when air quality improves (e.g., Starrett 1981, Sieg et al. 2004, Grainger 2012). The fear is that when amenities improve, all of the new value goes to the owner, who can then charge higher rent leaving the tenants no better off – and potentially worse off if the renters cannot afford the new rent and must re-locate into an area with worse air quality. While the sorting results in this paper suggest there is an increase in turnover, the results also show that the new occupants are no wealthier or better educated or whiter than the previous tenants.[25] A full welfare analysis is beyond the scope of this paper, but at a minimum the results suggest that air quality improvements had an ambiguous welfare effect on renters.

[25] Bayer et al. (2004) argue that a large part of house price capitalization (in response to school quality changes) is due to social multipliers. Wealthy households locate in good school districts and their presence further causes house prices to appreciate. Given the result in this paper that income does not change between old and new occupants, the capitalization rates are less likely to be affected by social multipliers and are thus a good indication of MWTP for air quality.

While the results as a whole inform policy by offering more evidence that households value clean air, more importantly the results inform methods of policy evaluation. Price changes in the housing market are frequently used to valuate preferences, measure benefits, and assess damages. The owner results, suggesting immediate and full capitalization of an imperfectly observed amenity, are a testament of the efficacy of the hedonic approach. However, the renter results suggest that if the timeframe of analysis is insufficient, then the valuation could be severely underestimated. 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 of MWTP could be as little as one quarter of the true valuation. Future work should continue to explore disparities in the owner and rental market in the response the amenity changes.

REFERENCES

Auffhammer, Maximilian, Antonio M. Bento, and Scott 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, Spencer and Randall P. Walsh. (2008). "Do people vote with their feet? An empirical test of Tiebout's mechanism." *American Economic Review* 98(3): 843-863.

Bayer, Patrick, Fernando Ferreira and Robert McMillan. (2004). "Tiebout Sorting, Social Multipliers and the Demand for School Quality." NBER Working Paper 10871.

Bayer, Patrick, Nathaniel Keohane and Christopher Timmins. (2009). "Migration and hedonic valuation: The case of air quality." *Journal of Environmental Economics and Management* 58(1): 1-14.

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

Bento, Antonio M., Matthew Freedman, and Corey Lang. (2012). "Spatial and Social Disparities in the Benefits from Air Quality Improvements." Cornell University Working Paper.

Cameron, Trudy Ann and Ian T. McConnaha (2006). "Evidence of environmental migration." *Land Economics* 82(2): 273-290.

Case, Bradford, Henry O. Pollakowski, and Susan M. Wachter. (1997). "Frequency of transaction and house price modeling." *Journal of Real Estate Finance and Economics* 14(1-2): 173-187.

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

Cellini, Stephanie R., Fernando Ferreira, and Jesse Rothstein. (2010). "The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design." *Quarterly Journal of Economics* 125(1): 215-261.

Chay, Kenneth Y. and Michael Greenstone. (2003). "Air Quality, Infant Mortality, and the Clean Air Act of 1970." NBER Working Paper 10053.

Chay, Kenneth Y. and Michael Greenstone. (2005). "Does air quality matter? Evidence from the housing market." *Journal of Political Economy* 113(2): 376-424.

Currie, Janet and Matthew Neidell. (2005). "Air pollution and infant health: What can we learn from California's recent experience?" *Quarterly Journal of Economics* 120(3): 1003-1030.

Daniels, Michael J., Francesca Dominici, Jonathan M. Samet, and Scott 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, Lucas. (2011). "The Effect of Power Plants on Local Housing Prices and Rents." *Review of Economics and Statistics*, 93(4): 1391-1402.

Dominici, Francesca, Michael J. Daniels, Jonathan M. Samet, and Scott L. Zeger. (2002). "Air pollution and mortality: Estimating regional and national dose-response relationships." *Journal of the American Statistical Association* 97(457): 100-111.

Freddie Mac. (2010). "Conventional Mortgage Home Price Index." Last retrieved November 1, 2010 from, www.freddiemac.com/finance/cmhpi/.

Genesove, David. (2003). "The nominal rigidity of apartment rents." *Review of Economics and Statistics* 85(4): 844-853.

Glaeser, Edward L. and Josheph Gyourko. (2007). "Arbistrage in Housing Markets." NBER Working Paper #13704.

