Massachusetts Institute of Technology
Department of Economics
Working Paper Series

Does Air Quality Matter? Evidence from the Housing Market

Kenneth Y. Chay
Michael Greenstone

Working Paper 04-19

February 2004

Room E52-251
50 Memorial Drive
Cambridge, MA 02142

DOES AIR QUALITY MATTER?
EVIDENCE FROM THE HOUSING MARKET*

Kenneth Y. Chay
University of California at Berkeley and NBER

Michael Greenstone
MIT, American Bar Foundation, and NBER

February 2004

* We thank countless colleagues and seminar participants for very helpful comments and suggestions. Justine Hastings, Mark Rodini, and Pablo Ibarraran provided excellent research assistance. Greenstone received funding from the Alfred P. Sloan Foundation and Resources for the Future. Chay received support from the Institute of Business and Economic Research, the Institute of Industrial Relations, and an UC-Berkeley faculty grant. Funding from NSF Grant No. SBR-9730212 is also gratefully acknowledged.

DOES AIR QUALITY MATTER?
EVIDENCE FROM THE HOUSING MARKET

ABSTRACT

We exploit the structure of the Clean Air Act to provide new evidence on the capitalization of total suspended particulates (TSPs) air pollution into housing values. This legislation imposes strict regulations on polluters in "nonattainment" counties, which are defined by TSPs concentrations that exceed a federally set ceiling. TSPs nonattainment status is associated with large reductions in TSPs pollution and increases in county-level housing prices. When nonattainment status is used as an instrumental variable for TSPs, we find that the elasticity of housing values with respect to particulates concentrations range from –0.20 to -0.35. These estimates of the average marginal willingness-to-pay for clean air are far less sensitive to model specification than cross-sectional and fixed effects estimates, which occasionally have the "perverse" sign. We also find modest evidence that the marginal benefit of pollution reductions is lower in communities with relatively high pollution levels, which is consistent with preference-based sorting. Overall, the improvements in air quality induced by the mid-1970s TSPs nonattainment designation are associated with a $45 billion aggregate increase in housing values in nonattainment counties between 1970 and 1980.

Kenneth Y. Chay
Department of Economics
University of California, Berkeley
549 Evans Hall
Berkeley, CA 94720
and NBER
kenchay@econ.berkeley.edu

Michael Greenstone
Department of Economics
MIT
50 Memorial Drive, E52-391B
Cambridge, MA 02142-1347
American Bar Foundation and NBER
mgreenst@mit.edu

Introduction

Federal air pollution regulations have been among the most controversial interventions mandated by the U.S. government. Much of this controversy is generated by an absence of convincing empirical evidence on their costs and benefits. Thus, the credible estimation of the economic value of clean air to individuals is an important topic to both policy makers and economists.

The hedonic approach to estimating the economic benefits of air quality uses the housing market to infer the implicit price function for this non-market amenity. Here, researchers estimate the association between property values and air pollution, usually measured by total suspended particulates (TSPs), regression-adjusted for differences across locations in observable characteristics. After over 30 years of research, the cross-sectional correlation between housing prices and particulates pollution appears weak. A meta-analysis of 37 cross-sectional studies suggests that a 1-$\mu g/m^3$ decrease in TSPs results in a 0.05-0.10% increase in property values, which implies only a -0.04 to -0.07 elasticity (Smith and Huang 1995). As a result, many conclude that either individuals place a small value on air quality or the hedonic approach cannot produce reliable estimates of the marginal willingness-to-pay (MWTP) for environmental amenities.

These weak results may be explained by two econometric identification problems that could plague the implementation of the hedonic method. First, it is likely that the estimated housing price-air pollution gradient is severely biased due to omitted variables. We show that the "conventional" cross-sectional and fixed effects approaches produce estimates of MWTP that are very sensitive to specification and occasionally have the perverse sign, indicating that TSPs and housing prices are positively correlated. Second if there is heterogeneity across individuals in tastes for clean air, then individuals may self-select into locations based on these unobserved differences. In this case, estimates of MWTP may only reflect the preferences of subpopulations that, for example, place a relatively low valuation on air quality.

This paper exploits the structure of the Clean Air Act Amendments (CAAAs) to provide new evidence on the capitalization of air quality into housing values. The CAAAs marked an unprecedented attempt by the federal government to mandate lower levels of air pollution. If pollution concentrations in a county exceed the federally determined ceiling, then the Environmental Protection Agency (EPA) designates the county as 'nonattainment'. Polluters in nonattainment counties face much greater regulatory oversight than their counterparts in attainment counties.

We use nonattainment status as an instrumental variable for TSPs changes in first-differenced equations for the 1970-1980 change in county-level housing prices. The instrumental variables estimates indicate that the elasticity of housing values with respect to particulates concentrations range from –0.20 to –0.35. Estimates from a random coefficients model that allows for nonrandom sorting provide evidence consistent with the self-selection of individuals across counties based on taste heterogeneity and suggest that the marginal benefit of a pollution reduction may be lower in communities with relatively high pollution levels. However, the self-selection bias in estimates of the average MWTP appears to be small relative to the influence of omitted variables.

The "reduced form" relationships between the instrument and changes in TSPs and housing prices provide direct estimates of the benefits of the nonattainment designation. We find that TSPs declined by roughly 10 $\mu g/m^3$ (12%) more in nonattainment than attainment counties. Further, the data reveals that housing prices rose by approximately 2.5% more in these counties.

Our estimates of the average MWTP for clean air are far less sensitive to specification than the cross-sectional and fixed effects estimates. For example, we find evidence that nonattainment status is uncorrelated with virtually all other observable determinants of changes in housing prices, including economic shocks. Thus, it is not surprising that our results are largely insensitive to the choice of controls.

The discrete relationship between regulatory status and the previous year's pollution levels provide two opportunities to gauge the credibility of our results. In particular, the structure of the rule that determines nonattainment status ensures that there are nonattainment and attainment counties with <u>identical</u> and <u>almost identical</u> average TSPs levels in the regulation selection year. This allows us to implement "matching" and quasi-experimental regression discontinuity validation tests. Both of these tests confirm the reduced form results and our basic finding of an important relationship between TSPs and housing values.

The analysis is conducted with the most detailed and comprehensive data available on pollution levels, EPA regulations, and housing values at the county-level. Through a Freedom of Information Act request, we obtained annual air pollution concentrations for each county based on the universe of state and national pollution monitors. These data are used to measure counties' prevailing TSPs concentrations and nonattainment status. We use the *County and City Data Books* data file, which is largely based on the 1970 and 1980 Censuses, to obtain measures of housing values and housing and county characteristics.

Overall, the timing and location of the changes in TSPs concentrations and housing values in the decade from 1970 to 1980 provide evidence that TSPs may have a causal impact on property values. Taken literally, our estimates indicate that the improvements in air quality induced by the mid-1970s TSPs nonattainment designation are associated with a $45 billion aggregate increase in housing values in nonattainment counties during this decade. The results also demonstrate that the hedonic method can be successfully applied in certain contexts.

I. The Hedonic Method and Econometric Identification Problems

An explicit market for clean air does not exist. The hedonic price method is commonly used to estimate the economic value of this non-market amenity to individuals.[1] It is based on the insight that the utility associated with the consumption of a differentiated product, such as housing, is determined by the utility associated with the individual characteristics of the good. For example, hedonic theory predicts that an economic bad, such as air pollution, will be negatively correlated with housing prices, holding all other characteristics constant. Here, we review the method and the econometric identification problems associated with its implementation.

A. The Hedonic Method

Economists have estimated the association between housing prices and air pollution at least since Ridker (1967) and Ridker and Henning (1967). However, Rosen (1974) was the first to give this

[1] Other methods for non-market amenity valuation include contingent valuation, conjoint analysis, and discrete choice models. See Smith (1996) for a review and comparison of these methods.

correlation an economic interpretation. In the Rosen model, a differentiated good can be described by a vector of its characteristics, $Q = (q_1, q_2, \ldots, qn)$. In the case of a house, these characteristics may include structural attributes (e.g., number of bedrooms), the provision of neighborhood public services (e.g., local school quality), and local amenities (e.g., air quality). Thus, the price of the i^{th} house can be written as:

(1) $P_i = P(q_1, q_2, \ldots, q_n)$.

The partial derivative of $P(\bullet)$ with respect to the n^{th} characteristic, $\partial P/\partial q_n$, is referred to as the marginal implicit price. It is the marginal price of the n^{th} characteristic implicit in the overall price of the house.

In a competitive market the housing price-housing characteristic locus, or the hedonic price schedule (HPS), is determined by the equilibrium interactions of consumers and producers.[2] The HPS is the locus of tangencies between consumers' bid functions and suppliers' offer functions. The gradient of the implicit price function with respect to air pollution gives the equilibrium differential that allocates individuals across locations and compensates those who face higher pollution levels. Locations with poor air quality must have lower housing prices in order to attract potential homeowners. Thus, at each point on the HPS, the marginal price of a housing characteristic is equal to an individual consumer's marginal willingness to pay (MWTP) for that characteristic and an individual supplier's marginal cost of producing it. Since the HPS reveals the MWTP at a given point, it can be used to infer the welfare effects of a marginal change in a characteristic for a given individual.

In principle, the hedonic method can also be used to recover the entire demand or MWTP function.[3] This would be of tremendous practical importance, because it would allow for the estimation of the welfare effects of nonmarginal changes. Rosen (1974) proposed a 2-step approach for estimating the MWTP function, as well as the supply curve.[4] In some recent work, Ekeland, Heckman and Nesheim (2004) outline the assumptions necessary to identify the demand (and supply) functions in an additive

---

[2] See Rosen (1974), Freeman (1993), and Palmquist (1991) for details.

[3] Epple and Sieg (1999) develop an alternative approach to value local public goods. Sieg, Smith, Banzhaf, and Walsh (2000) apply this locational equilibrium approach to value air quality changes in Southern California from 1990-1995.

[4] Brown and Rosen (1982), Bartik (1987), and Epple (1987) highlight the strong assumptions necessary to identify the structural parameters with this approach. There is a consensus that empirical applications have not identified a situation where these assumptions hold and that the second stage MWTP function for an environmental amenity has never been reliably estimated (Deacon et al. 1998).

version of the hedonic model with data from a single market. The estimation details are explored in further work.[5]

B. Econometric Identification Problems

This paper's goal is to estimate the hedonic price function for clean air and empirically assess whether housing prices rise with air quality. In some respects, this is less ambitious than efforts to estimate primitive preference parameters and, in turn, MWTP functions. However from a practical perspective, it is of at least equal importance, because the consistent estimation of equation (1) is the foundation upon which any welfare calculation rests. This is because the welfare effects of a marginal change in air quality are obtained directly from the HPS. Further, an inconsistent HPS will lead to an inconsistent MWTP function, invalidating any welfare analysis of non-marginal changes, regardless of the method used to recover preference or technology parameters.

Consistent estimation of the HPS in equation (1) is extremely difficult since there may be unobserved factors that covary with both air pollution and housing prices.[6] For example, areas with higher levels of TSPs tend to be more urbanized and have higher per-capita incomes, population densities, and crime rates.[7] Consequently, cross-sectional estimates of the housing price-air quality gradient may be severely biased due to omitted variables. This is one explanation for the wide variability in HPS estimates and the relatively frequent examples of perversely signed estimates from the cross-sectional studies of the last 30 years (Smith and Huang 1995).[8] Our first goal is to solve this problem of omitted variables.

Self-selection to locations based on preferences presents a second source of bias in estimation of

---

[5] Heckman, Matzkin, and Nesheim (2002) examine identification and estimation of nonadditive hedonic models. Heckman, Matzkin, and Nesheim (2003) examine the performance of estimation techniques for both types of models.

[6] See Halvorsen and Pollakowski (1981) and Cropper et al. (1988) for discussions of misspecification of the HPS due to incorrect choice of functional form for observed covariates.

[7] Similar problems arise when estimating compensating wage differentials for job characteristics, such as the risk of injury or death. The regression-adjusted association between wages and many job amenities is weak and often has a counterintuitive sign (Smith 1979; Black and Kneisner 2003). Brown (1980) assumes that the biases are due to permanent differences across individuals and focuses on job 'changers'.

[8] Smith and Huang (1995) find that a quarter of the reported estimates have perverse signs; that is, they indicate a positive correlation between housing prices and pollution levels.

the average MWTP for clean air in the population. In particular, if individuals with lower valuations for air quality sort to areas with worse air quality, then estimates of the average MWTP that do not account for this can be biased upward or downward depending on the structure of preferences and the amount of sorting. This is an especially salient issue for this paper, because its identification strategy is based on comparisons of different regions of the country and these regions are determined by the level of TSPs. So if individuals have sorted based on pollution levels, the approach may produce estimates of the average MWTP that are based on non-random subpopulations. Thus, our second goal is to estimate the average MWTP while accounting for self-selection and to probe how MWTP may vary in the population.

The consequences of the misspecification of equation (1) were recognized almost immediately after the original Rosen paper. For example, Small (1975) wrote:

> I have entirely avoided...the important question of whether the empirical difficulties, especially correlation between pollution and unmeasured neighborhood characteristics, are so overwhelming as to render the entire method useless. I hope that...future work can proceed to solving these practical problems....The degree of attention devoted to this [problem]...is what will really determine whether the method stands or falls..." [p. 107].

In the intervening years, this problem of misspecification has received little attention from empirical researchers[9], even though Rosen himself recognized it.[10] This paper's aims are to focus attention on this problem of misspecification and to demonstrate how the structure of the Clean Air Act may provide a quasi-experimental solution in the case of housing prices and TSPs.[11]

II. Background on Federal Air Quality Regulations

Before 1970 the federal government did not play a significant role in the regulation of air pollution; that responsibility was left primarily to state governments.[12] In the absence of federal legislation, few states found it in their interest to impose strict regulations on polluters within their

---

[9] Graves et al. (1988) is an exception.
[10] Rosen (1986) wrote, "It is clear that nothing can be learned about the structure of preferences in a single cross-section..." (p. 658), and "On the empirical side of these questions, the greatest potential for further progress rests in developing more suitable sources of data on the nature of selection and matching..." (p. 688).
[11] In an earlier version of this paper, we outlined and implemented a method that exploits features of the Clean Air Act to estimate MWTP functions for TSPs (Chay and Greenstone 2001). This method requires very strong assumptions that may not be plausible, so this material is not presented here.
[12] Lave and Omenn (1981) and Liroff (1986) provide more details on the CAAAs. In addition, see Greenstone (2002) and Chay and Greenstone (1998).

jurisdictions. Concerned with the detrimental health effects of persistently high concentrations of suspended particulates pollution, and of other air pollutants, Congress passed the Clean Air Act Amendments of 1970.

The centerpiece of the CAAAs is the establishment of separate federal air quality standards, known as the National Ambient Air Quality Standards (NAAQS), for five pollutants. The stated goal of the amendments is to bring all counties into compliance with the standards by reducing local air pollution concentrations. The legislation requires the EPA to assign annually each county to either nonattainment or attainment status for each of the pollutants, on the basis of whether the relevant standard is exceeded. The federal TSPs standard is violated if either of two thresholds is exceeded: 1) the annual geometric mean concentration exceeds 75 $\mu g/m^3$, or 2) the second highest daily concentration exceeds 260 $\mu g/m^3$ (see Appendix Table 1).[13]

The CAAAs direct the 50 states to develop and enforce local pollution abatement programs that ensure that each of their counties attains the standards. In their nonattainment counties, states are required to develop plant-specific regulations for every major source of pollution. These local rules demand that substantial investments, by either new or existing plants, be accompanied by installation of state-of-the-art pollution abatement equipment and strict emissions ceilings. The 1977 amendments added the requirement that any increase in emissions from new investment be offset by a reduction in emissions from another source within the same county.[14] States are also mandated to set emission limits on existing plants in nonattainment counties.

In attainment counties, the restrictions on polluters are less stringent. Large-scale investments, such as plant openings and large expansions at existing plants, require less expensive (and less effective) pollution abatement equipment; moreover offsets are not necessary. Smaller plants and existing plants are essentially unregulated.

---

[13] In addition to the TSPs standard, Appendix Table 1 lists the industrial and non-industrial sources, abatement techniques, and health effects of TSPs. The TSPs standard prevailed from 1971 until 1987, when, instead of regulating all particulates with an aerodynamic diameter less than 100 micrometers, the EPA shifted its focus to fine particles. The regulations were changed to apply only to emissions of PM-10s (particles with an aerodynamic diameter of at most 10 micrometers) in 1987 and to emissions of PM-2.5s in 1997.

