The (Non)Separability of Air Quality: Evidence from Millions of Households Across the United States

Christos Makridis, Stanford University*

October 23, 2015

PRELIMINARY AND INCOMPLETE, do not distribute without permission.

Abstract

The costs and benefits of environmental policy depend crucially on the assumed micro-elasticities between market and non-market goods. In their absence, general equilibrium models have assumed environmental amenities are perfect substitutes with market goods, such as consumption and leisure, producing qualitatively different welfare assessments of environmental policy under even a narrow range of micro-elasticities. I estimate these elasticities using over 40 million observations from Census micro-data, together with weather and air quality measures at the county-level, between 2000-2014, finding that the elasticities between air quality and consumption, housing, and leisure are 7.14, .54, and .2, respectively. These estimates are identified from county-industry-specific deviations in air quality from the county averages after conditioning on shocks common to all counties within a state. Under simulated counterfactual distributions for 2010, these elasticities imply that the Clean Air Act Amendments had very large negative effects.

*Stanford University, Working Paper. Email: cmakridi@stanford.edu; LinkedIn: www.linkedin.com/in/christosmakridis/. PRELIMINARY AND INCOMPLETE, do not cite. I am grateful to detailed discussions with Don Fullerton, Han Hong, Nicolai Kuminoff, Kerry Smith, and Constantine Yannelis, as well as Spencer Banzhaf, Lawrence Goulder, Wayne Gray, Charles Kolstad, Mar Reguant, and seminar participants at Arizona State University, the Heartland Conference, Stanford University (Applied Microeconomics Lunch and SEEPAC). Partially funded by the NSF Graduate Research Fellowship.

1. Introduction

Pollution in the United States has been cut in half over the past forty years, despite continued growth in per capita consumption and leisure (Figure 1), driven by a battery of factors, ranging from structural transformation in the energy sector (Makridis, 2014) to more stringent regulation (Shapiro and Walker, 2014). While the declines in pollution have unambiguously led to a higher quality of life (Costa and Kahn, 2003; Chay and Greenstone, 2005, 2003; Currie and Neidell, 2005), the *net* welfare effects of both the observed decline in pollution and future environmental policy interventions hinge on quantitative estimates on the micro-elasticities that govern household's substitution between market and non-market goods. And yet, all general equilibrium models used in evaluating environmental policies impose the restrictive assumption that non-market and market goods are perfect substitutes with each other. Mathematically, this takes the form of additive separability between market goods, such as consumption (c) and leisure (l), and the non-market good, like air quality (S): $u(c,l) + \Phi(S)$ where u and Φ are both increasing and weakly concave.[1]

While the assumption is inconsistent with the data, current models (almost without exception) have deferred to it in the absence of quantitative estimates of behavioral elasticities between market and non-market goods.[2] The assumption has the interpretation that changes in the demand for an externality occur independently of changes in the demand for private goods (Davis and Whinston, 1962). Although the assumption was initially attractive because Diamond and Mirrlees (1973) argued it was necessary to rule out unstable equilibria, (Sandmo, 1980) demonstrated that unstable equilibria can emerge even without income effects.[3] The primary contributions of this paper

[1] Additive separability implies that there is an affine transformation of the non-market good. Since households take the non-market good as given in the competitive equilibrium, the externality drops out of the first-order conditions under additive separability; $u(c,l)$ does not depend on S. Additive separability differs from weak complementarity, i.e., $U(g(c,l),S)$, which would imply that S and c or l are average substitutes. The imposition of additive separability is less restrictive within a separate literature in environmental economics involving hedonic pricing, travel cost, and equilibrium sorting models (Kuminoff et al., 2013).

[2] There is an emerging body of empirical evidence, e.g., the effects of pollution on defensive investments (Greenstone et al., 2013), infant mortality (Chay and Greenstone, 2003; Currie and Neidell, 2005), labor productivity (Graff Zivin and Neidell, 2012; Hanna and Oliva, 2015), and health (Moretti and Neidell, 2009; Neidell, 2007; Schlenker and Walker, 2012) and human capital (Bharadwaj et al., 2014; Sanders, 2012; Neidell and Graff Zivin, 2013). See Currie et al. (2014) for a review of the literature on the link between health and pollution. See Guojun (2013) for a regression discontinuity approach from the 2008 Beijing Olympic games (their requirement to meet air quality standards) that found that a 10% decrease in mean PM10 concentrations decreased monthly cardiovascular mortality by 13.6% and Van Hee and Pope (2012) for a randomized medical experiment.