Grainger, Corbett. (2012). "The Distributional Effects of Pollution Regulations: Do Renters Fully Pay for Cleaner Air?" forthcoming in *Journal of Public Economics*.

Greenstone, Michael and Justin Gallagher (2008). "Does hazardous waste matter? Evidence from the housing market and the superfund program." *Quarterly Journal of Economics* 123(3): 951-1003.

Hall, Jane, Arthur Winer, Michael Kleinman, Frederick Lurman, Vic Brajer, and Steven Colonme. (1992). "Valuing the Health Benefits of Clean-Air." *Science* 255(5046): 812-817.

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.

Henderson, J. Vernon. (1996). "Effects of air quality regulation." *American Economic Review* 86(4): 789-813.

Imbens, Guido W. and Joshua D. Angrist (1994). "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62(2): 467-475.

Kahn, Matthew E. and Erin T. Mansur. (2010). "How Do Energy Prices, and Labor and Environmental Regulations Affect Local Manufacturing Employment Dynamics? A Regression Discontinuity Approach." NBER Working Paper 16538.

Los Angeles Housing Department. (2012). "Rent Stabilization Ordinance Information." Retrieved February 12, 2012, from http://lahd.lacity.org/.

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

Pope, Jaren C. (2008). "Buyer information and the hedonic: The impact of a seller disclosure on the implicit price for airport noise." *Journal of Urban Economics* 63(2): 498-516.

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

Sieg, Holger, V. Kerry Smith, H. Spencer Banzhaf, and Randy Walsh. (2004). "Estimating the general equilibrium benefits of large changes in spatially delineated public goods." *International Economic Review* 45(4): 1047-1077.

Starrett, David A. (1981). "Land-Value Capitalization in Local Public-Finance." *Journal of Political Economy* 89(2): 306-327.

Tiebout, Charles M. (1956). "A Pure Thoery of Local Expenditures." *Journal of Political Economy* 64(5): 416-424.

Unites States Environmental Protection Agency. (2009). "Air Quality Standards Database." Last retrieved September 1, 2009, from www.epa.gov/aqspubl1/.

Unites States Environmental Protection Agency. (2010). "Particulate Matter (PM)." Last retrieved November 1, 2010, from www.epa.gov/.

Walker, W. Reed. (2011). "Environmental Regulation and Labor Reallocation: Evidence from the Clean Air Act." *American Economic Review* 101(3): 442-447.

Data Appendix

List of variables used as covariates in regressions

Unit Characteristics (from the American Housing Survey)

New roof between beginning and end of interval (1=yes)[*]
Kitchen remodeled between beginning and end of interval (1=yes) [*]
New or remodeled bathroom between beginning and end of interval (1=yes) [*]
Other addition between beginning and end of interval (1=yes) [*]
Rent control at beginning of interval (1=yes)[†]
Rent control at end of interval (1=yes)[†]

Neighborhood Characteristics (from the Decennial Census)

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[±]
ln(median family income)[≠]
Total housing units
Share of units occupied
Share of occupied units that are owner occupied

County Economic Characteristics (from County Business Patterns and the Bureau of Economic Analysis)