[14] Offsets could be purchased from a different facility or could be generated by tighter controls on existing operations at the same site (Vesilind, Peirce, and Weiner 1988).

Both the states and the federal EPA are given substantial enforcement powers to ensure that the CAAAs' statutes are met. For instance, the federal EPA must approve all state regulation programs in order to limit the variance in regulatory intensity across states. On the compliance side, states run their own inspection programs and frequently fine noncompliers. The 1977 legislation made the plant-specific regulations both federal and state law, which gives the EPA legal standing to impose penalties on states that do not aggressively enforce the regulations and on plants that do not adhere to them.

Nadeau (1997) and Cohen (1998) document the effectiveness of these regulatory actions at the plant level. Henderson (1996) provides direct evidence that the regulations are successfully enforced. He finds that ozone concentrations declined more in counties that were nonattainment for ozone than in attainment counties. Greenstone (2004) finds that sulfur dioxide nonattainment status is associated with modest reductions in sulfur dioxide concentrations. In this paper and Chay and Greenstone (2003a), we find striking evidence that TSPs levels fell substantially more in TSPs nonattainment counties than in attainment counties during the 1970s.[15]

III. Data Sources and Descriptive Statistics

To implement our analysis, we compile the most detailed and comprehensive data available on pollution levels, EPA regulations, and housing values for the 1970s. Here, we describe the data sources and provide some descriptive statistics. More details on the data are provided in the Data Appendix.

A. Data Sources

TSPs Pollution Data and National Trends. The TSPs data were obtained by filing a Freedom of Information Act request with the EPA that yielded the *Quick Look Report* file, which comes from the EPA's *Air Quality Subsystem* (AQS) database. This file contains annual information on the location of and readings from every TSPs monitor in operation in the U.S. since 1967. Since the EPA regulations are

[15] Greenstone (2002) provides further evidence on the effectiveness of the regulations. He finds that nonattainment status is associated with modest reductions in the employment, investment, and shipments of polluting manufacturers. Interestingly, the regulation of TSPs has little association with changes in employment. Instead, the overall employment declines are driven mostly by the regulation of other air pollutants.

applied at the county level, we calculated the annual geometric mean TSPs concentration for each county from the monitor-level data. For counties with more than one monitor, the county mean is a weighted average of the monitor-specific geometric means, with the number of observations per monitor used as weights. The file also reports the four highest daily monitor readings.

Our 1970 and 1980 county-level measures of TSPs are calculated with data from multiple years. In particular, the 1970 (1980) level of TSPs is the simple average over a county's nonmissing annual averages in the years 1969-72 (1977-80). These formulas mitigate measurement error in the TSPs measures. Further the EPA's monitoring network was still growing in the late 1960s, so the 1969-72 definition allows for a larger sample.

There are two primary reasons for our exclusive focus on TSPs rather than on other forms of air pollution. First, TSPs is the most visible form of air pollution and has the most pernicious health effects of all the pollutants regulated by the CAAAs.[16] Second, the EPA's monitoring network for the other air pollutants was in its nascent stages in the early 1970s and the inclusion of these pollutants in our models severely restricts the sample size.[17]

TSPs Attainment/Nonattainment Designations. The EPA did not begin to publicly release the annual list of TSPs nonattainment counties until 1978. We contacted the EPA but were informed that records from the early 1970s "no longer exist." Consequently, we used the TSPs monitor data to determine which counties exceeded either of the federal ceilings and assigned these counties to the nonattainment category; all other counties are designated attainment. We allowed these designations to vary by year and based them on the previous year's concentrations. This is likely to be a reasonable approximation to the EPA's actual selection rule, because it is based on the same information that was available to the EPA. The Data Appendix provides more details on our assignment rule.

---

[16] See Dockery et al. (1993), Ransom and Pope (1995), and Chay, Dobkin, and Greenstone (2003) and Chay and Greenstone (2003a and 2003b) for evidence on the effects of TSPs on adult and infant health, respectively.

[17] Only 34 out of our sample of 988 counties were monitored for all of the other primary pollutants regulated by the 1970 CAAAs at the beginning and end of the 1970s. Alternatively when the sample is limited to counties monitored for TSPs and one other pollutant, the sample sizes are 135 (carbon monoxide), 49 (ozone), and 144 (sulfur dioxide). We separately examined the relationship between housing values and levels of ozone, sulfur dioxide, and carbon monoxide in the 1970s and found no association. Chay and Greenstone (1998) find modest evidence that changes in ozone concentrations during the 1980s were capitalized into housing prices.

Housing Values and County Characteristics. The property value and county characteristics data come from the 1972 and 1983 *County and City Data Books* (*CCDB*). The *CCDB*s are comprehensive, reliable, and contain a wealth of information for every U.S. county. Much of the data is derived from the 1970 and 1980 *Censuses of Population and Housing*.

Our primary outcome variable is the log-median value of owner-occupied housing units in the county. The control variables include demographic and socioeconomic characteristics (population density, race, education, age, per-capita income, poverty and unemployment rates, fraction in urban area), neighborhood characteristics (crime rates, doctors and hospital beds per-capita), fiscal/tax variables (per-capita taxes, government revenue, expenditures, fraction spent on education, welfare, health, and police), and housing characteristics (e.g., year structure was built and whether there is indoor plumbing). The Data Appendix contains a complete set of the controls used in the subsequent analysis.

The Census data contains fewer variables on the characteristics of homes and neighborhoods than is ideal. For example, these data do not contain information on square feet of living space, garages, air conditioning, lot size, crime statistics, or schooling expenditures per student. We explain our identification strategy in greater detail below, but we believe that it overcomes some of the limitations of the Census data. In particular, we include county fixed effects to control for permanent, unobserved variables and use the indicator for nonattainment status as an instrumental variable in an effort to isolate changes in TSPs that are orthogonal to changes in the unobserved determinants of housing prices.

We note that a number of studies have used census tract level data or even house-level price data and focused on local markets (e.g., Ridker and Henning 1967, Harrison and Rubinfeld 1978, and Palmquist 1984). In contrast, the unit of observation in our data is the county. Two practical reasons for the use of these data are that TSPs regulations are enforced at the county-level and census tracts are difficult to match between the 1970 and 1980 Censuses.

The use of this county-level data raises a few issues. First in the absence of arbitrary assumptions about which counties constitute separate markets, it is necessary to assume that there is a national housing market.[18] The benefit of this is that our estimates of MWTP will reflect the preferences of the entire US population, rather than the subpopulation that lives in a particular city or local market. The cost is that we are unable to explore the degree of within-county taste heterogeneity and sorting. If taste heterogeneity and sorting are greater within counties than between counties, as is likely the case, then the subsequent results will understate the individual-level dispersion in MWTP.

Second, the hedonic approach as originally conceived is an individual level model and aggregation to the county-level may induce some biases. For example if the individual relationship is nonlinear, the aggregation will obscure the true relationship. We suspect that the aggregation to the county-level may not be an important source of bias. Notably, our cross-sectional estimates from this county-level data are very similar to the estimates in the previous literature that rely on more disaggregated data and are summarized in Smith and Huang (1995).

Further, the aggregation does not lead to the loss of substantial variation in TSPs. Using the availability of readings from multiple monitors in most counties, we find that only 25% of the total variation in 1970-80 TSPs changes is attributable to within-county variation, with the rest due to between-county variation. Finally, a census tract-level (or individual-house level) analysis introduces inference problems that a county-level analysis avoids because there are substantially fewer monitors than census tracts (or houses). [19]

B. Descriptive Statistics

Figure 1 presents trends from 1969-1990 in average particulates levels across the counties with monitor readings in each year.[20] Air quality improved dramatically over the period, with TSPs levels falling from an average of 85 $\mu g/m^3$ in 1969 to 55 $\mu g/m^3$ in 1990. Most of the overall pollution reduction

---

[18] Given this assumption, it is sensible to also explore whether TSPs affect wages as in Roback (1982). We find no association between TSPs and wages and briefly discuss these results in Section VII.

[19] For example, Harrison and Rubinfeld's (1978) analysis of 506 census tracts relies on only 18 TSPs monitors. As noted by Moulton (1986), the treatment of these correlated observations as independent can lead to incorrect inferences.

[20] These are weighed averages of the county means, with the county's population in 1980 used as weights. The sample consists of 169 counties with a combined population of 84.4 million in 1980. The unweighted figure is qualitatively similar.

occurred in two punctuated periods. While the declines in the 1970s correspond with the implementation of the 1970 CAAA, the remaining improvements occurred during the 1981-82 recession. As heavily polluting manufacturing plants in the Rust Belt permanently closed due to the recession, air quality in these areas improved substantially (Kahn 1997, Chay and Greenstone 2003b). This implies that local economic shocks could drive both declines in TSPs and declines in housing prices. Below, we find that fixed-effects estimates of the HPS may be seriously biased by these shocks.

Table 1 presents summary statistics on the variables that we control for in the subsequent regressions. The means are calculated as the average across the 988 counties with nonmissing data on TSPs concentrations in 1970, 1980, and 1974 or 1975, as well as nonmissing housing price data in 1970 and 1980. These counties form the primary sample, and they account for approximately 80 percent of the U.S. population. All monetary figures are denoted in 1982-84 dollars. During the 1970s the mean of the counties' median housing price increased from roughly \$40,300 to \$53,168, while TSPs declined by 8 $\mu g/m^3$. Per-capita incomes rose by approximately 15%, and unemployment rates were 2.2 percentage points higher at the end of the decade. The increase in educational attainment during this period is also evident. The population density and fraction of people residing in urbanized areas are roughly constant at the beginning and end of the decade.[21]

IV. The CAAAs as a Quasi-Experimental Approach to the Hedonic Identification Problems

In the ideal analysis of individuals' valuation of air quality, air pollution concentrations would be orthogonal to all determinants of housing prices and tastes for clean air. Since this orthogonality condition does not hold, this section describes how we exploit the differences in regulatory intensity introduced by the CAAAs to address the identification problems described in Section I B. The first subsection demonstrates that TSPs nonattainment status is strongly correlated with declines in TSPs concentrations and increases in housing prices. These findings appear robust to exploiting the discreteness of the function that determines TSPs nonattainment status.

[21] Since the definition of the vacancy variables changes over time, it is impossible to include the first difference of these variables in the subsequent regressions. Consequently, the regressions separately control for the 1970 and 1980 levels of these variables.

The second subsection provides theoretical and statistical rationales for using mid-decade TSPs nonattainment status as an instrumental variable, rather than beginning of decade nonattainment status. It also highlights the likely problems with the conventional cross-sectional and fixed effects estimation strategies.

A. TSPs Nonattainment Status and Changes in TSPs Concentrations and Housing Prices

This subsection explores the relationship between TSPs nonattainment status and changes in TSPs concentrations and housing prices. Figure 2A examines the initial impact of the 1970 CAAAs on TSPs concentrations. The counties with continuous monitor readings from 1967-1975 are stratified by their regulatory status in 1972, which is the first year that the CAAAs were in force.[22] The horizontal line at 75 $\mu g/m^3$ is the annual federal standard and the vertical line separates the pre-regulation years (1967 through 1971) from the years that the regulations were enforced. The exact TSPs concentration is reported at each data point.

Before the CAAAs, TSPs concentrations are approximately 35 $\mu g/m^3$ higher in the nonattainment counties. The pre-regulation time-series patterns of the two groups are virtually identical. From 1971 to 1975, the set of 1972 nonattainment counties had a stunning 22-$\mu g/m^3$ reduction in TSPs, while TSPs fell by only 6 $\mu g/m^3$ in attainment counties, continuing their pre-1972 trend. This implies that virtually the entire national decline in TSPs from 1971-75 in Figure 1 is attributable to the regulations.

Figure 2B demonstrates that mid-1970s nonattainment status is also associated with reductions in TSPs concentrations. Here, counties are divided into those that are nonattainment in either or both 1975 and 1976 and those that are attainment in both years. TSPs concentrations for both sets of counties are plotted for the years 1970-1980 for the 414 counties with readings in every year. Average TSPs concentrations decline by approximately 17 $\mu g/m^3$ in both sets of counties between 1970 and 1974. This is surprising since the 1975-6 nonattainment counties are more likely to be nonattainment in 1972, but it implies that, at least as it relates to pre-existing trends, the attainment counties may form a valid

[22] The sample consists of 228 counties with a total population of 89 million in 1970. As the Data Appendix describes, 1972 nonattainment status is determined by 1971 TSPs concentrations.

counterfactual for the nonattainment ones.[23] Specifically, mean reversion and differential trends are not likely sources of bias. Between 1974 and 1980, mean TSPs concentrations declined by 6.3 $\mu g/m^3$ in nonattainment counties and increased by 4.1 $\mu g/m^3$ in attainment counties. Consequently, mid-decade nonattainment status is associated with a roughly 10 $\mu g/m^3$ relative improvement in TSPs.[24] [25]

The structure of the federal regulations lends itself to the application of two validity tests of the relationship between nonattainment status and changes in TSPs concentrations and housing prices. Recall, nonattainment status is a discrete function of the annual geometric mean and second highest daily concentrations of TSPs in the previous year. The first test is a comparison of the outcomes of nonattainment and attainment counties with selection year mean TSPs concentrations "near" 75 $\mu g/m^3$. If the unobservables are similar at this regulatory threshold then a comparison of these nonattainment and attainment counties will control for all omitted factors correlated with TSPs. This test has the features of a quasi-experimental regression-discontinuity design (Cook and Campbell 1979).[26]

The second test compares nonattainment and attainment counties with selection year mean TSPs concentrations less than 75 $\mu g/m^3$. Here, the variation in regulatory status is based on violations of the daily standard. The key assumption is that, conditional on selection year mean TSPs, the assignment of nonattainment status does not depend on the potential outcome, or is "ignorable" (Rubin 1978). This approach has the features of a matching design.

---

[23] 73 of the 265 (117 of the 149) 1975-76 TSPs attainment (nonattainment) counties were TSPs nonattainment in 1972. Thus, over 25 percent of the counties switched their nonattainment status between the beginning and middle of the decade.

[24] The results are qualitatively similar when Figure 2B is based on the 988 counties in our primary sample.

[25] When a figure similar to Figure 2B is constructed for the 1980s, the reduction in TSPs attributable to mid-decade regulations cannot be distinguished from differential responses to the 1981-82 recession. This finding is not surprising given the geographic variation in the effect of the 1981-82 recession (Chay and Greenstone 2003b) and the termination of the TSPs regulatory program in 1987. For these and others reasons, this study focuses solely on the 1970s. Chay and Greenstone (1998) provide a fuller discussion of these issues and present the results from strategies that address the problems in estimating the HPS for the 1980s.

[26] In some contexts, leveraging a discontinuity design may accentuate selection biases if economic agents know about the discontinuity point and change their behavior as a result. Given the wide variety of factors that determine local TSPs concentrations ranging from wind patterns to industrial output, we suspect that counties were unable to engage in non-random sorting near the TSPs regulatory ceiling. Similarly we suspect that during the 1970s, individual homeowners were unaware of the proximity of their county's TSPs concentration to the annual threshold, making it unlikely that they would move as a consequence.

Figure 3 graphically implements these checks on the validity of mid-decade TSPs nonattainment status as an instrument. Separately for the 1975 attainment and nonattainment counties, each panel graphs the bivariate relation between an outcome of interest and the geometric mean of TSPs levels in 1974, the selection year for the 1975 nonattainment designation. The plots come from the estimation of nonparametric regressions that use a uniform kernel density regression smoother. Thus, they represent a moving average of the raw changes across 1974 TSPs levels.[27] The difference in the plots for attainment and nonattainment counties can be interpreted as the impact of 1975 nonattainment status, both for the counties exceeding just the daily concentration threshold (i.e., counties with 1974 concentrations below 75 $\mu g/m^3$) and for the counties exceeding the annual threshold.

Panel A presents the 1970-80 change in mean TSPs by the level of mean TSPs in 1974. First, compare the nonattainment counties with selection year TSPs concentrations just above 75 $\mu g/m^3$ (demarcated by the vertical line in the graph) and attainment counties just below this threshold. The figure reveals that right at the threshold nonattainment counties had an approximately 5 $\mu g/m^3$ larger decline in TSPs than attainment counties over the course of the decade. Further an examination of the counties with selection year TSPs concentration in the 65-85 $\mu g/m^3$ range reveals that the nonattainment counties had anywhere from a 10-14 $\mu g/m^3$ greater reduction in mean TSPs. The size of the TSPs reductions declines for counties with 1974 mean concentrations greater than 90 $\mu g/m^3$. Further, the slight downward slope in the plot for attainment counties is consistent with some reversion in TSPs.