[3] Diamond and Mirrlees (1973) were concerned with ruling out instances these so-called anomalies—instances where corrective taxes could induce more of the externality—by deriving a narrow class of preferences. Sandmo (1980) found that these "anomalies" can occur even without income effects (e.g., even when utility is linear in income). Both Sandmo (1975) and Kopczuk (2003) caution the assumption of additive separability. For example Sandmo (1975) (p. 92-93) acknowledges that marginal damages are not independent of income and relative price effects under externality targeting in general equilibrium. Similarly, Kopczuk (2003) (p. 84) says "this conclusion [the generality of the principle of targeting] may break down when issues involving tax avoidance, evasion, and administrative cost are introduced. The approach also ignores general equilibrium considerations." The reason why additive separability between, for example, consumption and leisure is much less restrictive arises from the fact that

are to develop an analytical framework for characterizing substitution between market and non-market goods, credibly identify these structural elasticities, and conduct a simple welfare analysis to illustrate how these elasticities can be introduced into general equilibrium models.

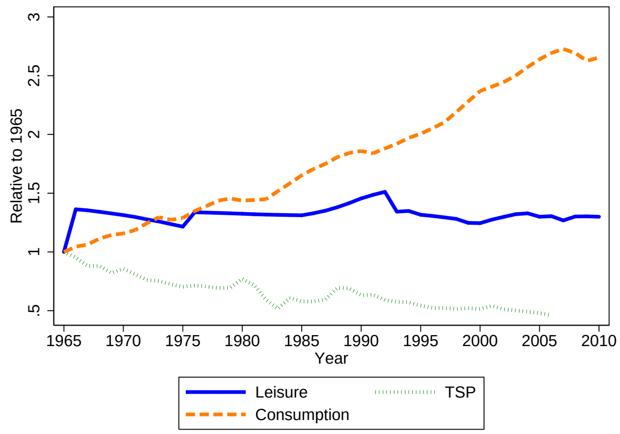

Figure 1: Trends in Consumption, Leisure, and Total Suspended Particulates, 1965 Normalized
Notes.–Source: EPA, OECD, ATUS, World Bank. The left panel plots the historical trend of particulate matter of 10 micrograms per cubic meter or less (PM10) using data initially collected from Smith (2012). The right panel plots the historical trends of average leisure and per capita consumption (obtained from the World Bank in 2005 constant dollars). Leisure is measured according to the first definition in Aguiar and Hurst (2007) (activities providing direct utility), including: measure 1 terms and time in sleeping, eating, personal care, and time eating out. Leisure micro-data comes from the IPUMS historical time use project and observations are weighted according to their recommended value based on data quality and national representativeness.

The quantity and quality of non-market goods, like air quality, affects individuals' consumption of market goods by influencing their locational choice decisions (Banzhaf and Walsh, 2008). Unfortunately, there is no empirical on the elasticity of substitution between environmental amenities and private goods and services. Unlike willingness to pay estimates, these elasticities govern household's dynamic behavioral responses to policy through *cross-substitution* among private goods. In fact, $1.7 of the $2 trillion estimated net benefits in the Environmental Protection Agency's (EPA, p. 3, 2011) Second Prospective evaluation of the Clean Air Act Amendments are driven by a failure to model household's behavioral responses and relative preferences for environmental amenities over private goods/services.[4] Establishing credible estimates of parameters governing behavioral responses is a prerequisite to further analysis, especially given the passage of the 2014 Clean Power Plan.[5] Policy interventions that affect the provision of non-market goods will necessarily affect household's demand for market goods since they change the relationship between marginal willingness to pay (MWTP) and marginal cost for other market goods—that is, if costs of environmental compliance rise, and subsequently affect the relative price of consumption, it is now relatively

market clearing conditions induce prices on market goods. Smith and Carbone (2008) for a more detailed discussion.

[4]Similarly, in their prospective study of "The Benefits and Costs of the Clean Air Act 1990-2010: EPA Report to Congress", 90% of the estimated $110 billion in benefits were driven by reductions in mortality.