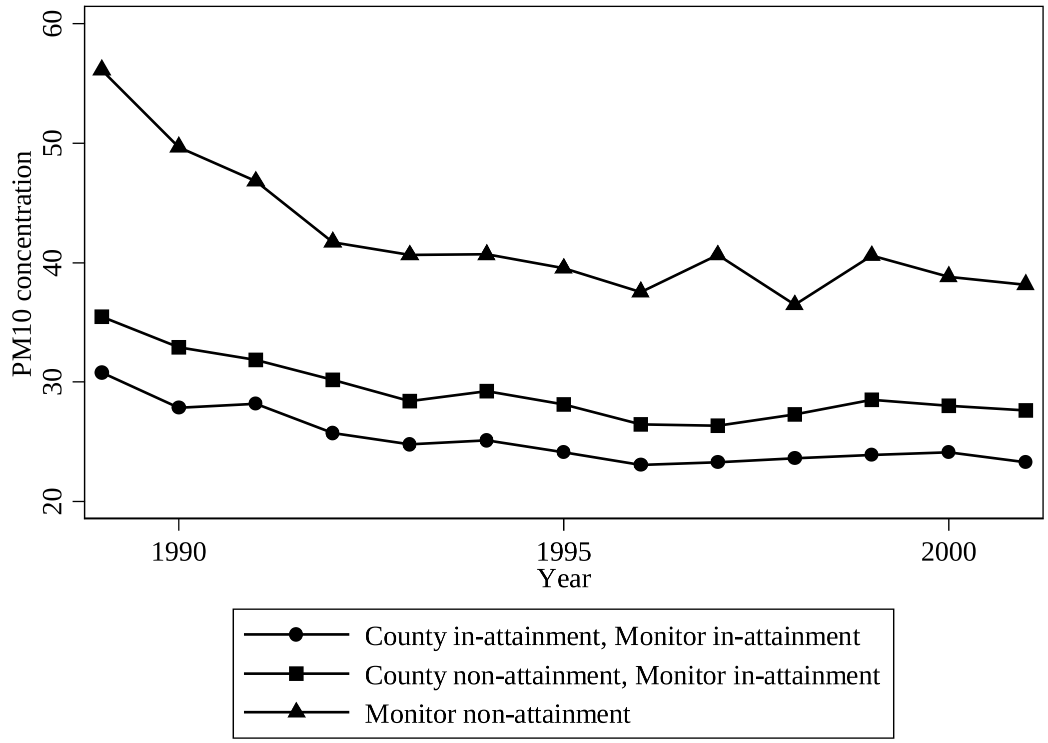
ln(number of jobs in construction +1)
ln(number of jobs in manufacturing +1)
ln(number of jobs in mining +1)
ln(number of jobs in agriculture or forestry +1)
ln(number of jobs in all other sectors +1)
ln(average county income)

---

[*] Only available for owner-occupied units.
[†] Only included for renter-occupied specifications.
[±] These tract level variables were replaced with MSA averages for the sorting analyses. Additionally, the education variables were replaced with a single variable, average educational attainment, equal to the share of college graduates minus the share of high school dropouts.
[≠] Removed for sorting analysis since county average income is available at higher frequency from BEA.

Figures and Tables

Figure 1: PM10 concentrations across time by attainment status



Notes: Each air quality monitor in the sample is split into attainment categories by the following rule. If a monitor ever exceeds the NAAQS PM10 standards 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, Monitor in-attainment' group. All other monitors are put in the 'County in-attainment, Monitor in-attainment' group.

Table 1: Housing and Occupant characteristics in 1989

Variable	Difference of in-sample and out-of sample: owners	Difference of in-sample and out-of sample: renters	Difference of owners and renters for in-sample units	Difference of non-attainment and in-attainment: owners	Difference of non-attainment and in-attainment: renters
	(1)	(2)	(3)	(4)	(5)
PM_{10}			-0.6	20.5*	17.8*
Market value (2000$)	-1975			-11835	
Annual rent (2000$)		-68			483
Household income (2000$)	-1676	-1134	17750*	-395	6584*
Percent of households with children under 18	-4.5*	-2.5	-1.2	0.1	1
Percent of households with head over 65	3.5*	1	15.2*	1	-4.3
Percent of households with Black or Hispanic head	5.9*	9.3*	-18.2*	1.9	-6.2
Percent of households with high school graduate as head	0.4	-5.8*	4.9*	-2.5	-7.4*
Percent of households with college graduate as head	-1.4	1.8	3.8*	3.5	9.6*
Number of bedrooms	-0.1*	-0.2*	1.2*	-0.1*	-0.1
Number of bathrooms	-0.1*	-0.1*	0.5*	0.1*	0
Square feet	-57.5	-85.8	967.8*	-409.6*	209.7
Percent of units with a dishwasher	-7.8*	-11.7*	29.9*	8.1*	3.7
Percent of units with central AC	-7.4*	-12.5*	18.8*	6.9*	3.6
Percent of units under rent control		4.2*			17.8*
Change in log market value, 1987-1989	0.002			-0.041	
Change in log annual rent, 1987-1989		0.02*			-0.031
Change in household income, 1987-1989	-295	735	221	2446	2197
Percent of units with turnover, 1987-1989	-0.9	-3.9*	-32.3*	2	-4

Notes: Each entry is the difference in means among units and occupants in the two groups given. Sample sizes are as follows: 1712 for in-sample owners, 17991 for out-of-sample owners, 1381 for in-sample renters, 7655 for out-of-sample renters, 1471 for in-attainment owners, 241 for non-attainment owners, 1139 for in-attainment renters, and 242 for non-attainment renters. * indicates a statistically significant difference of means at the 5% level.