Second, consider the counties with annual mean concentrations below 75 $\mu g/m^3$. Here, the nonattainment counties received this designation for having as few as 2 "bad days". There are 67 such counties. A comparison of these nonattainment counties to the attainment ones with selection year mean TSPs concentrations in the same range suggests that nonattainment status is associated with about a 5-unit greater reduction in mean TSPs over the decade.[28]

[27] The smoothed scatterplots are qualitatively similar for several different choices of bandwidth – e.g., bandwidths that use between 10 to 20 percent of the data to calculate local means.

[28] The finding from Figures 2 and 3A that nonattainment status is associated with reductions in TSPs concentrations contradicts recent claims that the Clean Air Act had no effect on air quality (Goklany 1999).

Panel B plots the conditional change in log-housing values from 1970-80 separately for nonattainment and attainment counties. Both validity checks indicate a striking association between 1975 nonattainment status and greater increases in property values over the decade.[29] They suggest that nonattainment counties had about a 0.02-0.04 log point relative increase in housing prices.

In summary, Figures 2 and 3 document that nonattainment status is strongly associated with reductions in TSPs and increases in housing prices. The discrete differences in TSPs changes and housing price changes that are visible with the regression discontinuity and matching validity tests provide convincing support for a causal interpretation of the effect of 1975 TSPs nonattainment status on these outcomes.

Finally, Figure 3 foreshadows our results on the relationship between housing prices and TSPs. In particular, the strong correspondence between the patterns in Panels A and B suggests that this quasi-experiment may also be detecting a casual relationship between air pollution and property values through the mechanism of regulation. The panels also imply that MWTP may vary with the level of TSPs concentrations. Specifically, the ratio of the nonattainment-attainment difference in housing price increases to the difference in mean TSPs declines is lower in magnitude in dirtier counties (i.e., those with 1974 TSPs concentrations above 75 $\mu g/m^3$). One explanation for this difference is that individuals with strong preferences for clean air systematically sorted to the relatively clean areas. In order to explore this possibility further, the below implements an estimation approach that estimates the average MWTP, while attempting to control for sorting on heterogeneous tastes.

B. Mid-Decade TSPs Nonattainment Status

There are theoretical and statistical reasons that mid-decade (e.g., 1975-6) nonattainment status is a better candidate for an instrumental variable than the beginning of the decade nonattainment designation. First, consider the theoretical rationale. Since annual county-level housing values are

---

[29] It should be noted that there are no nonattainment counties with 1974 geometric mean TSPs concentrations below 40 $\mu g/m^3$. The counties with mean TSPs concentrations below 40 $\mu g/m^3$ have much smaller populations and noticeably different characteristics from the counties with higher TSPs concentrations. Therefore, these counties may not be comparable to the other monitored counties.

unavailable, we rely on 1970 and 1980 Census data. Additionally, Figures 2A and 2B suggest that there is at least a 2-3 year lag before the full effects of TSPs nonattainment status on pollution reductions are realized. A focus on 1972 nonattainment status would leave 5-6 years until the 1980 census, which is likely enough time for individuals to sort themselves into new locations and supply to respond to these air quality changes. Consequently when housing prices are measured in 1980, the composition of people and homes may differ from the composition in 1970. In other words if local housing markets are integrated over this time horizon, it may be invalid to use 1970-80 housing price changes to measure the MWTP for a reduction in TSPs in the early 1970s.[30]

A focus on mid-decade nonattainment status may mitigate this problem of compensatory responses. The intuition is that the changes in air quality due to mid-decade regulation are not evident until the end of the 1970s, which is roughly the same time that the Census Bureau asks about housing values. This may provide an opportunity to observe how TSPs reductions are capitalized without substantial contamination by general equilibrium adjustments in demand and supply. It is evident that the virtually identical "pre-period" change in TSPs concentrations in nonattainment and attainment counties (recall Figure 2B) is a necessary condition for this approach to be valid.

The second reason for our focus on mid-decade TSPs nonattainment status is that this designation is uncorrelated with most observable determinants of housing prices, including economic shocks. This is not true when comparing beginning of decade nonattainment and attainment counties. Further, we find evidence that there is likely to be substantial confounding in the conventional cross-sectional and fixed effects approaches to estimating the HPS. Table 2 presents evidence on these points.

Table 2 shows the association of TSPs levels, TSPs changes, and TSPs nonattainment status with numerous determinants of housing prices. In the first four columns, the sample is our base set of 988 counties. The Column 5 entries are derived from our "regression discontinuity" sample of 475 counties that are near the annual threshold. A county is included in this sample if it meets two criteria: 1) its 1974 geometric mean TSPs concentrations is in the 50-100 $\mu g/m^3$ range, and 2) if its 1974 geometric mean

[30] For example, Blanchard and Katz (1992) find that local housing prices fall in the first five years following a negative local employment shock but rebound fully within about a decade of the shock. This rebound in prices is likely due to the general equilibrium responses of consumers and suppliers of housing.

concentration is below 75 $\mu g/m^3$, it must be designated attainment (i.e., all counties that are nonattainment solely due to the bad day rule are dropped). The column 6 entries are from our bad day sample, which is limited to the 419 counties with 1974 geometric mean concentrations between 50 and 75 $\mu g/m^3$.

The entries in each column are the differences in the means of the variables across two sets of counties and the standard errors of the differences (in parentheses). A '*' indicates that the difference is statistically significant at the 5% level, while '**' indicates significance at the 1% level. If TSPs levels were randomly assigned across counties, one would expect very few significant differences.

Column 1 presents the mean difference in the 1970 values of the covariates between counties with 1970 TSPs concentrations greater and less than the median 1970 county-level TSPs concentration. There are significant differences across the two sets of counties for several key variables, including income per capita, population, population density, urbanization rate, the poverty rate, the fraction of houses that are owner occupied, and the share of government spending on education. It is interesting that mean housing values are higher in the dirtier counties, although this difference is not statistically significant at conventional levels. Although it is not presented here, an analogous examination of the means in 1980 leads to similar conclusions. Overall, these findings suggest that "conventional" cross-sectional estimates may be biased due to incorrect specification of the functional form of the observables variables and/or omitted variables.

Column 2 performs a similar analysis for the 1970-1980 TSPs changes. Here, the entries are the mean difference in the change in the covariates between counties with a change in TSPs that is less (i.e., larger declines) and greater than the median change in TSPs. Reductions in TSPs are highly correlated with economic shocks. The counties with large pollution declines experienced substantially less growth in per-capita income, smaller population growth, a bigger increase in unemployment rates, a larger decline in manufacturing employment, and less new home construction. These entries demonstrate that TSPs concentrations are pro-cyclical and suggest that unless it is possible to perfectly control for the economic cycle, the fixed-effects estimator of the HPS will have a positive bias. For example, the second row shows that housing values grew more in the counties that had a relative <u>increase</u> in TSPs levels.

Column 3 compares the 1970-80 change in the same set of variables in 1970-2 TSPs nonattainment and attainment counties. Here, a county is designated nonattainment if it exceeds the federal standards in any of the years 1970, 1971, or 1972; all other counties are in the attainment category. The nonattainment counties had a smaller increase in per-capita income growth and larger and statistically significant declines in population, manufacturing employment, new home construction, and population density than the attainment counties. We suspect that the population flows reveal a substantial worsening of economic conditions in nonattainment counties. This worsening is likely due to non-neutral impacts of the 1974-75 recession and/or the economic effects of the regulations themselves (e.g., Greenstone 2002). Just as with the fixed effects results, these entries suggest that estimates that rely on 1970-2 TSPs nonattainment status as an instrument may be positively biased due to the confounding of changes in economic activity with the regulation-induced change in TSPs. These findings imply that beginning of decade nonattainment status is not an attractive candidate for an instrumental variable.[31]

Columns 4 repeats this analysis among 1975-6 TSPs nonattainment and attainment counties and finds that the observable determinants of housing prices are better balanced across these counties. Remarkably, the mid-decade nonattainment instrument purges the non-neutral economic shocks apparent above. For example, the differences in the changes in per-capita income, total population, unemployment rates, manufacturing employment, and new home construction among the two sets of counties are all smaller in magnitude than in the other columns and statistically indistinguishable from zero. Also, nonattainment and attainment counties had almost identical changes in urbanization rates during the 1970s, suggesting that differential 'urban sprawl' within counties is not a source of bias.[32] Notably, nonattainment counties had both a greater reduction in TSPs and a greater increase in housing values from 1970-80, foreshadowing our instrumental variable results.

Finally, columns 5 and 6 compare the covariates across 1975 TSPs nonattainment and attainment counties for the regression discontinuity and bad day samples. It is evident that the observable variables

[31] Counties that were 1973-4 TSPs nonattainment also had statistically significant larger declines in population and increases in the poverty rate.

[32] While there are some significant differences between nonattainment and attainment counties in 1970 values of the variables, the hypothesis of equal population densities in 1970 cannot be rejected at conventional levels.

are well balanced by nonattainment status in these columns. In fact, none of the mean differences in column 5 are statistically different from zero and only two of them are in column 6.[33]

Although a direct test of the validity of the exclusion restriction is as always unavailable, it is reassuring that our instrumental variable is largely uncorrelated with observable determinants of housing prices. Overall, the results in this table suggest that using mid-decade nonattainment status as an instrumental variable has some important advantages over "conventional" estimation strategies and also over the use of beginning of decade nonattainment status as an instrumental variable.

Figure 4 provides a graphical overview of the location of the 1975-6 TSPs nonattainment and attainment counties. A county's shading indicates its regulatory status; light gray for attainment, black for nonattainment, and white for the counties without TSPs pollution monitors. The pervasiveness of the regulatory program is evident. For example, 45 of the 51 states had at least one county designated nonattainment. This is important if there are regional or state-specific determinants of the change in housing prices between 1970 and 1980.

In summary, there are several reasons why our approach to estimating the HPS may be attractive. First, TSPs nonattainment status is strongly associated with large differential reductions in particulates levels across counties. The timing and location of the changes provide convincing evidence that the estimated pollution impact may be causal. Second, the nonattainment designation is largely uncorrelated with observable determinants of housing price changes, including economic shocks. In fact, the instrument appears to purge the local demand and supply shocks that contaminate estimates based on 'fixed-effects' analyses. Third, since the regulations are federally mandated, their imposition is presumably orthogonal to underlying economic conditions and the local political process determining the supply of non-market amenities.[34]

---

[33] Further, the differences in the 1970 levels of the variables are also smaller in these restricted samples. This is especially true in the regression discontinuity sample, where the hypothesis of equal means cannot be rejected for all the covariates, except % Urban (p-value=.037).

[34] Scientific evidence provides additional support for the credibility of regulation instruments that depend on pollution levels. Cleveland et al. (1976) and Cleveland and Graedel (1979) document that wind patterns often transport air pollution hundreds of miles and that the ozone concentration of air entering into the New York region in the 1970s often exceeded the federal standards. A region's topographical features can also affect pollution concentrations. Counties located in valleys (e.g., Los Angeles, Phoenix, Denver, the Utah Valley) are prone to weather inversions that lead to prolonged periods of high TSPs concentrations.

V. Econometric Models for the HPS and Average MWTP

Here, we discuss the econometric models used to estimate the hedonic price locus. First, we focus on the constant coefficients version of these models. We then discuss a random coefficients model that allows for self-selection bias arising from taste sorting. We show how this model identifies the average MWTP in the population while providing a simple statistical test of sorting based on preferences for clean air.

A. Estimation of the HPS Gradient

The cross-sectional model predominantly used in the literature is:

(2) $$y_{c70} = X_{c70}'\beta + \theta T_{c70} + \varepsilon_{c70}, \qquad \varepsilon_{c70} = \alpha_c + u_{c70}$$

(3) $$T_{c70} = X_{c70}'\Pi + \eta_{c70}, \qquad \eta_{c70} = \lambda_c + v_{c70},$$

where y_{c70} is the log of the median property value in county c in 1970, X_{c70} is a vector of observed characteristics, T_{c70} is the geometric mean of TSPs across all monitors in the county, and ε_{c70} and η_{c70} are the unobservable determinants of housing prices and TSPs levels, respectively.[35] The coefficient θ is the 'true' effect of TSPs on property values and is interpreted as the average gradient of the HPS. For consistent estimation, the least squares estimator of θ requires $E[\varepsilon_{c70}\eta_{c70}] = 0$. If there are omitted permanent (α_c and λ_c) or transitory (u_{c70} and v_{c70}) factors that covary with both TSPs and housing prices, then the cross-sectional estimator will be biased.

With repeated observations over time, a 'fixed-effects' model implies that first-differencing the data will absorb the county permanent effects, α_c and λ_c. This leads to:

(4) $$y_{c80} - y_{c70} = (X_{c80} - X_{c70})'\beta + \theta(T_{c80} - T_{c70}) + (u_{c80} - u_{c70})$$

(5) $$T_{c80} - T_{c70} = (X_{c80} - X_{c70})'\Pi + (v_{c80} - v_{c70}).$$

[35] For each county, T_{c70} (T_{c80}) is calculated as the average across the county's annual mean TSPs concentrations from 1969-72 (1977-80). Each county's annual mean TSPs concentration is the weighted average of the geometric mean concentration of each monitor in the county, using the number of observations per monitor as weights. Averaging over more than one year reduces the impact of temporary perturbations on our measures of pollution.

For identification, the least squares estimator of θ requires $E[(u_{c80} - u_{c70})(v_{c80} - v_{c70})]=0$. That is, there are no unobserved shocks to pollution levels that covary with unobserved shocks to housing prices.

Suppose there is an instrumental variable (IV), Z_c, that causes changes in TSPs without having a direct effect on housing price changes. One plausible instrument is mid-1970s TSPs regulation, measured by the attainment-nonattainment status of a county. Here, equation (5) becomes:

(6) $T_{c80} - T_{c70} = (X_{c80} - X_{c70})'\Pi_{TX} + Z_{c75}\Pi_{TZ} + (v_{c80} - v_{c70})^{\circ}$, and

(7) $Z_{c75} = 1(T_{c74} > \overline{T}) = 1(v_{c74} > \overline{T} - X_{c74}'\Pi - \lambda_c)$,

where Z_{c75} is the regulatory status of county c in 1975, $1(\bullet)$ is an indicator function equal to one if the enclosed statement is true, and $\overline{T}$ is the maximum concentration of TSPs allowed by the federal regulations.[36] Nonattainment status in 1975 is a discrete function of TSPs concentrations in 1974. In particular, if T^{avg}_{c74} and T^{max}_{c74} are the annual geometric mean and 2nd highest daily TSPs concentrations, respectively, then the actual regulatory instrument used is $1(T^{avg}_{c74} > 75\ \mu g/m^3$ <u>or</u> $T^{max}_{c74} > 260\ \mu g/m^3)$.

An attractive feature of this approach is that the reduced-form relations are policy relevant. In particular, Π_{TZ} from equation (6) measures the change in TSPs concentrations in nonattainment counties relative to attainment ones. In the other reduced-form equation,

(8) $y_{c80} - y_{c70} = (X_{c80} - X_{c70})'\Pi_{yX} + Z_{c75}\Pi_{yZ} + (u_{c80} - u_{c70})^{\circ}$,

Π_{yZ} captures the relative change in log-housing prices. Since the IV estimator, (θ_{IV}), is exactly identified, it is a simple ratio of the two reduced form parameters, that is $\theta_{IV} = \Pi_{yZ}/\Pi_{TZ}$.

Two sufficient conditions for the IV estimator (θ_{IV}) to provide a consistent estimate of the HPS gradient are $\Pi_{TZ} \neq 0$ and $E[v_{c74}(u_{c80} - u_{c70})] = 0$. The first condition clearly holds. The second condition requires that unobserved price shocks from 1970-80 are orthogonal to transitory shocks to 1974 TSPs levels. In the simplest case, the IV estimator is consistent if $E[Z_{c75}(u_{c80} - u_{c70})] = 0$.