[5]http://www2.epa.gov/carbon-pollution-standards/fact-sheet-clean-power-plan-benefits#benefits

more expensive to consume market goods over non-market goods, like air quality. Households can only choose their desired quantity of non-market goods imperfectly through decisions over where to work and live—that is, by selecting into a geographical location with specified amenities—and the elasticities between market and non-market goods are precisely the parameters the pin down household's behavioral responses to changes in the provision of amenities. Ignoring the interaction between market and non-market goods can lead to gross underestimates of the excess burden of environmental policy in general equilibrium.[6]

In order to recover preferences for air quality, I build the most comprehensive database to date with over 42 million records of individual consumption, leisure, demographic attributes, together with county-level information on air quality, pollution, and weather between 2000-2014.[7] My modeling framework builds upon the hedonics and labor market sorting literature (Roback, 1982; Rosen, 1974) and the quasi-experimental literature (Greenstone and Gayer, 2009) by including nonseparabilities between market and non-market goods and deriving equilibrium conditions over household's locational choice and consumption-labor-housing decisions that can be taken directly to the data. While this period experiences a great deal of within-county variation in air quality due to the impact of the Great Recession on production, there are two major sources of endogeneity to confront. First, adverse local demand shocks increase households' leisure availability (Aguiar et al., 2013) and decrease their consumption (Mian et al., 2013) activities. Second, household's locational sorting is driven by their preference for local amenities (e.g., air quality) (Kuminoff et al., 2013; Banzhaf and Walsh, 2008; Epple, 1987). Motivated by insights from the quasi-natural experimental literature (Greenstone and Gayer, 2009), I instrument for leisure using interactions between temperature and age brackets (Graff Zivin and Neidell, 2014), for consumption using interactions between electricity expenditures and the age of the house, for housing services using interactions between the housing supply elasticity (Saiz, 2010) and county leave-one-out averages of the property tax rate, and for air quality (or pollution) using wind speeds. My benchmark results exploit county-industry-specific deviations in consumption, leisure, and air quality from their county averages after adjusting for common shocks across all counties within a state. In addition to

[6]Using Goulder and Williams III (2003) as a benchmark, Carbone and Smith (2008) found that even a small 5% environmental tax could understate the excess burden by nearly 100% when leisure and air quality are complements and overstate it by 50% when they are substitutes. The severity of the bias depends on how far away reality (the nonseparable case) is from the assumption of additive separability. Berry et al. (2014) emphasizes that complementarities may easily arise between a non-market good and unobserved variables, which in this case are other rationed private goods/services.

[7]Since I am not estimating the effects of pollution on health, my annual frequency suffices. To put my sample selection in perspective, Chay and Greenstone (2005, 2003) exploit sharp changes in TSP between 1980-1982; Currie and Neidell (2005) use 1989-2000; Moretti and Neidell (2009) use 1993-2000 in Los Angeles; Neidell (2007) uses 1989-1997 in Los Angeles; Graff Zivin and Neidell (2012) use 2009 and 2010 in Central Valley of California (orange pickers).

a number of robustness checks, I remarkably obtain similar elasticity estimates from reduced-form regressions that project the market good (e.g., consumption) on air quality, instrumented for using wind speeds.

My paper enriches three main veins of research at the intersection of macro-environmental economics. First, I produce a comprehensive database of county and individual -level outcomes from publicly available Census micro-data that can be used for an array of future micro-to-macro and empirical exercises. In order to produce measures of consumption and leisure, I draw on auxiliary micro-data from the American Time Use Survey (ATUS) and Consumption Expenditure Survey (CES). Building on prior literature in labor economics that imputed consumption using the Panel Study of Income Dynamics (see Blundell et al. (2008) and Orazio and Pistaferri (2014)), I use a highly flexible sieve estimator to estimate elasticities between the observed determinants of non-durables consumption and leisure and actual consumption and leisure that vary by family size, year, and educational attainment. I conduct a variety of exercises to show that these predictions are reliable and representative of individuals' likely consumption and leisure bundles. An advantage of the Census data is its large sample size that I use to uncover sources of heterogeneity and state-dependence.[8]

Second, although my analytical framework is disciplined according to structural sorting and hedonic models (see Kuminoff et al. (2013) for a survey), I provide a new identification strategy for recovering causal relationships between market and non-market goods by exploiting quasi-natural variation in climate and weather variables. quasi-natural methods have risen in popularity because they offer solutions to classic omitted variables bias problems (Greenstone, 2004; Greenstone et al., 2013; Greenstone and Gayer, 2009; Chay and Greenstone, 2005, 2003; Graff Zivin and Neidell, 2014; Hanna and Oliva, 2015; Neidell, 2007). The advantage of combining these methods is that structural parameters can be consistently estimated and readily introduced into a suite of macroeconomic models, which is important since the magnitude of these estimates both qualitatively and quantitatively affects welfare analysis (Carbone and Smith, 2008, 2013; Klaiber and Smith, 2012).[9]