Table 2: IV first difference results for owner occupied units

	Interval				
	2	4	6	8	10
First Stage					
Monitor non-attainment	-5.75	-12.10	-19.21	-22.03	-16.22
	(0.83)***	(1.44)***	(2.23)***	(2.53)***	(3.51)***
F-stat	47.80	70.47	74.11	75.61	21.36
Second Stage					
ΔPM_{10} (1/100)	-0.66	-1.01	-0.59	-0.83	-1.39
	(0.28)**	(0.32)***	(0.3)**	(0.39)**	(0.7)**
R-squared	0.001	0.007	0.010	0.070	0.099
Observations	17485	8961	5276	2972	1700

Notes: The dependent variable in the first stage is the change in annual PM10 concentrations, and the dependent variable in the second stage is the change in the natural log of house value. Each regression uses the full set of controls listed in the Data Appendix. 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: IV first difference results for renter occupied units

	Interval				
	2	4	6	8	10
First Stage					
Monitor non-attainment	-4.75	-12.34	-15.83	-17.53	-13.23
	(0.86)***	(1.12)***	(2.3)***	(2.61)***	(3.18)***
F-stat	30.44	122.36	47.46	45.23	17.34
Second Stage					
ΔPM_{10} (1/100)	-0.21	-0.25	-0.76	-1.18	-2.03
	(0.35)	(0.2)	(0.32)**	(0.58)**	(0.98)**
R-squared	0.032	0.088	0.147	0.110	0.041
Observations	15464	7394	4148	2166	1022

Notes: The dependent variable in the first stage is the change in annual PM10 concentrations, and the dependent variable in the second stage is the change in the natural log of annual rent. Each regression uses the full set of controls listed in the Data Appendix. 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 4: Additional second stage results for owner occupied units

	Interval				
	2	4	6	8	10
Panel A: Geographic overlap for owner and renter units					
ΔPM_{10} (1/100)	-0.67	-1.11	-0.72	-0.89	-1.41
	(-0.29)**	(0.34)***	(0.32)**	(0.42)**	(0.73)**
Panel B: Control for prior unit-specific price trends					
ΔPM_{10} (1/100)	-0.84	-0.85	-0.50	-0.54	-0.99
	(-0.32)***	(0.33)***	(0.31)	(0.36)	(0.59)*
Panel C: Only includes intervals that begin in 1989 or 1991					
ΔPM_{10} (1/100)	-0.80	-0.85	-0.59	-0.87	-1.39
	(-0.38)**	(0.33)**	(0.29)**	(0.39)**	(0.7)**
Panel D: Include only unit-interval observations without turnover					
ΔPM_{10} (1/100)	-0.59	-1.02	-0.56	-0.52	-1.54
	(-0.27)**	(0.35)***	(0.32)*	(0.44)	(0.86)*
Panel E: Match units to air quality monitors at a 2 mile radius					
ΔPM_{10} (1/100)	-0.73	-0.84	-0.10	-0.69	-1.62
	(-0.43)*	(0.37)**	(0.37)	(0.68)	(1)
Panel F: Match units to air quality monitors at a 4 mile radius					
ΔPM_{10} (1/100)	-0.37	-0.73	-0.40	-0.84	-1.39
	(-0.24)	(0.25)***	(0.24)*	(0.37)**	(0.61)**

Notes: Each coefficient comes from a separate regression, identical to those in Table 2, except for the sample. Sample sizes are censored to reduce disclosure risk. Panel A includes only monitors which are matched to both owner and renter units. Panel B includes the unit-specific change in the natural log of house value 1987-1989 as a covariate. 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 5: Additional second stage results for renter occupied units