We implement the IV estimator in a number of ways. First, we use all of the available data and the mid-decade nonattainment indicator as the instrument to obtain θ_{IV}. We also calculate IV estimates in two other ways that allow for the possibility that $E[v_{c74}(u_{c80} - u_{c70})] \neq 0$ over the entire sample. The first

---

[36] In practice, our preferred instrument equals 1 if a county is nonattainment in 1975 or 1976 and 0 otherwise. In this section, we denote it with Z_{c75} for ease of exposition.

leverages the regression discontinuity (RD) design implicit in the 1(•) function that determines nonattainment status. For example if $[v_{c74}(u_{c80} - u_{c70})]=0$ in the neighborhood of the annual regulatory ceiling (i.e., 75 $\mu g/m^3$), then a comparison of changes in nonattainment and attainment counties in this neighborhood will control for all omitted variables. In the case where this assumption is invalid but the relationship between v_{c74} and $(u_{c80} - u_{c70})$ is sufficiently smooth, then causal inference is possible by including smooth functions of T_{c74} in the vector of covariates.

The second approach exploits the matching design that is feasible due to the $T_{c74}^{max} > 260\ \mu g/m^3$ discontinuity in the regulation selection rule. Here, we obtain θ_{IV} from the sample of counties with "identical" T_{c74}^{avg}. We implement this by limiting the sample to counties with T_{c74}^{avg} between 50 $\mu g/m^3$ and 75 $\mu g/m^3$ and continuing to use Z_{c75} as an instrumental variable. It is even possible to "nonparametrically" control for T_{c74}^{avg} with a series of indicators so that the comparisons between nonattainment and attainment counties are within narrow ranges of this variable. This approach will produce consistent estimates of θ_{IV} if the number of 'bad' days in a county does not independently affect price changes, holding constant the annual mean of TSPs, which seems plausible.

Before proceeding, it is worth noting that in most applications of Rosen's model, the vector of controls, denoted by X, is limited to housing and neighborhood characteristics. Income and other similar variables are generally excluded on the grounds that they are "demand shifters" and are needed to obtain consistent estimates of the MWTP function. However, if individuals believe that there are spillovers, then the presence of wealthy individuals or high levels of economic activity is an amenity and the exclusion restriction is invalid. In our analysis, we are agnostic about which variables belong in the X vector and instead show unadjusted estimates, as well as estimates adjusted for all the variables listed in Table 1. Importantly, our IV estimates are largely insensitive to the choice of control variables.

B. Random Coefficients, Self-Selection, and the Average MWTP

Each point on the HPS provides the marginal consumer's MWTP for a marginal change in TSPs. If individual tastes for clean air are identical, then the average gradient of the HPS, θ, gives the average marginal rate of substitution for all consumers. However, if there is sorting arising from taste dispersion, then θ may differ from the average MWTP in the population. We use a random coefficients model to illustrate this and derive a test for a negative assortive matching equilibrium.

Suppose preferences for air quality can be summarized at the county level. Simplifying notation, define $y_i \equiv y_{c70}$, $\Delta y_{it} \equiv (y_{c80} - y_{c70})$, $\Delta T_{it} \equiv (T_{c80} - T_{c70})$, and $Z_i \equiv Z_{c75}$. Ignoring the observables, the random coefficients version of equation (2) is $y_i = \theta_i T_i + \varepsilon_i$, where θ_i represents heterogeneity in the MWTP across individuals/counties and $E(\theta_i)=\bar{\theta}$ is the average MWTP in the population. Here, the least squares estimator of $\bar{\theta}$ will be biased if either $E(\varepsilon_i T_i) \neq 0$ due to omitted variables or $E(\theta_i T_i) \neq 0$ due to self-selection. If individuals sort across counties based on their tastes for air quality, then $E(\theta_i T_i) > 0$, i.e., individuals with a high valuation for clean air select to counties with low TSPs levels.

If θ_i is stationary over time, then the random coefficients analogs of equations (4) and (6) are:

(9) $\Delta y_{it} = \bar{\theta}\Delta T_{it} + (\theta_i - \bar{\theta})\Delta T_{it} + \Delta u_{it},$

(10) $\Delta T_{it} = \Pi Z_i + \Delta v_{it},$

Suppose that ΔT_{it} is monotonically related to T_i, in that the size of 1970-80 TSPs reductions is (weakly) increasing in 1970 TSPs levels. (Recall, Figure 3A suggests that this is the case, except for the dirtiest counties.) Then θ_i and ΔT_{it} may be correlated through either a nonconstant marginal utility or self-selection due to taste heterogeneity.

In the presence of correlated random coefficients, identification of $\bar{\theta}$ requires stronger assumptions than the orthogonality conditions above. Heckman and Vytlacil (1998) and Wooldridge (1997) specify conditions under which the inclusion of the first-stage residuals, Δv_{it}, as a control variable in the outcome equation will purge both the omitted variables and selection biases. Since this is numerically identical to 2SLS, θ_{2sls} will be consistent. The key condition in each of these papers turns on whether θ_i is a function of Z_i. In our case, Z_i is a function of TSPs concentrations and if individuals with a high valuation for clean air select to counties with low TSPs levels, then $E(\theta_{2sls}) \neq \bar{\theta}$. In this situation, 2SLS may identify the average MWTP for a nonrandom subpopulation.

There is an alternative two-step approach to estimating $\bar{\theta}$ that allows for separate 'control functions' for the omitted variables and self-selection biases. This procedure also provides a simple statistical test for sorting based on tastes.[37] Consider the following assumptions:

<u>A1</u>: $E(\Delta u_{it}|Z_i) = E(\theta_i|Z_i) = 0$

<u>A2</u>: $E(\Delta u_{it}|\Delta T_{it}, Z_i) = \lambda_T \Delta T_{it} + \lambda_Z Z_i$

<u>A3</u>: $E(\theta_i|\Delta T_{it}, Z_i) = \psi_T \Delta T_{it} + \psi_Z Z_i$

A2 and A3 allow the conditional expectations of both Δu_{it} and θ_i to depend linearly on ΔT_{it} and Z_i. When these assumptions are combined with A1, they imply $E(\Delta u_{it}|\Delta T_{it}, Z_i) = \lambda_T \Delta v_{it}$ and $E(\theta_i|\Delta T_{it}, Z_i) = \psi_T \Delta v_{it}$. This results in the regression model:

(11) $\Delta y_{it} = \bar{\theta}\,\Delta T_{it} + \lambda_T \Delta\hat{v}_{it} + \psi_T \Delta T_{it} \cdot \Delta\hat{v}_{it} + \Delta\varepsilon_{it},$

where $\Delta\hat{v}_{it}$, the estimated residuals from the 1st-stage equation, is the potentially endogenous component of TSPs changes.

The estimation of equation (11) has several attractive features. First, under A1-A3, least squares fitting of (11) will produce a consistent estimate of the average MWTP in the population. Second, $\lambda_T = \frac{\mathrm{Cov}(\Delta u_{it}, \Delta v_{it})}{\mathrm{Var}(\Delta v_{it})}$, so it provides a convenient measure of the importance of omitted variables bias in the conventional fixed effects estimator (i.e., equation (4)). The cyclical nature of TSPs concentrations and the results in Table 2 imply that the estimated λ_T will be positive.

Third the coefficient $\psi_T = \frac{\mathrm{Cov}(\theta_i, \Delta v_{it})}{\mathrm{Var}(\Delta v_{it})}$, so it measures the importance of heterogeneity in MWTP. The sign and significance of the estimated ψ_T provide a test of sorting. First, note that homogeneous preferences and marginal utilities that do not increase in air quality imply $\psi_T \leq 0$. Only if there is taste heterogeneity <u>and</u> individuals sort across counties based on this heterogeneity can ψ_T be greater than 0 (i.e., individuals who prefer clean air sort into low pollution counties). As a result, an estimated $\psi_T > 0$ is consistent with negative assortative matching under the weak restriction of non-increasing marginal utilities.

[37] The control function approach outlined here follows from Garen (1984) who examined selectivity bias in the returns to education. See Heckman and Vytlacil (1999), Vytlacil (1999), Manski and Pepper (2000), Blundell and Powell (2000), and Florens, et al (2000) for a discussion of these and related issues.

This test may have important implications for the optimal design of regulatory policy. If tastes are homogeneous, then a diminishing marginal utility implies that the marginal benefit of a pollution reduction is greater in communities with higher pre-regulation TSPs levels. This is consistent with the CAAAs' annual threshold of 75 $\mu g/m^3$. However if there is taste heterogeneity and sorting based on this heterogeneity, then those with greater distaste for pollution will sort to areas with lower TSPs levels. Here, the welfare gain from a TSPs reduction may be greater in communities with lower pollution levels, a possibility that the current design of the CAAAs effectively ignores.

VI. Empirical Estimates of the HPS Gradient and the Average MWTP

Here, we present the estimates of the HPS gradient and the average MWTP from the econometric models discussed above. There are three main findings. First, conventional regression analysis produces unreliable estimates of the HPS gradient. Second, the 1975-6 TSPs nonattainment instrumental variable produces robust estimates that imply individuals place greater value on clean air than previously recognized. Finally, the random coefficients results provide evidence of taste based sorting in equilibrium, but the overall variation in county-level MWTP is not large.

A. 'Conventional' Estimates of the HPS Gradient

Table 3 presents 'conventional' estimates of the capitalization of TSPs into property values from the 1970 and 1980 cross-sections and the 1970-80 first differences. These estimates provide a useful benchmark since they are based on regression specifications typically used in the literature. For the 1970 and 1980 cross-sections, Column 1 gives the unadjusted correlation; Column 2 allows the observables to enter linearly; Column 3 adds cubic polynomials and interactions of the control variables; and Column 4 adds unrestricted region effects for each of the nine Census Bureau divisions so that the identification comes from within region comparisons of counties. These four specifications are used throughout the remainder of the analysis.

For 1970 the unadjusted correlation between housing prices and TSPs has a counterintuitive sign but is statistically insignificant. However, the correlation adjusted for a linear combination of the observables suggests that a 1-unit decline in TSPs leads to an approximately 0.06% increase in housing values. While the implied elasticity is only -0.04, the estimate would be judged statistically significant at conventional levels. Also, it is in the middle of the range of estimates summarized in the Smith and Huang (1995) meta-analysis. This is noteworthy since it is based on a time period and regression specification similar to those used in the bulk of the previous research.

The estimate implies that if Allegheny county, which is in Pittsburgh, reduced its 1970 TSP levels by 50% (a 65-$\mu g/m^3$ reduction), housing prices would increase by only 4% or $1,800 ($1982-84), all else equal. Further, the estimate is reduced when the analysis adjusts for a flexible functional form and interactions for the covariates in column 3 and becomes even smaller and statistically insignificant when a full set of region indicators are included in column 4. Note that the fit of the regressions is quite good.

The 1980 results also bring into question the reliability of cross-sectional analysis. Here, a linear covariate adjustment leads to the perverse result that a 1-unit TSPs reduction is associated with a 0.10% <u>decrease</u> in housing prices. This is particularly disturbing given the estimate's precision and the excellent fit of the regression equation (R^2=0.82). Further, this is the same specification that produced an estimate similar to the "meta-estimate" from the previous literature. Controlling for nonlinearities and interactions in the covariates and unrestricted region effects reduces the magnitude of the estimate but it is still perversely signed.

The third panel of the table contains the 1980-70 first-differenced results. First-differencing the data eliminates the bias in the cross-sectional estimates attributable to permanent differences across counties. However, this approach will be biased if it is not possible to adequately control for shocks that drive both pollution and price changes. In the first column, the unadjusted correlation between changes in housing prices and TSPs has the perverse positive sign and is highly significant. This finding was foreshadowed by the results in column 2 of Table 2. Adjustment for the covariates in columns 2-4 causes the estimate to become economically small and statistically indistinguishable from zero.

These results represent our effort to replicate the previous literature's approach. Overall, with county-level data on almost 1000 counties, the cross-sectional and fixed effects correlation between TSPs and property values is weak and very sensitive to the choice of specification. Based on these conventional estimation procedures, we conclude that either individuals place a small value on air quality or the HPS gradient is plagued by substantial omitted variables bias.

B. Instrumental Variables Estimates of the HPS Gradient

Reduced Form Relations. Table 4 contains the regression results from estimating equations (6) and (8). The regulation variable is an indicator equal to one if the county was nonattainment in either 1975 or 1976 (or both years). Column 1 presents the unadjusted estimates and columns 2-4 present the estimates from the same specifications as in Table 3.

The first panel shows that mid-decade nonattainment is associated with a 9-10 $\mu g/m^3$ (11-12%) reduction in TSPs. This estimate is insensitive to a wide set of controls, including region fixed effects as in column 4. Further, it is highly significant with an F-statistic ranging between 22 and 31 depending on the specification, suggesting that it is the most important (observable) determinant of 1980-70 TSPs changes. Thus, the first-stage impact of regulation is very powerful and appears valid.

The second panel reveals another striking empirical regularity. The TSPs nonattainment variable is associated with a 2-3.5% relative increase in housing values from 1980-70. These estimates are also highly significant. The adjusted estimates are on the low end of this range, but after the linear adjustment in column 2 further controls have little effect on the estimate, even as there is a large improvement in the regression fit (e.g., R^2=0.73 in column 4).

Taken literally, these results imply that the federal TSPs nonattainment designation resulted in important improvements in air quality and property values in these counties. These findings are important in their own right, because they indicate that the Clean Air Act's regulation of TSPs had substantial benefits during the 1970s.

Instrumental Variables Estimates. Table 5 contains the IV estimates of the HPS gradient derived from three different instruments. In the first panel the instrument is the 1975-76 nonattainment indicator, so the reported θ_{IV} is simply the ratio of the reduced form estimates in Table 4. The estimates suggest that a 1-$\mu g/m^3$ reduction in mean TSPs results in a 0.2-0.4% property value increase, which is a -0.20 to -0.35 elasticity. This is roughly 5-8 times larger than the largest cross-sectional estimate in Table 3. Further these estimates are largely insensitive to adjustment, which is not surprising given the findings in Table 2.[38] [39] Thus, concerns about which of the measured variables belong in the HPS equation are unfounded in this setting.

The second panel presents IV estimates when the instrument is 1975 TSPs nonattainment status. This provides the statistical analog to the plots in Figure 3. The parameter estimates are very similar to the estimates in the first panel. Overall, these results also suggest that there is a robust association between changes in TSPs and housing prices.

The third panel presents IV results when the instrument is based on 1970-72 TSPs nonattainment status. As Section IV B details, we suspect that these estimates are biased upwards, but they are presented for completeness. The unadjusted estimate in column 1 has a perverse sign. As we control for more covariates, the sign reverses. In fact, the column 4 estimates implies that a 1-$\mu g/m^3$ reduction in mean TSPs results in a statistically significant increase of 0.07% in property values. This pattern of the coefficients is consistent with our concerns about the validity of the instrument due to non-neutral economic shocks and general equilibrium responses to the regulation-induced changes in air quality documented in Table 2.[40] [41]

---

[38] The IV point estimates are larger when we adjust for the levels of the variables in 1970, rather than the 1980-1970 changes, and when the regressions are weighted by the square root of population. For example with the 1975-76 TSPs nonattainment instrument, the IV estimates (standard errors) are –0.481 (0.183) and –0.400 (0.174) when the analysis adjusts linearly for the 1970 levels (instead of the changes) of the controls (i.e., column 2) and when ln 1970 housing prices is added to this specification, respectively. When the regression is weighted by the square root of the sum of the 1970 and 1980 county-level populations, the same instrument and the columns 1-4 specifications yield estimates (standard errors) of -.576 (.526), -.364 (.204), -.498 (.259), and -.379 (.186). In this context, the standard errors from the weighted equations are all at least twice as large as those from the unweighted ones.

[39] We also explored the consequences of limiting the sample to the 428 counties located in metropolitan areas or SMSAs. The estimates across the four specifications are -.297 (.206), -.417 (.213), -.462 (.300), and -.412 (.269). The standard errors are similar when the variance-covariance matrix allows for correlation within SMSAs, rather than the Eicker-White standard errors reported here. There is little variation in nonattainment status within SMSAs so the inclusion of SMSA fixed effect causes the standard errors to increase dramatically making meaningful inference impossible. Finally when the sample is limited to the counties that are not part of a SMSA, the estimates across the four specifications are -.975 (.513), -.302 (.170), -.222 (.119), and -.135 (.104). In the context of the standard errors, the qualitative findings are similar to those in the SMSA sample.

[40] The same concerns about general equilibrium effects and similar evidence of non-neutral economic shocks apply to an instrument based on 1973-4 TSPs nonattainment status. When this variable is used as an instrument, we also find evidence that is consistent with an upward bias in the estimated effect of TSPs on housing prices.