Third, I show how these elasticities can be used for general equilibrium analysis on the welfare effects of environmental policy, which is important since environmental policy necessarily affects relative prices and households' cross-substitution patterns between market and non-market goods. Using my estimated elasticities, I find that the CAAA provided \$70 billion in benefits—much lower than the \$2 trillion that the estimated. Generally, my results are also related with the health

[8]Heterogeneity and state-dependence are becoming increasingly important in macroeconomic (Guvenen, 2012; Guvenen and Smith, 2014; Heathcote et al., 2010b, 2014) and health (Finkelstein et al., 2013) literatures.

[9]For example, the elasticity of substitution between consumption and environmental amenities determines whether environmental degradation is decreasing or increasing in wealth (Shibayama and Fraser, 2014)—a literature on the Environmental Kuznets Curve dating back to Grossman and Krueger (1995).

effects of pollution (Chay and Greenstone, 2003; Currie and Neidell, 2005; Currie et al., 2014) and theoretical literature on nonseparability in the context of environmental amenities (Maler, 1974; Bockstael and McConnell, 2007; Flores and Graves, 2008).

Section 2 presents a structural model for disciplining the elasticity estimates and mapping them into macroeconomic models. Section 3 describes the data sources and presents both the methodology and results involving the imputation of non-durables consumption and leisure. Section 4 provides motivating evidence over additional identification problems in the canonical hedonic model. Section 5 estimates the structural model and subjects the results to a battery of robustness checks. Section 6 implements a welfare analysis of the benefits/costs of the rise in air quality over the Great Recession by applying my estimated elasticities in a calibrated model with simulated counterfactual distributions for consumption, housing, and leisure. Section 7 concludes.

2. The Demand for Air Quality

The basic framework for understanding the demand for air quality is based on a simple refinement to the neoclassical growth model incorporating preferences for non-market goods. Households will choose market goods, defined as a triple $X_{j,t} = (C_{j,t}, L_{j,t}, H_{j,t})$, consisting of non-durables consumption, leisure, and housing services, indexed by location j in period t. Households are able to choose among j differentiated locations defined by different levels of environmental amenities (e.g., quality); there are no mobility costs.[10] The production side will remain simple since relative prices are exogenous when taken to the data. To keep the model simple, I omit subscripts on location and time. The crucial insight from the setup is the conversion of a spatially and time varying public good into a private good whose price can be identified using geographical-specific variation in local private goods/services, individual-level time allocation, and locational sorting decisions. Appropriate instruments will allow for consistent identification of preference parameters (Ekeland et al., 2004).

A. Households: Following Rosen (1974), and in particular Roback (1982) who imposes homogeneity in household preferences, suppose that households have preferences over and choose private consumption (C), leisure (L), and housing, (H), taking environmental quality, (S), in a location as exogenous conditional on the choice of the triple, X, generated by a constant elasticity

[10] A reasonable theoretical concern is that individuals working and living in geographical locations sufficiently far away from each other might introduce measurement error air quality since the worker is implicitly consuming it in two locations. While the net effect is only likely to be attenuation since it will merely dull the signal-to-noise relationship between air quality and the market goods, I also conduct a robustness exercise in the empirical section leveraging only the subset of the sample that commutes less than half an hour to work, meaning that inference is only over individuals not susceptible to this potential identification problem; the results remain the same.

of substitution utility function

$$
\begin{aligned}
U(C,L,H,S) \quad &= \quad \exp(\rho)\left\{\alpha_C \log\left[\mu(g_C C)^{\phi} + (1-\mu)(g_S S)^{\phi}\right]^{\frac{1}{\phi}}\right.\\
&+ \qquad \alpha_L \log\left[\gamma(g_L L)^{\psi} + (1-\gamma)(g_S S)^{\psi}\right]^{\frac{1}{\psi}} \\
&+ \qquad \left.\alpha_H \log\left[\pi(g_H H)^{\zeta} + (1-\pi)(g_S S)^{\zeta}\right]^{\frac{1}{\zeta}}\right\}
\end{aligned} \tag{1}
$$