	Interval				
	2	4	6	8	10
Panel A: Geographic overlap for owner and renter units					
ΔPM_{10} (1/100)	-0.18	-0.25	-0.76	-0.87	-1.71
	(0.36)	(0.2)	(0.33)**	(0.58)	(0.94)*
Panel B: Control for prior unit-specific price trends					
ΔPM_{10} (1/100)	0.02	-0.31	-0.60	-1.36	-2.02
	(0.42)	(0.2)	(0.4)	(0.7)*	(1.13)*
Panel C: Only includes intervals that begin in 1989 or 1991					
ΔPM_{10} (1/100)	-0.64	-0.30	-0.77	-1.22	-2.03
	(0.43)	(0.24)	(0.27)***	(0.59)**	(0.98)**
Panel D: Include only unit-interval observations with turnover					
ΔPM_{10} (1/100)	-0.16	-0.08	-0.89	-1.19	-2.12
	(0.5)	(0.36)	(0.35)***	(0.57)**	(1)**
Panel E: Exclude rent controlled units					
ΔPM_{10} (1/100)	-0.40	-0.50	-0.80	-1.19	-1.99
	(0.38)	(0.23)**	(0.36)**	(0.58)**	(0.98)**
Panel E: Match units to air quality monitors at a 2 mile radius					
ΔPM_{10} (1/100)	0.41	0.19	-0.56	-1.03	-1.71
	(0.5)	(0.22)	(0.27)**	(0.68)	(1.48)
Panel F: Match units to air quality monitors at a 4 mile radius					
ΔPM_{10} (1/100)	0.04	-0.29	-0.70	-0.81	-1.30
	(0.32)	(0.21)	(0.28)**	(0.38)**	(0.59)**

Notes: Each coefficient comes from a separate regression, identical to those in Table 3, except for the sample. Sample sizes are censored to reduce disclosure risk. Panel A includes only monitors which are matched to both owner and renter units. Panel B includes the unit-specific change in the natural log of annual rent 1987-1989 as a covariate. 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 6: IV first difference results for changes in turnover frequency and demographic characteristics of occupants for renter occupied units

	Interval				
	2	4	6	8	10
Panel A: Turnover					
ΔPM_{10} (1/100)	0.52	1.57	-3.33	-3.85	0.56
	(2.10)	(0.91)	(1.41)**	(2.08)*	(5.29)
Marginal effect (%)	0.20	0.58	-1.11	-1.14	0.13
Panel B: Presence of children under 18					
ΔPM_{10} (1/100)	-0.74	-2.03	-8.49	-7.51	-13.40
	(4.13)	(3.32)	(3.63)**	(3.69)**	(4.55)***
Marginal effect (%)	-0.29	-0.78	-3.04	-2.65	-4.30
Panel C: Household head over the age of 65					
ΔPM_{10} (1/100)	3.65	10.62	8.33	14.07	23.20
	(8.73)	(5.56)*	(5.54)	(5.08)***	(7.44)***
Marginal effect (%)	1.24	3.33	2.69	4.22	6.72
Panel D: Household head Black or Hispanic					
ΔPM_{10} (1/100)	-3.90	-0.88	4.26	-8.81	0.81
	(3.04)	(2.27)	(3.73)	(3.98)*	(5.59)
Marginal effect (%)	-0.29	-0.78	-3.04	-2.65	-4.30
Panel E: Educational attainment					
ΔPM_{10} (1/100)	-2.22	1.14	-4.67	0.88	-2.03
	(4.06)	(1.81)	(2.47)*	(3.34)	(6.27)
Marginal effect (%)	-0.87	0.44	-1.76	0.33	-0.75
Panel F: Natural log of household income					
ΔPM_{10} (1/100)	4.11	-0.85	-0.60	2.99	-0.90
	(2.08)**	(1.00)	(1.52)	(1.91)	(3.29)

Notes: Each coefficient represents a different regression. Each specification includes all covariates listed in the data appendix. For Panel A, the sample is the same as in Table 3b, and the dependent variable is binary and equal to one if new occupants moved in to the unit at some point during the interval. For Panels B-F, samples include only unit-intervals that experienced turnover and where the demographic characteristic of choice changed. For Panels B-E, the dependent variable is binary and equals one if the unit gained in the given characteristic and equals zero if it lost. For example, if the out-moving occupant does not have children and the in-moving occupant does, then this would be coded as one. If both out- and in-moving occupants have children, then that unit-interval would be excluded from the sample. For Panel E, maximum educational attainment of all household heads is classified into high school dropout, high school graduate, and college graduate, and the dependent variable takes the value one if there is an increase in educational attainment 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 educational attainment declines. The results shown in Panels A-E are estimated using an IV probit specification. Panel F is estimated using least squares, still in the IV first difference framework, as the dependent variable is continuous. 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.