[41] As footnote 23 detailed, beginning of decade TSPs nonattainment status covaries with 1975-76 TSPs nonattainment status. Due to this confounding and the possibility that 1970-72 nonattainment status may capture unobserved trends in housing prices, it may be appropriate to include the 1970-72 TSPs nonattainment indicator as a control in the equations where the 1975-6 nonattainment indicator is the instrumental variable. When this covariate

Robustness Tests. Table 6 presents the statistical versions of the regression discontinuity and bad day matching tests and is intended as a check of the robustness of the results in Table 5. The first panel reports the regression discontinuity I results. The 181 counties that are nonattainment due to the bad day rule only in either 1975 or 1976 are dropped from the sample and the instrument is 1975-6 TSPs nonattainment status. All specifications include the geometric mean of TSPs in 1974 and 1975 and their squares as controls. These terms control for "smooth" functions of the variables that determine regulatory status, thus the TSPs coefficient is identified from the discontinuous jump in the probability of nonattainment status at the annual threshold.

Once we adjust for covariates as in columns 2 and 3, the estimates indicate that a 1 unit decline in TSPs is associated with a roughly 0.65-0.75% increase in housing prices.[42] These estimates have associated t-statistics that are greater than one, but they would not be judged statistically significant by conventional criteria. The loss of precision is a consequence of trying to separately identify the effect of the instrument from the smooth functions of 1974 and 1975 TSPs concentrations.

The second panel presents the results from our regression discontinuity II approach. Here, 1975 TSPs nonattainment status is the instrument, but the sample is limited to counties with 1974 TSPs concentrations in the 50-100 $\mu g/m^3$ range. We add the further sample restriction that counties that are nonattainment for the bad day rule only are dropped, which leaves a sample of 475 counties. The identifying assumption is that unobservable determinants of the change in housing prices are balanced across the counties above and below the annual threshold in this subsample. The estimates suggest that a 1-$\mu g/m^3$ reduction in mean TSPs results in a 0.1-0.3% increase in housing values, and the adjusted estimate are at the low end of this range.

The results in the third panel are also derived from the 1975 instrument, but the sample is restricted to counties with 1974 TSPs concentrations below the annual threshold of 75 $\mu g/m^3$ and above 50 $\mu g/m^3$. This test exploits the bad day feature of the regulations. The parameter estimates imply that a

is added to our four primary specifications, the estimates are -2.159 (1.927), -.612 (.418), -.673 (.425), and -.466 (.359). The loss of precision is predictable, but the qualitative results are unchanged.

[42] The parameter estimates are modestly larger in magnitude when the squares of 1974 and 1975 TSPs concentrations are dropped from these specifications. Also, note the percentage change is larger than the point estimate because the ln approximation understates the percentage change.

1-$\mu g/m^3$ reduction in mean TSPs results in a 0.50-0.65% increase in house prices. Further, the point estimates are virtually identical when separate dummy variables are included for each 5 unit interval between 50 and 75 so that identification is obtained within these intervals.

Overall, Table 6 supports the evidence of a meaningful relationship between TSPs concentrations and housing prices that was presented in Table 5. This is noteworthy because these tests are derived directly from the structure of the regulations. Thus, they are pre-specified and rule out ex-post rationalizations of unexpected findings. It is relevant though that each of these tests is very demanding of the data, consequently none of the individual parameters would be judged to differ from zero by conventional criteria.

C. Random Coefficients Estimates of the Average MWTP and Evidence on Taste Sorting

If tastes for clean air are homogeneous, then the 2SLS estimates of the HPS gradient in Table 5 are consistent estimates of the average MWTP in the population. However, a comparison of the estimates in the second and third panels of Table 6 that ignores the standard errors indicates that the housing price-TSPs gradient may be steeper at lower concentrations of TSPs. This was also suggested by Figure 3. If there were homogeneous county-level tastes for air quality, this finding would violate the assumption of declining marginal utility. Consequently, we believe that there is at least modest evidence of taste heterogeneity and that the Table 5 estimates may represent the average MWTP for a nonrandom subpopulation.

To examine this issue more formally, we estimate the random coefficients regression model specified in equation (11). The model relaxes the single-index restriction of 2SLS and includes separate control functions for the omitted variables and self-selection biases. As long as A1 and the linear conditional expectations restrictions from A2 and A3 hold, the model will consistently identify the population average MWTP and provide a simple test of county-level taste sorting.

Table 7 presents the results from estimating this model using the TSPs 1975-6 nonattainment variable as the instrument.[43] There are several important findings. First, the estimates of the average MWTP are only slightly higher than the 2SLS estimates in Table 5. Thus, the single control function underlying 2SLS appears to do a reasonable job of absorbing both sources of bias. Second, the estimated coefficient of the first control function, λ_T, is positive and highly significant. This implies that the omitted variables bias in the conventional first-differences estimate is substantial, even after regression adjustment. Third, the selection bias control function also has a positive coefficient estimate ($\psi_T>0$), which is highly significant in the first column. Under the assumptions of the model, this provides direct statistical evidence of nonrandom taste sorting. The fact that the estimated ψ_T is reduced substantially by regression adjustment implies that much of the county-level taste sorting behavior can be explained by observable differences across counties.

These results suggest that negative assortive matching may be a relevant phenomenon in the housing market. However, Table 7 suggests that the overall heterogeneity in county-level MWTP across the population is not large. Further, the relative magnitudes of the λ_T and ψ_T estimates imply that omitted variables bias is a much bigger issue than selectivity bias in estimating the HPS and MWTP. To probe the robustness of the results to the linear conditional expectations restrictions from A2 and A3, we also estimated a model that allows polynomials of both control functions to enter the outcome equation. This led to average MWTP estimates that are identical to those in Table 7, suggesting that the linear 'approximations' in A2 and A3 may be robust.

VII. Interpretation and Welfare Calculations

It appears that mid-decade TSPs regulation is causally related to both air pollution reductions and housing price increases during the 1970s. Here, we use the findings above to develop measures of the economic benefits of the regulations, and, more generally, the WTP for air quality. The estimated gradient of the hedonic price function provides the average MWTP for a 1-unit decline in air pollution.

[43] Some of the regressors in equation (11) are generated from 1st-stage estimation. As a result, we calculated the standard errors of this sequential estimator using the bootstrap with 1,000 replications.

Thus, welfare analysis of the non-marginal TSPs reductions induced by the mid-decade regulations requires identification of the MWTP function

An ad hoc approach to obtaining this function is to make strong assumptions on its shape. A popular assumption, proposed by Freeman (1974), presumes a constant MWTP for clean air.[44] It is straightforward to calculate WTP from this function.[45] The 1975-76 TSPs nonattainment counties had about a 10-unit reduction in mean TSPs, and this decline was capitalized into property values at a rate of about 0.28% per unit.[46] Since the average value of a house in mid-decade nonattainment counties in 1970 was $86,900 (in $2001), mean housing values increased by roughly $2,400 in these counties. The Census PUMS data indicate that there were about 19 million houses in nonattainment counties. This implies that the WTP for the late 1970s TSPs reductions was approximately $45 billion.

The $45 billion figure is also an estimate of the increase in local property values attributable to the mid-1970s TSPs regulation. Over longer time periods, it may be reasonable to expect the value of the tax base to increase by even more as supply responds. Nevertheless, this figure is potentially useful for local governments, since it provides a monetary measure of the local benefits of regulation. By this metric, the CAAAs' mid-decade regulation of TSPs provided substantial benefits.

All of the above calculations follow the previous literature's convention for calculating the value of a unit decline in TSPs, however this convention is flawed. This approach assumes that the entire increase in housing prices is due to the change in air quality in 1980 and that after 1980 the regulation-induced gains in air quality in nonattainment counties would disappear. A more realistic view is that nonattainment status changed individual's expectations about the future path of TSPs concentrations in

[44] This is equivalent to assuming that preferences are homogeneous and linear with respect to air quality.

[45] Setting aside the validity of the constant MWTP assumption, there are some important differences between this measure and an ideal measure of welfare change. First, this measure will tend to overstate the welfare gain relative to one derived from a compensated MWTP function that holds utility constant. Second, we assume that consumers and suppliers have not had time to respond to the TSPs change by moving or changing the supply or quality of the housing stock. However, at the existing HPS, some individuals are likely to be made better off by making these changes. Our measure of the welfare change does not account for this type of compensatory behavior and will thus tend to understate the true welfare gain. See Bartik (1988) for a clear discussion of these issues.

[46] The summary estimate of 0.28% is equal to the weighted average of the estimates from the first panel of Table 5, where the weights are the inverse of the standard errors.

both sets of counties and that the change in housing prices reflects the expected stream of utility associated with this change in expected TSPs concentrations.

We make two alternative calculations that account for expectations and the long-lived nature of housing assets. First, we assume that individuals expected the relative gain in air quality in nonattainment counties would remain constant at 10 $\mu g/m^3$ forever.[47] In the case of the homogeneous MWTP assumption, this implies that a permanent 1 $\mu g/m^3$ decline in TSPs concentrations increases housing prices by roughly 0.28% or $240 ($2001). Second when we add an assumption about the discount rate, it is possible to back out the value of a 1 $\mu g/m^3$ decline in TSPs that lasts a single year. With a 5% discount rate, our results imply that such a decline is worth approximately $12. Of course, these numbers will differ with changes in assumptions about the future path of TSPs or the discount rate.

Finally, we have focused on the effect of TSPs on land values, but, according to the canonical Roback (1982) model, the full implicit price of an amenity is the sum of the land price differential plus the negative of the wage differential. Appendix Table 2 reports the results of fitting our preferred specifications with the mid-decade instrument when the dependent variable is the change in log income. The parameter estimates on the TSPs variable are a precisely estimated zero. Consequently, the above welfare calculations are unchanged by accounting for changes in income.

VIII. Conclusion

This study has used the air pollution reductions induced by the Clean Air Act Amendments to provide new evidence on the capitalization of air quality into housing values. The evidence strongly suggests that TSPs nonattainment status is causally related to both air pollution declines and housing price increases during the 1970s. Using the county-level regulations as instruments, we estimate that a 1-$\mu g/m^3$ reduction in TSPs results in a 0.2-0.4% increase in mean housing values, which is a -0.20 to -0.35 elasticity. Importantly, this estimate is remarkably stable across a variety of specifications. By contrast,

---

[47] If individuals had perfect foresight, then this might not be a bad assumption. In particular, of the set of counties with continuous monitoring from 1970 through 1990, the mid-1970s TSPs nonattainment counties only experienced a relative improvement in TSPs of 4 $\mu g/m^3$ between 1980 and 1990.

estimates from cross-sectional and fixed-effects approaches are very sensitive to specification and sometimes are perversely signed.

Using a random coefficients econometric model, we find that the IV approach provides robust estimates of the average MWTP for clean air. The evidence also suggests that the marginal benefit of a TSPs reduction may be lower in communities with higher pollution levels. This finding is consistent with self-selection across locations due to taste heterogeneity. However, the overall variation in county-level MWTP is not large, and the results imply that omitted variables bias is a much bigger issue than selectivity bias in estimating the HPS and MWTP. Welfare calculations suggest that the mid-1970s TSPs nonattainment designation provided a $45 billion aggregate gain for homeowners in those counties.

This study has a few lessons for future research. First, it reveals that markets can be used to value environmental amenities and that it is possible to successfully estimate the HPS and MWTP in certain contexts. However, it is important to cross-validate our findings in other situations in which the hedonic method is an appropriate tool for analysis. Second, it demonstrates that quasi-experimental approaches can be used to estimate parameters derived from economic models. Third, the ultimate promise of hedonic theory is that it provides a framework to obtain MWTP functions. These functions are of tremendous practical import, because they can be used to calculate the welfare effects of non-marginal changes in goods for which explicit markets are missing and forecast the consequences of alternative policies. Future research should integrate the credible estimation of the HPS with strategies to estimate MWTP functions.

DATA APPENDIX

Determining Attainment/Nonattainment Status at the County Level

The ability to accurately determine the EPA's assignment of counties to attainment/nonattainment status for TSPs is crucial for implementing this paper's quasi-experiment. In the 1972-1977 period, the EPA did not publicly release the names of the counties that were designated nonattainment. To learn the identity of these counties, we contacted the EPA but were informed that records from that period "no longer exist." However, the readings from the air pollution monitoring system were used by the EPA and the states to determine which counties were in violation of the federal air quality standards. Consequently, for the years 1972-77, we use our pollution data to replicate the EPA's selection rule. Counties with monitor readings exceeding the NAAQS for TSPs were assigned nonattainment status; all other counties were designated attainment.

Beginning in 1978, the *Code of Federal Regulations* (Title 40, Part 80) published annually the identity of all nonattainment counties. We collected these annual county-level designations for each of the 3,063 U.S. counties. There is a close correspondence between our "constructed" measure of 1978 TSPs nonattainment status with the actual designations. This suggests that our constructed nonattainment designations are likely to be a good approximation to the counties that the EPA treated as nonattainment in the earlier part of the 1970s.

The Siting of TSPs Monitors and the "Reliability" of the TSPs Pollution Data

Central to the credibility of the analysis is that the pollution concentration readings used accurately reflect the "true" air quality faced by individuals. Since readings from the TSPs monitors are used to determine nonattainment status, it is possible that states or counties strategically placed the monitors to fabricate the appearance of low (or improving) pollution concentrations. To explore the likelihood of this, we examined the CFR and found that the Amendments contain very precise criteria that govern the siting of a monitor.[48] In particular, the legislation forbids states from siting a monitor in a location that does not meet one of the scientific criteria outlined for monitors.[49]

Moreover, the Amendments provided the EPA with a number of enforcement tools to ensure that the states complied with the criteria for siting a monitor. First, the part of the CFR that lists the criteria for monitor placements is incorporated into the SIPs. Since the SIPs are both federal and state law, the EPA can sue states for violating federal law. Second, the usual process for siting is that the states propose a monitor network, and the EPA's district office either approves it or suggests alterations. The federal EPA can also review and reject the siting program, resulting in two layers of oversight. Third, the district offices often require photographs of sites to verify a monitor's placement. Fourth, it is illegal to move many of the monitors. For the monitors that can be moved, the relocation can only be done to better meet the scientific criteria outlined in the CFR. Finally, the district offices are cognizant of which states do not put resources into their siting programs. One district officer said that in these situations they are willing to "play dictator."[50]

---

[48] The substance of this discussion results from the Code of Federal Regulations (CFR) 1995, title 40, part 58 and a conversation with Manny Aquilania and Bob Palorino of the EPA's District 9 Regional Office. Using a recent CFR is not a problem, because the hierarchical control over monitor placement specified in the 1995 CFRs is consistent with previous monitor siting guidelines.

[49] These criteria require that the monitors be placed so that they determine: the highest concentration expected in the area, the representative concentrations in areas of high population density, the impact on ambient pollution levels of significant fixed and mobile categories, and the general background concentration level due to geographic factors. Moreover, the CFR specifically requires that the monitors be a minimum distance from stationary sources of pollution. Using the Landview CD-ROM to examine maps of counties giving the location of pollution monitors, the location of stationary pollution sources, and the location and demographics of the population confirmed the above.

[50] The county-level measures of mean TSP pollution levels used in the analysis are based on averaging the annual geometric mean reading of every monitor in the county over 4 years. Consequently, any idiosyncratic shocks to pollution levels in a county in a short time span will not pose any problems.

Matching the 1980 Census PUMS to the TSPs Pollution and Regulation Data

From the 5-percent PUMS sample of the 1980 Census, we extracted all heads of households who resided in the United States in 1980. This results in a micro data set containing 4.023 million household heads; 2.91 million men and 1.12 million women. The household heads locations of residence in 1980 and 1975 are derived from questions on the state and county group (PUMA) of residence in 1980 and 1975. About 31 thousand of the over 4 million household heads living in the U.S. in 1980 resided outside of the 50 states in 1975.

We merged this Census extract to the TSPs pollution and regulation data. Based on the 1980 County Group Equivalency File for the 1980 Census, we wrote code for three different schemes for matching the PUMA/County Group identifier provided in the PUMS to the FIPS county codes. The schemes matched PUMAs to counties, which accounted for 100%, at least 75%, and at least 50% of the populations in the PUMAs.[51] We then matched the Census data to the TSPs regulation data by the FIPS county identifiers. The results in Table 8 are insensitive to the choice of matching algorithms and to the mid-decade regulation variable used to stratify the sample.