where U denotes utility, ρ denotes a preference ("taste") shock, g_C, g_L, g_H, and g_S denote technological trends that affect the value of consumption, leisure, housing, and environmental quality in preferences, α, μ, γ are share parameters, and ϕ, and ψ are elasticity parameters.,[11] Letting $X \in \{C,L,H\}$, $\omega \in \{\phi,\psi,\zeta\}$, and $\theta \in \{\mu,\gamma,\pi\}$, the elasticity of substitution between the private good/service and air quality is given by $\varepsilon_X \in (0,\infty) = -\mathrm{d}\ln(X/S)/\mathrm{d}\ln(U_X/U_S)$ and where $\omega = (\varepsilon_X - 1)/\varepsilon_X$ such that the three extreme cases are given by $\varepsilon_X \equiv 1/(1-\omega) \rightarrow 0$ $(\omega \rightarrow -\infty)$, $\varepsilon_X \equiv 1/(1-\omega) \rightarrow \infty$ $(\omega \rightarrow 1)$, and $\varepsilon_X \equiv 1/(1-\omega) \rightarrow 1$ $(\omega \rightarrow 0)$, which imply perfect complementarity, perfect substitutability, and Cobb Douglas elasticities, respectively (equivalent for leisure). The private good/service and air quality are gross substitutes when $\varepsilon_X > 1$ $(\omega > 0)$ and gross complements when $\varepsilon_X < 1$ $(\omega < 0)$. The technological parameters (g_C, g_L, g_S) characterize underlying trends that affect household preferences.[12] These preferences are sufficiently parsimonious to capture the nonseparability between consumption-leisure and air quality, but also tractable enough through additive separability between the consumption and leisure aggregates in order to obtain closed form expressions that map into the data. An important abstraction is that the triple $X = (C,L,H)$ is continuous, driven by the homogeneity of preferences after controlling for heterogeneity in individual-level observable tastes.

B. Firms: Suppose that firms use capital and labor to produce a homogeneous output using a constant returns to scale technology

$$
Y = F(K, 1-L) = K^{\theta}(1-L)^{1-\theta} \tag{2}
$$

and where the wage and price of capital are equal to their marginal products: $w = F_{1-L}$ and $r = F_H$. Wages and housing rents map into the hedonic framework for inferring household's valuation of environmental amenities. Capital produces pollution, which reduces air quality, given by an arbitrarily defined concentration response function, $S = g(K)$, where S is decreasing in K.

[11]The assumption of additive separability between consumption and leisure is widely applied in the elasticity of labor supply literature (e.g., Altonji, 1986). Any mis-specification in preferences is unlikely to induce bias because of quasinatural variation generated via the instruments.

[12]For example, Hall and Jones (2007) suggest that the rise in healthcare spending can be explained by a saturation of marginal utility for consumption goods; likewise, Aguiar and Hurst (2007) suggest that households have experienced a steady increase in leisure since the 1950s. Importantly, trends in environmental quality (g_S)—governed plausibly by environmental policy (Shapiro and Walker, 2014)—will affect the value of consumption and leisure depending on the share and elasticity parameters.

C. Equilibrium: Under the assumption that the representative household maximizes their utility subject to a simple budget constraint, households solve

$$V(K_t) = \max_{C_t, L_t, H_t, K_{t+1}} \{U(C_t, L_t, H_t; S) + V_{t+1}(K_{t+1})\}$$

subject to his budget constraint

$$C_t + p_t H_t + K_{t+1} = w_t(1 - L_t) + K_t(1 + r_t - \delta)$$

where p is the price of housing, and the price of consumption is the numeraire. The production function and relatives prices are given by Equation 2. Dropping time subscripts, optimizing behavior implies that the following three equilibrium conditions hold

$$U_C(C, S) = \beta \mathbb{E}\left[U_{C'}(C', S')(1 - \delta + r')\right] \tag{3}$$

$$U_L(L, S)/U_C(C, S) = w \tag{4}$$

$$U_H(H, S)/U_L(C, S) = p/w \tag{5}$$

together with the resource constraint: $Y = C + p_t H_t + K_{t+1} - (1 - \delta)K_t$. When the household is not working, he allocates his time exclusively to leisure. The first condition is the intertemporal Euler, which equates the marginal utilities of consumption over time, together with the externality that is a function of savings decisions since housing is a long-lived asset. The second condition is the intratemporal Euler on consumption-leisure, which equates the ratio of the marginal utilities of leisure to consumption with the wage. Equation 4 will provide a characterization of the equilibrium differential that allocates individuals across locations and compensates those who face lower consumption of environmental amenities. Similarly, the third condition is the intratemporal on housing-leisure, which equates the ratio of marginal utilities of housing and leisure to the price ratio.[13] The nonseparabilities in these equations allow me to identify a virtual price on S—that is, a price associated with air quality as if households could purchase it directly; see Perroni (1992)