Variables from the 1972 and 1983 *County and City Data Books*

The following are the list of variables taken from the 1972 and 1983 *County and City Data Books* (*CCDB*) and used in the housing value regressions. Most of the information comes from the 1970 and 1980 *Censuses of Population and Housing.* The crime data comes from the U.S. Federal Bureau of Investigation; the medical data comes from the American Hospital Association and the American Medical Association; the spending and tax variables come from the *Census of Governments*. See "Source Notes and Explanations" in the *CCDB* for more detailed explanations of the variables and their sources. We start with the variables used in the 1980 analysis from the 1983 CCDB.

outcome variable

log-median value of owner occupied housing units in 1980
(deflated to $1982-84 by the total shelter component of the CPI)

economic conditions variables

per-capita money income in 1979
civilian labor force (aged 16 or older) unemployment rate
% of employment in manufacturing in 1980

demographic and socioeconomic variables

population per square mile in 1980
% of population white in 1980
% of population female in 1980
% of population aged 65 and over in 1980
% of population over 25 with at least a high school degree in 1980
% of population over 25 with at least a college degree in 1980
% of population in urban area
% of families below the poverty level in 1979

housing variables

---

[51] The matches are comprehensive. The total U.S. population in 1980 was 226,545,805. For the match requiring that a county account for 100% of the PUMA population, 603 PUMAs accounting for 369 counties with a population of 143,462,851 could be matched. For the match in which the county accounts for at least 50% of the PUMA population, the numbers were 831 PUMAs, 561 counties, and 178,872,025 people. It appears that the PUMAs/counties that can be matched are relatively large.

% of year round housing built in last 10 years
% of year round housing built 10-20 years ago
% of year round housing built before 1939
% of occupied housing units lacking complete plumbing in 1980
% of housing units vacant in 1980
% of housing units owner occupied in 1980

neighborhood variables
crime rate per 100,000 in 1981
all serious crimes known to police per 100,000 in 1981
property crimes per 100,000 in 1981
physicians per 100,000 in 1980
hospital beds per 100,000 in 1980

spending and tax variables
per-capita government revenue in 1977
per-capita total taxes in 1977
per-capita property taxes in 1977
per-capita general expenditures in 1977
% of spending on education in 1977
% of spending for police protection in 1977
% of spending on public welfare in 1977
% of spending on health in 1977
% of spending on highways in 1977

For 1970 the following variables were unavailable:
% of year round housing built in last 10 years
% of year round housing built 10-20 years ago
% of year round housing built before 1939
crime rate per 100,000
all serious crimes known to police per 100,000
property crimes per 100,000
physicians per 100,000
hospital beds per 100,000
per-capita total taxes
% of spending for police protection

For the 1980-70 First-Differences and Instrumental Variables Regressions
"First differences" in all of the variables that are in both the 1972 and 1983 *CCDB*'s are included as control variables.

REFERENCES

Bartik, Timothy J. 1987. "The Estimation of Demand Parameters in Hedonic Price Models." Journal of Political Economy, 95: 81-88.

Bartik, Timothy J. 1988. "Measuring the Benefits of Amenity Improvements in Hedonic Price Models." Land Economics, 64: 172-83.

Black, Dan A. and Thomas J. Kneisner. 2003. "On the Measurement of Job Risk in Hedonic Wage Models." Journal of Risk and Uncertainty, 27(3): 205-20.

Blanchard, Olivier Jean and Lawrence F. Katz. 1992. "Regional Evolutions." Brookings Papers on Economic Activity, 1-75.

Blundell, Richard and James Powell. 2000. "Endogeneity in Nonparametric and Semiparametric Regression Models." Mimeograph, UC-Berkeley.

Brown, Charles. 1980. "Equalizing Differences in the Labor Market." Quarterly Journal of Economics, 94: 113-134.

Brown, James N. and Harvey S. Rosen. 1982. "On the Estimation of Structural Hedonic Price Models." Econometrica, 50: 765-68.

Chay, Kenneth Y., Carlos Dobkin, and Michael Greenstone. 2003. "The Clean Air Act of 1970 and Adult Mortality." Journal of Risk and Uncertainty, 27(3): 279-300.

Chay, Kenneth Y. and Michael Greenstone. 1998. "Does Air Quality Matter? Evidence from the Housing Market." NBER Working Paper No. 6826.

Chay, Kenneth Y. and Michael Greenstone. 2001. "Does Air Quality Matter? Evidence from the Housing Market." Center for Labor Economics Working Paper No. 33.

Chay, Kenneth Y. and Michael Greenstone. 2003a. "Air Quality, Infant Mortality, and the Clean Air Act of 1970." NBER Working Paper No. 10053.

Chay, Kenneth Y. and Michael Greenstone. 2003b. "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession." Quarterly Journal of Economics, CXVIII: 1121-1167.

Cleveland, William S. and T.E. Graedel. 1979. "Photochemical Air Pollution in the Northeast United States." Science, 204: 1273-78.

Cleveland, William S., B. Kleinder, J. E. McRae, and J. L. Warner. 1976. "Photochemical Air Pollution: Transport from the New York City Area into Connecticut and Massachusetts." Science, 191: 179-81.

Cohen, Mark A. 1998. "Monitoring and Enforcement of Environmental Policy." Unpublished Paper, Vanderbilt University, Owen Graduate School of Management.

Cook, Thomas D. and Donald T. Campbell. 1979. Quasi-Experimentation: Design and Analysis Issues for Field Settings. Boston, MA: Houghton Mifflin.

Cropper, Maureen L., Leland B. Deck, and Kenneth E. McConnell. 1988. "On the Choice of Functional Form for Hedonic Price Functions." Review of Economics and Statistics, 70: 668-75.

Deacon, Robert T., David S. Brookshire, Anthony C. Fisher, Allen V. Kneese, Charles D. Kolstad, David Scrogin, V. Kerry Smith, Michael Ward, and James Wilen. 1998. "Research Trends and Opportunities in Environmental and Natural Resource Economics." Environmental and Resource Economics, 11: 383-97.

Dockery, Douglas W., et al. 1993. "An Association Between Air Pollution and Mortality in Six U.S. Cities." The New England Journal of Medicine, 329: 1753-9.

Ekeland, Ivar, James J. Heckman, and Lars Nesheim. 2004. "Identification and Estimation of Hedonic Models." Journal of Political Economy, forthcoming.

Epple, Dennis. 1987. "Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products." Journal of Political Economy, 95: 59-80.

Epple, Dennis and Holger Sieg. 1999. "Estimating Equilibrium Models of Local Jurisdictions." Journal of Political Economy, 107: 645-81.

Florens Jean-Pierre, James Heckman, Costas Meghir, Edward Vytlacil. 2000. "Instrumental Variables, Local Instrumental Variables and Control Functions." Manuscript, University of Toulouse.

Freeman, A. Myrick III. 1974. "On Estimating Air Pollution Control Benefits from Land Value Studies." Journal of Environmental Economics and Management, 1(1): 74-83.

Freeman, A. Myrick III. 1993. The Measurement of Environmental and Resource Values: Theory and Methods. Washington, D.C.: Resources for the Future.

Garen, John. 1984. "The Returns to Schooling: A Selectivity Bias Approach with a Continuous Choice Variable." Econometrica, 52: 1199-1218.

Goklany, Indur. 1999. Clearing the Air: The Real Story of the War on Air Pollution, Cato Institute: Washington, DC.

Graves, Phil, James C. Murdoch, Mark A. Thayer, and Don Waldman. 1988. "The Robustness of Hedonic Price Estimation: Urban Air Quality." Land Economics, 64: 220-33.

Greenstone, Michael. 2002. "The Impacts of Environmental Regulations on Industrial Activity: Evidence from the 1970 and 1977 Clean Air Act Amendments and the Census of Manufactures." Journal of Political Economy, 110: 1175-1219.

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

Halvorsen, Robert and Henry O. Pollakowski. 1981. "Choice of Functional Form for Hedonic Price Equations." Journal of Urban Economics, 10: 37-49.

Harrison, David Jr. and Daniel L. Rubinfeld. 1978. "Hedonic Housing Prices and the Demand for Clean Air." Journal of Environmental Economics and Management, 5: 81-102.

Heckman, James and Edward Vytlacil. 1998. "Instrumental Variables Methods for the Correlated Random Coefficient Model." Journal of Human Resources, 33: 974-87.

Heckman, James and Edward Vytlacil. 1999. "Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects." Proceedings of the National Academy of Science, 96: 4730-34.

Heckman, James, Rosa Matzkin, Lars Nesheim. 2002. "Non-parametric Estimation of Nonadditive Hedonic Models." Mimeograph.

Heckman, James, Rosa Matzkin, Lars Nesheim. 2003. "Simulation and Estimation of Nonadditive Hedonic Models." NBER Working Paper 9895.

Henderson, J. Vernon. 1996. "Effect of Air Quality Regulation." American Economic Review, 86: 789-813.

Kahn, Matthew E. 1997. "The Silver Lining of the Rust Belt Manufacturing Decline." Journal of Urban Economics, 46: 360-76.

Lave, Lester B. and Gilbert S. Omenn. 1981. Clearing the Air: Reforming the Clean Air Act. Brookings Institution: Washington, DC.

Liroff, Richard A. 1986. Reforming Air Pollution Regulations: The Toil and Trouble of EPA's Bubble. The Conservation Foundation: Washington, D.C.

Manski, Charles F. and John V. Pepper. 2000. "Monotone Instrumental Variables: With and Application to the Returns to Schocling." Econometrica, 68: 997-1010.

Moulton, Brent R. 1986. "Random Group Effects and the Precision of Regression Estimates." Journal of Econometrics, 32: 385-97.

Nadeau, Louis W. 1997. "EPA Effectiveness at Reducing the Duration of Plant-Level Noncompliance." Journal of Environmental Economics and Management, 34: 54-78.

Palmquist, Raymond B. 1984. "Estimating the Demand for the Characteristics of Housing." Review of Economics and Statistics, 66: 394-404.

Palmquist, Raymond B. 1991. "Hedonic Methods." Measuring the Demand for Environmental Improvement, John B. Braden and Charles D. Kolstad, eds., Amsterdam: Elsevier.

Ransom, Michael R. and C. Arden Pope III. 1995. "External Health Costs of a Steel Mill." Contemporary Economic Policy. 8: 86-97.

Ridker, Ronald G. 1967. Economic Costs of Air Pollution: Studies in Measurement. New York: Praeger.

Ridker, Ronald G. and John A. Henning. 1967. "The Determinants of Residential Property Values with Special Reference to Air Pollution." Review of Economics and Statistics, 49: 246-57.

Roback, Jennifer. 1982. "Wages, Rents, and the Quality of Life." Journal of Political Economy, 90: 1257-78.

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

Rosen, Sherwin. 1986. "The Theory of Equalizing Differences." Chapter 12 in Handbook of Labor Economics, Orley Ashenfelter and Richard Layard, eds., Amsterdam: North-Holland.

Rubin, Donald B. 1978. "Bayesian Inference for Causal Effects: The Role of Randomization." The Annals of Statistics, 6: 34-58.

Sieg, Holger, V. Kerry Smith, H. Spencer Banzhaf, and Randy Walsh. 2000. "Estimating the General Equilibrium Benefits of Large Policy Changes: The Clean Air Act Revisited." NBER Working Paper No. 7744.

Small, Kenneth A. 1975. "Air Pollution and Property Values: A Further Comment." Review of Economics and Statistics, 57: 105-7.

Smith, Robert S. 1979. "Compensating Wage Differentials and Public Policy: A Review." Industrial and Labor Relations Review, 32: 339-52.

Smith, V. Kerry. 1996. "Pricing What is Priceless: A Status Report on Non-Market Valuation of Environmental Resources." Mimeograph, Duke University Department of Economics, August.

Smith, V. Kerry and Ju-Chin Huang. 1995. "Can Markets Value Air Quality? A Meta-Analysis of Hedonic Property Value Models." Journal of Political Economy, 103: 209-27.

Vesilind, P. Aarne, J. Jeffrey Peirce, and Ruth Weiner. 1988. Environmental Engineering, Butterworth-Heinemann: Boston.

Vytlacil, Edward. 2000. "Semiparametric Identification of the Average Treatment Effect in Nonseparable Models." Mimeograph, University of Chicago.

Wooldridge, Jeffrey M. "On Two Stage Least Squares Estimation of the Average Treatment Effect in a Random Coefficient Model." Economics Letters, 56: 129-33.

Table 1: Summary Statistics, 1970 and 1980

	1970	1980
Mean Housing Value	40,290	53,166
Mean TSPs	64.1	56.3
Economic Condition Variables		
Income per Capita ($82-84)	7,122	8,186
Total Population	161,889,646	177,192,574
Unemployment Rate	0.046	0.068
% Employment in Manufacturing	0.249	0.226
Demographic and Socioeconomic Variables		
Population Density	608	585
% >= High School Graduate	0.504	0.646
% >= College Graduate	0.097	0.147
% Urban	0.576	0.593
% Poverty	0.124	0.097
% White	0.901	0.877
% Senior Citizens	0.100	0.113
Housing Variables		
% of Houses Built in Last 10 Years	-----	0.285
% of Houses Built 10-20 Years Ago	-----	0.187
% Overall Vacancy Rate	-----	0.078
% Vacancy Rate Owners Units	0.014	-----
% Vacancy Rate Renter's Units	0.077	-----
% Owner Occupied	0.676	0.620
% of Houses Built Before 1939	-----	0.267
% of Houses without Plumbing (*100)	0.003	0.028
Tax and Expenditure Variables		
Per Capita Government Revenue	747	1098
Per Capita Total Taxes	-----	422
Per Capita Property Taxes	170	354
Per Capital General Expenditures	768	1072
% of Spending on Education	0.548	0.509
% of Spending on Highways	0.091	0.070
% of Spending on Welfare	0.046	0.037
% of Spending on Health	0.048	0.067
% of Spending on Police	-----	0.043
Neighborhood Variables		
Crime Rate per 100,000	-----	4619
Physicians per 100,000	-----	125
Hospital Beds per 100,000	-----	642

Notes: Calculations are based on the 988 counties with data on TSPs concentrations in 1970, 1980, and 1974 or 1975. The housing and overall CPI series are used to deflate all monetary entries to $1982-4. The TSPs data is derived from the EPA's network of pollution monitors. The 1970 (1980) mean TSPs concentration is the average across all counties' mean TSPs concentration from 1969-1972 (1977-80). Each county's annual mean TSPs concentration is calculated as the weighted average of the geometric mean concentrations of each monitor in the county, using the number of observations per monitor as weights. The county-level mean across multiple years (e.g., 1969-1972) is the average of the annual means. The other entries are derived from the 1972 and 1983 County and City Data Books. An entry of '-----' means that the variable was not collected in the relevant year. See the Data Appendix for more details.

Table 2: Sample Means by TSPs Categories

	Cross-Section 1970 (1)	1st Difference 1980 -1970 (2)	TSPs Nonattain in 1970, 1971 or 1972 (3)	TSPs Nonattain in 1975 or 1976 (4)	TSPs in RD
Total (Nonattainment) Counties	988 (-----)	988 (-----)	608 (380)	708 (280)	35
Housing Value	1092 (918)	-3237** (713)	-517 (710)	2609** (795)	200
Mean TSPs	39.2** (1.2)	-30.9** (1.0)	-19.6** (1.3)	-10** (1.5)	-12.
Economic Condition Variables					
Income per Capita ($82-84)	378** (95)	-160** (41)	82* (42)	49 (46)	4
Total Population (% Change)	142,016** (24,279)	-.057** (.013)	-.046** (.014)	-.001 (.015)	.00
Unemployment Rate	-.001 (.001)	.005** (.001)	.002 (.001)	.000 (.001)	.00
% Employment in Manufacturing	.010 (.008)	-.012** (.003)	-.008** (.003)	-.000 (.003)	-.00
Demographic and Socioeconomic Variables					
Population Density	602** (193)	-66.9** (24.8)	-100.5** (25.4)	-18.0 (27.7)	1.0
% Urban	.141** (.017)	-.005 (.005)	-.009 (.005)	-.001 (.006)	-.00
% Poverty	-.012** (.005)	.011** (.002)	.015** (.002)	.014** (.003)	.00
% White	.012 (.008)	-.007* (.003)	-.022** (.003)	-.020** (.003)	-.00
Housing Stock Variables					
% of Houses Built in Last 10 Years		-.034** (.007)	-.025** (.007)	-.006 (.008)	-.00
% Owner Occupied	-.013* (0.006)	.008* (.004)	.013** (.004)	.008* (.004)	.00
% Houses No Plumbing (*100)	-.0005** (.0001)	-.006** (.002)	-.007** (.002)	-.007** (.002)	.00
Tax and Expenditure Variables					
Per Capita Government Revenue	23.8 (24.7)	77.3* (34..2)	60.2 (35.2)	44.6 (38.1)	10.
Per Capita Property Taxes	8.5 (11.7)	26.0** (10.0)	7.2 (9.9)	-1.1 (10.7)	-1.
% of Spending on Education	-.030** (.008)	-.006 (.006)	-.009 (.006)	.012 (.007)	.00

Notes: See the Notes to Table 1. The entries in each column are the differences in the means of the determinants of housing price
and the standard errors of the differences (in parentheses). Column (1) presents the mean difference in the 1970 values of the cov
1970 TSPs concentrations greater and less than the median 1970 county-level TSPs concentration, respectively. Column (2) repo
for 1980-70 TSPs changes. Here, the entries are the mean difference in the change in the covariates between counties with a chan
greater than the median change in TSPs. The entries in columns (3) and (4) are the mean difference of the 1980-1970 change in th
and 1975-6 nonattainment and attainment counties, respectively. The columns (5) and (6) entries compare 1980-1970 changes for
attainment counties. In these columns the samples are restricted to allow for an examination of the assumptions underlying the "R
"Bad Day" tests as they relate to observable variables. See the text for more details. * and ** indicate significance at the 5% and

Table 3 Cross-Sectional and First-Difference Estimates of the Effect of TSPs Pollution on Log-Housing Values (estimated standard errors in parentheses)

	(1)	(2)	(3)	(4)
1970 Cross Section				
Mean TSPs	0.032	-0.062	-0.040	-0.024
(1/100)	(0.038)	(0.018)	(0.017)	(0.017)
R-Squared	0.00	0.79	0.84	0.85
Sample Size	988	987	987	987
1980 Cross Section				
Mean TSPs	0.093	0.096	0.076	0.027
(1/100)	(0.066)	(0.031)	(0.030)	(0.028)
R-Squared	0.00	0.82	0.89	0.89
Sample Size	988	984	984	984
1980-1970 (1st Differences)				
Mean TSPs	0.102	0.024	0.004	-0.006
(1/100)	(0.032)	(0.020)	(0.016)	(0.014)
R-Squared	0.02	0.55	0.65	0.73
Sample Size	988	983	983	983
County Data Book Covariates	N	Y	Y	Y
Flexible Form of County Covs	N	N	Y	Y
Region Fixed Effects	N	N	N	Y

Notes: See notes to Tables 1 and 2. For 1970 and 1980, the Mean TSPs variable is the 1969-72 and 1977-80 average of the annual geometric mean concentrations, respectively. See the Data Appendix for a full list of the control variables. The flexible functional form includes quadratics, cubics, and interactions of the variables as controls. The mean of the ln of 1970 housing prices is 10.55. The means of the dependent variables in the second and third panels are 10.82, and 0.27, respectively. Standard errors are estimated using the Eicker-White formula to correct for heteroskedasticity.

Table 4: Estimates of the Impact of Mid-Decade TSPs Nonattainment on 1970-1980 Changes in TSPs Pollution and Log-Housing Values (estimated standard errors in parentheses)

	(1)	(2)	(3)	(4)
Mean TSPs Changes				
TSPs Nonattain in 1975 or 1976	-9.96	-10.41	-9.57	-9.40
	(1.78)	(1.90)	(1.94)	(2.02)
F-stat. Regulation (numerator dof)	31.3 (1)	29.9 (1)	24.4 (1)	21.5 (1)
F-stat Census Vars (numerator dof)	-----	3.2 (26)	7.8 (86)	7.7 (94)
R-Squared	0.04	0.10	0.19	0.20
Log-Housing Changes				
TSPs Nonattain in 1975 or 1976	0.036	0.022	0.026	0.019
	(0.012)	(0.009)	(0.008)	(0.008)
F-stat. Regulation (numerator dof)	8.5 (1)	6.2 (1)	9.3 (1)	6.4 (1)
F-stat Census Vars (numerator dof)	-----	40.2 (26)	19.2 (84)	26.4 (92)
R-Squared	0.01	0.56	0.66	0.73
County Data Book Covariates	N	Y	Y	Y
Flexible Form of County Covs	N	N	Y	Y
Region Fixed Effects	N	N	N	Y
Sample Size	988	983	983	983

Notes: See notes to previous tables. In the first panel the dependent variable is the difference between the 1977-80 and 1969-72 averages of mean TSPs concentrations. The mean is -7.82 $\mu g/m^3$. In the second panel the dependent variable is the difference between 1980 and 1970 log-housing values and its mean is 0.27. Standard errors are estimated using the Eicker-White formula to correct for heteroskedasticity.

Table 5: Instrumental Variables Estimates of the Effect of 1970-80 Changes in TSPs Pollution on Changes in Log-Housing Values (estimated standard errors in parentheses)

	(1)	(2)	(3)	(4)
TSPs Nonattain in 1975 or 1976				
Mean TSPs	-0.362	-0.213	-0.266	-0.202
(1/100)	(0.152)	(0.096)	(0.104)	(0.090)
Sample Size	988	983	983	983
TSPs Nonattain in 1975				
Mean TSPs	-0.350	-0.204	-0.228	-0.129
(1/100)	(0.150)	(0.099)	(0.102)	(0.084)
Sample Size	975	968	968	968
TSPs Nonattain in 1970, 1971 or 1972				
Mean TSPs	0.072	-0.032	-0.050	-0.073
(1/100)	(0.058)	(0.042)	(0.041)	(0.035)
Sample Size	988	983	983	983
County Data Book Covariates	N	Y	Y	Y
Flexible Form of County Covs	N	N	Y	Y
Region Fixed Effects	N	N	N	Y

Notes: See notes to previous tables. The coefficients are estimated using two-stage least squares. The first row of the three panels indicates which instrument is used. From panel 1 to 3, the instruments are an indicator equal to one if the county was nonattainment for TSPs in either 1975 or 1976, an indicator equal to one if the county was nonattainment for TSPs in 1975, and an indicator that equals one if the county was nonattainment for TSPs in either 1970, 1971 or 1972, respectively. Standard errors are estimated using the Eicker-White formula to correct for heteroskedasticity.

Table 6: Robustness of the Instrumental Variables Estimates of the Effect of 1970-80 Changes in TSPs Pollution on Changes in Log-Housing Values (estimated standard errors in parentheses)

	(1)	(2)	(3)
Regression Discontinuity I TSPs Nonattain in 1975 or 1976			
Mean TSPs	-0.931	-0.500	-0.570
(1/100)	(0.870)	(0.439)	(0.508)
Sample Size	807	802	802
Regression Discontinuity II TSPs Nonattain in 1975			
Mean TSPs	-0.285	-0.133	-0.122
(1/100)	(0.165)	(0.101)	(0.119)
Sample Size	475	472	472
Bad Day/Matching TSPs Nonattain in 1975			
Mean TSPs	-0.498	-0.486	-0.394
(1/100)	(0.788)	(0.580)	(0.516)
Sample Size	419	416	416
Census Covariates	N	Y	Y
Flexible Form of Census Covs	N	N	Y

Notes: See notes to previous tables. The coefficients are estimated using two-stage least squares. In the first panel, counties that are TSPs nonattainment for the "Bad-Day" rule only in either 1975 or 1976 are dropped from the sample. The instrumental variable is an indicator equal to one if the county was nonattainment for TSPs in either 1975 or 1976. All specifications include the 1974 and 1975 county-level geometric means of TSPs and their squares in order to control for "smooth functions" of the selection variables. In the second and third panels, the instrumental variable is an indicator equal to one if the county was nonattainment for TSPs in 1975. In the second panel, the sample is limited to counties with 1974 county-level geometric means of TSPs between 50 and 100 $\mu g/m^3$. Further, counties that are TSPs nonattainment for the violating the "Bad-Day" rule in 1975 but are below the annual threshold are dropped from the sample. In the third panel, the sample is restricted to counties with 1974 county-level geometric means of TSPs between 50 and 75 $\mu g/m^3$. Standard errors are estimated using the Eicker-White formula to correct for heteroskedasticity.

Table 7: Control Function Estimates of the Capitalization of 1970-80 Changes in TSPs Pollution, with Correction for Selectivity Bias Due to Random Coefficients
(estimated standard errors in parentheses)

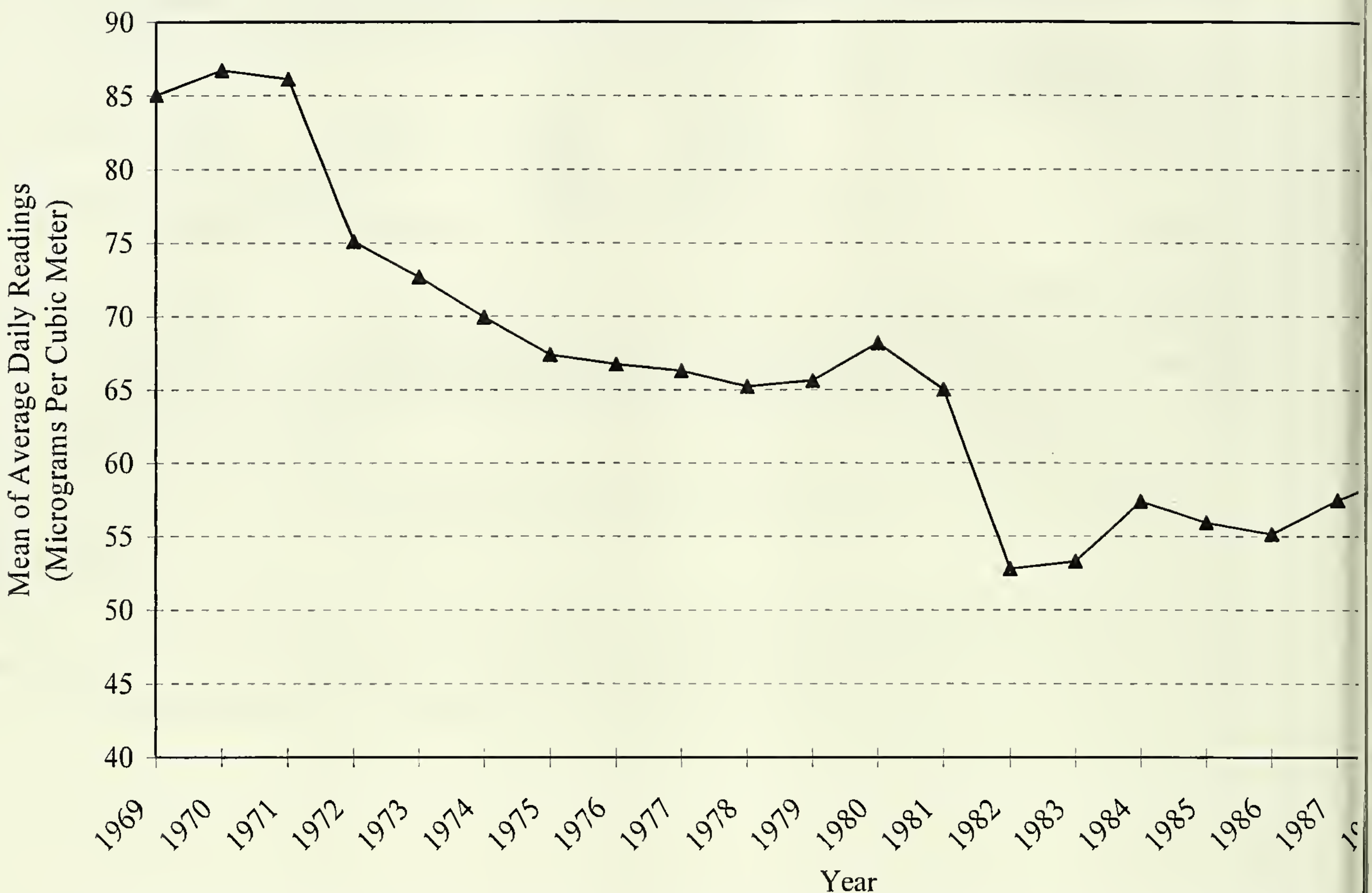
	(1)	(2)	(3)	(4)
Mean TSPs	-0.320	-0.196	-0.256	-0.200
(1/100)	(0.157)	(0.101)	(0.110)	(0.099)
v_i (1st-stage residual)	0.500	0.256	0.289	0.205
(1/100)	(0.164)	(0.103)	(0.112)	(0.100)
v_i* Mean TSPs	0.116	0.049	0.032	0.007
(1/10,000)	(0.043)	(0.026)	(0.023)	(0.021)
Sample Size	988	983	983	983
Census Covariates	N	Y	Y	Y
Flexible Form of Census Covs	N	N	Y	Y
Region Fixed Effects	N	N	N	Y

Notes: The standard errors are calculated based on 1,000 bootstrap replications of the sequential estimator. See text for details on the selectivity bias correction when the endogenous variable is continuous. Estimates are insensitive to including polynomials of the arguments of the two control functions.

Appendix Table 1: Regulation, Sources, Control Technologies, and Health Effects of Total Suspended Particulates (TSPs) Pollution

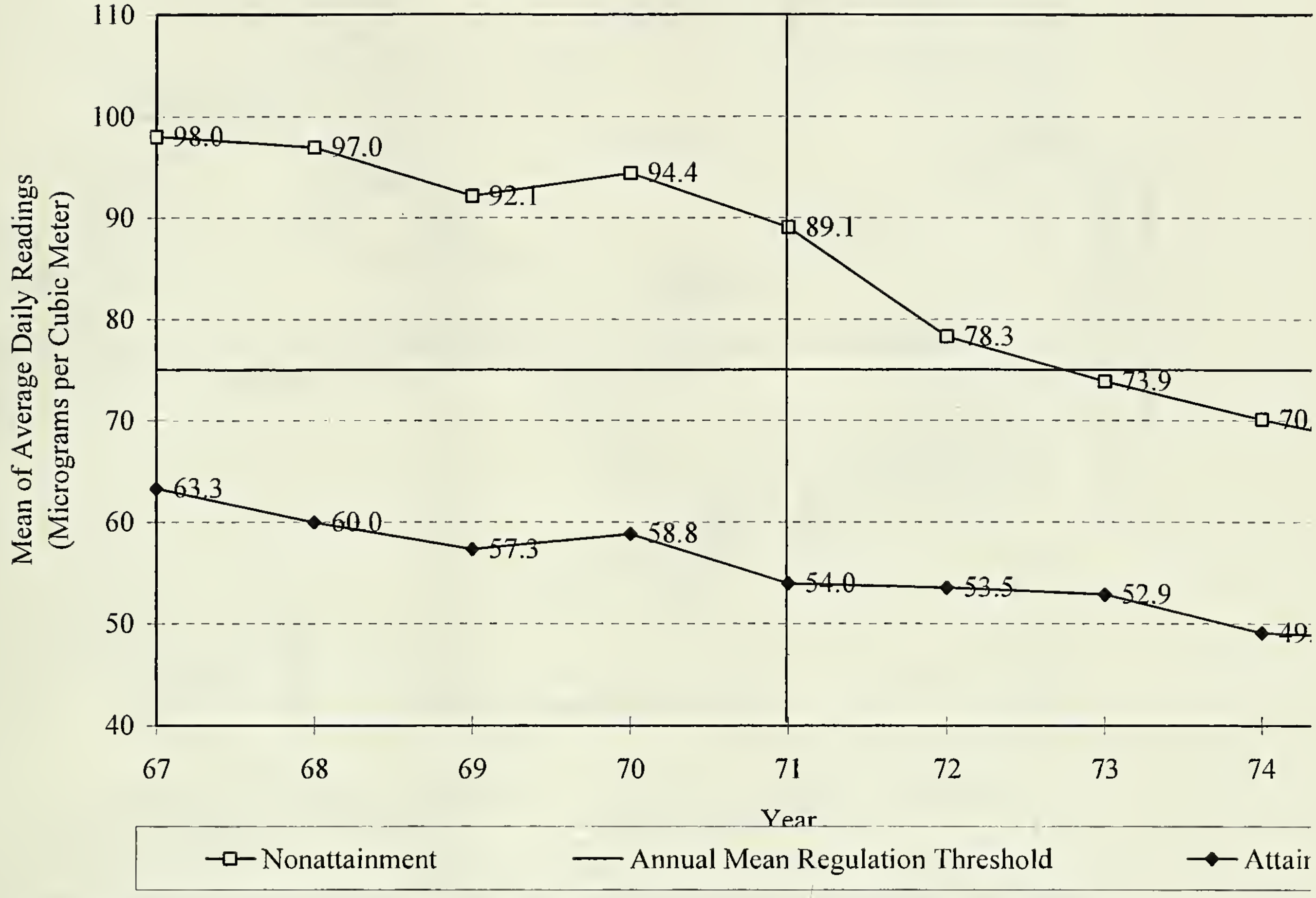
National Ambient Air Quality Standards Maximum Allowable Concentration (Primary Standard):	
Annual Geometric Mean (never to be exceeded)	75 Micrograms per Cubic Meter
Maximum 24 Hour Concentration (not to be exceeded more than once a year)	260 Micrograms per Cubic Meter
Sources Industrial Processes (e.g., Pulp and Paper; Stone, Clay, Glass, and Concrete Products; Iron and Steel), Smelters, Automobiles, Burning Industrial Fuels, Woodsmoke, Dust from Paved and Unpaved Roads, Construction, and Agricultural Ground Breaking.	
Techniques to Control Emissions The control of TSPs is frequently accomplished by directing the polluted air through a "bag" filter, which captures the pollutants, or a wet "scrubber" that increases the mass of the particulates, causing their separation from the "clean" air (Vesilind, et al. 1988).	
Health Effects TSPs can affect breathing and respiratory systems, causing increased respiratory disease and lung damage. Children, the elderly and people suffering from heart or lung disease (e.g., asthma) are especially at risk. Recent research has linked particulates pollution to increased mortality rates (Dockery, et al. 1993; Ransom and Pope 1995; Chay and Greenstone 2003a & 2003b).	

Figure 1: National Trends in Total Suspended Particulates Pollution, 1969-90



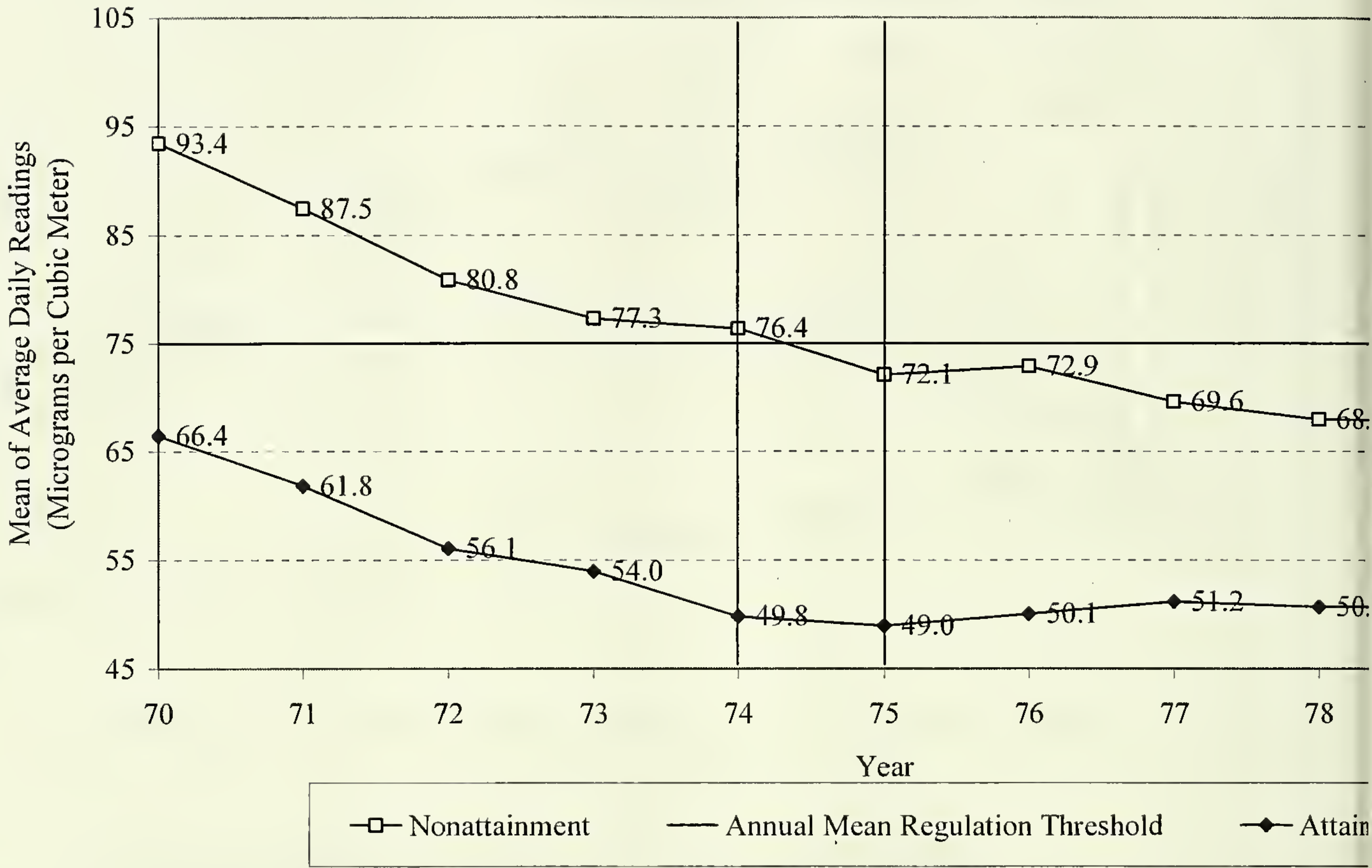
Note: The data points are derived from the 169 counties that are continuously monitored in this period. Tl a total population of approximately 84.4 million in 1980. The annual county means were calculated as the of the monitor-specific geometric means, where the weight is the number of monitor observations. The ye average is calculated as the weighted average of the county-specific means, where the weight is the 1980 p

Figure 2a: Trends in Total Suspended Particulates Pollution from 1967-1975, by 1972 Attainment



Notes: The data points are derived from the 228 counties that were continuously monitored in this period. Th counties had a 1970 population of approximately 25.8 million people, while about 63.4 million people lived ir nonattainment counties in the same year. Each data point is the unweighted mean across all counties in the rel category.

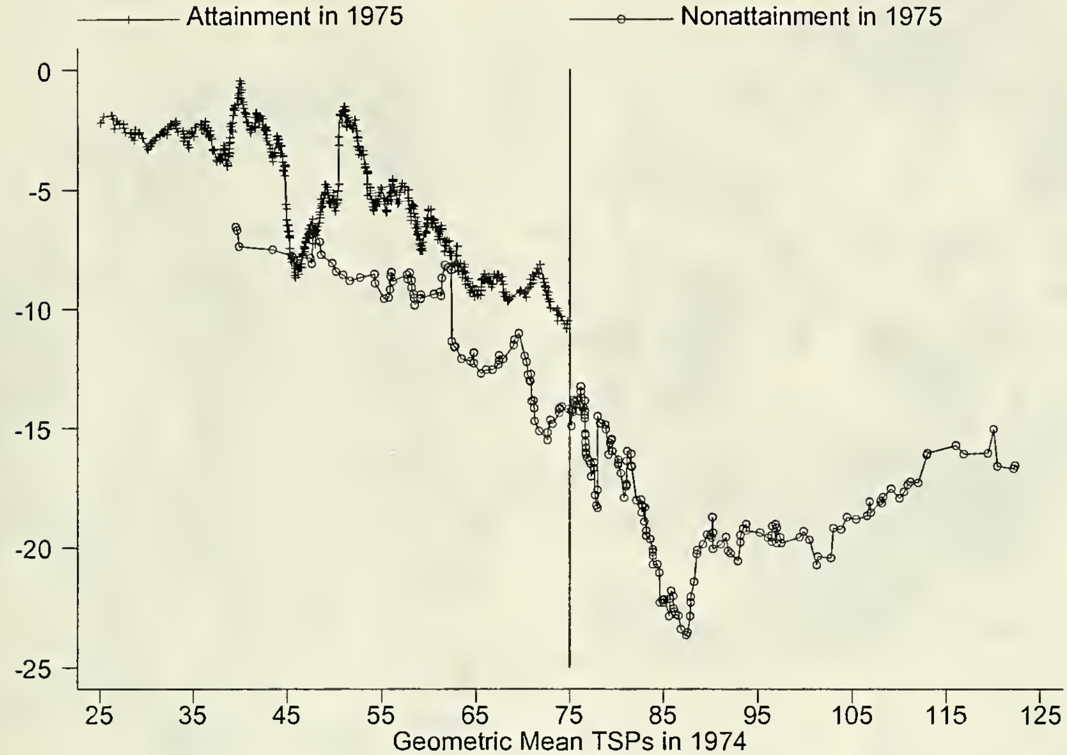
Figure 2b: Total Suspended Particulates Pollution Trends from 1970-80, by 1975-76 Nonattainm



Notes: The data points are derived from the 414 counties that were continuously monitored in this period. counties had a 1970 population of approximately 56.1 million people, while roughly 67.5 million people li nonattainment counties in the same year.

Figure 3: Change in Mean TSPs and Log-Housing Values by 1975 Nonattainment Status,
By the Geometric Mean of TSPs in 1974

A. 1970-1980 Change in Mean TSPs by 1975 Nonattainment Status



B. 1970-1980 Change in Log-Housing Values by 1975 Nonattainment Status

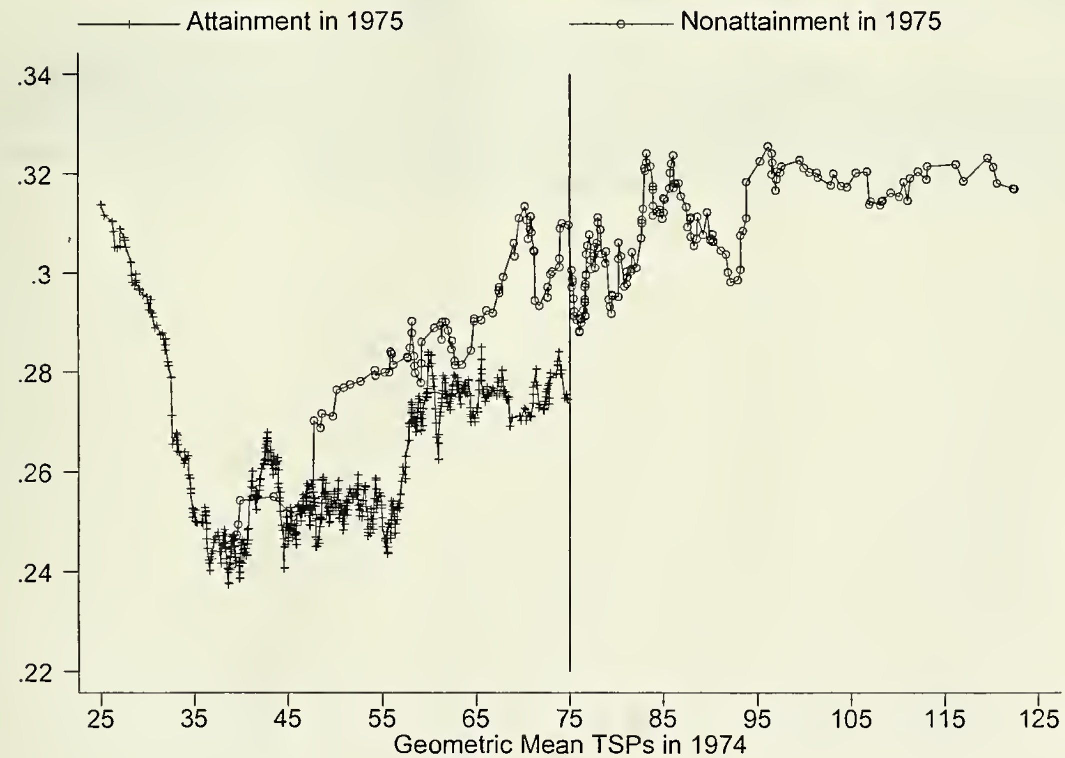
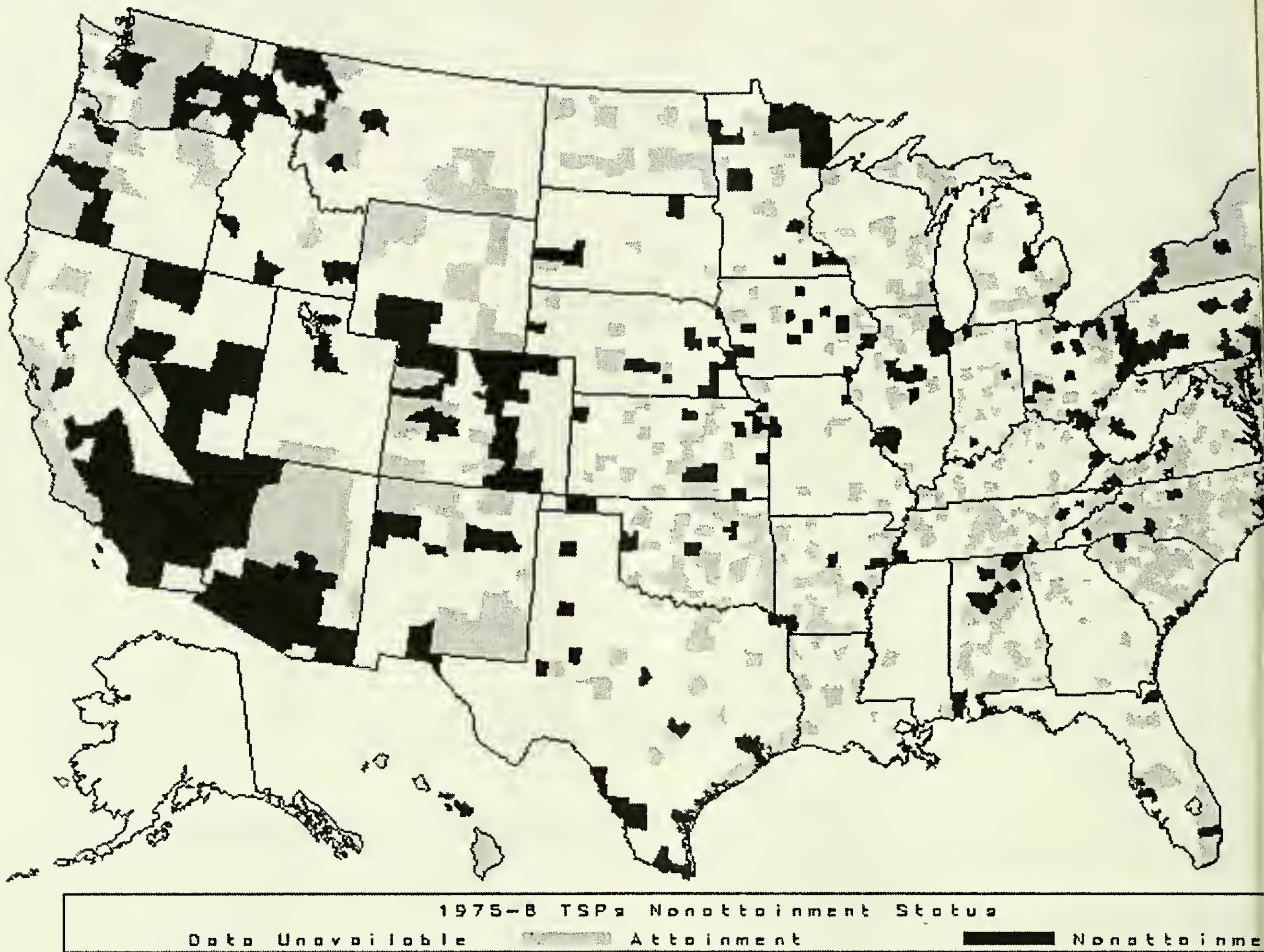


Figure 4: Incidence of 1975-6 TSPs Nonattainment Status



Notes: In our primary sample of 988 counties, there are 280 nonattainment and 708 attainment counties. They are pic
respectively. The 2,169 counties without complete data are depicted in white. The nonattainment designations are de
"Air Quality Subsystem Database." See the Data Appendix for further details.

Date Due