[13] While housing is modeled as a flow—meaning that households are renters—modeling housing as an asset would induce an intertemporal Euler between housing services and consumption, instead of an intratemporal Euler on housing and leisure. Since I use cross-sectional data in this paper, adding another intertemporal arbitrage condition would not allow me to take it to the data. Quantitatively, the only modification that an intertemporal (rather than intratemporal) induces is that housing and air quality appear as the change in logs, rather than purely logs, and the rental rate of housing is included. To the extent that my identification strategy exploits exogenous variation, modeling the change in housing & air quality and/or omitting the the return on housing will not bias the model estimation.

and Amiran and Hagen (2014) for theoretical justification.

D. Comparison to the Literature: There are a couple of conceptual issues relating to hedonic price theory that need to be contrasted with prior work in canonical hedonic models (Roback, 1982; Rosen, 1974). Equations 4 and 5 are the objects of endogenous sorting processes that under-gird the equilibrium in local labor and housing markets. Kuminoff (2012) provides a unified model for households to sort across jobs and housing locations without one implying the other, whereas the canonical models assume that every time a person changes houses, they change jobs, and vice versa. Given that housing and employment opportunities are fundamentally linked—for example, U.S. counties became less stratified by public goods provision and housing demographics as moving costs declined between 1850-1990 (Rhode and Strumpf, 2003)—obtaining unbiased estimates of preferences for local public goods hinges on how these decisions are jointly modeled. Specifically, Equation 5 captures the simultaneity of the labor-housing decision by equating their marginal utilities equal to the price ratio.

While the coefficient on air quality in the regression is analogous to the Roback (1982) measure of willingness to pay—and, indeed, these elasticities are "inputs" to willingness to pay—there are three main differences. First, WTP estimates cannot be readily inserted into the current class of dynamic macroeconomic models. In addition to the fact that my companion work (Cai and Makridis, 2015) is, to my knowledge, the only example of a DSGE setup featuring both housing and environmental quality, there is no concept of spatially heterogeneous non-market goods in standard macroeconomics. This makes it impossible to calibrate to target WTP estimates from the micro-data. Second, Flores and Carson (1997) show that there are at least three plausible reasons that the income elasticity (of environmental quality) provides insufficient information to evaluate the welfare impacts of a change in public goods. One condition, for example, is when the income elasticity depends on goods that might be rationed, which is always the case in the context of environmental policy interventions since they induce higher prices and declines in employment. Third, many previous studies make restrictive assumptions about the exogeneity of certain variables (e.g., labor/leisure; see Flores and Graves (2008)), which effectively behaves as an omitted variables bias problem. Although quasinatural experimental methods offer a solution, they face separate identification challenges (see Kuminoff and Pope (2014)).[14]

E. Empirical Implementation: While the equilibrium conditions produce non-linear equations that are difficult to take directly to the data, the two intratemporal Euler equations can be simplified into the following

[14]To put this in perspective, the average household sacrifices over $5,000 per year to consume non-market amenities in their geographical location (Bieri et al., 2014), bigger cities tend to have differences in availability of consumption goods (Handbury and Weinstein, 2014) and higher wages (Glaeser and Resseger, 2009; Davis and Dingel, 2014).

$$\log w = (\psi - 1)\log L + (1-\phi)\log C + (\phi - \psi)\log S + \rho(X) + \epsilon \quad (6)$$

$$\log(p/w) = (\zeta - 1)\log H + (1-\psi)\log L + (\psi - \zeta)\log S + \rho(X) + \nu \quad (7)$$

where ϵ and ν include the constants from T_L, T_C, and other unobserved heterogeneity (including measurement error), and $\rho(X)$ contain a vector of shocks (e.g., household, state, and/or county -level controls). The derivation is documented in the Appendix. Broadly speaking, after the equilibrium conditions are log-linearized around the steady state, a sufficiently high level of the public good with respect to the atomistic individual's level of the private good will make some terms disappear. The elasticities (dlog$w$/dlog$S$) and (dlog(p/w)/dlog$S$) identify the labor demand and home capitalization effects. The intuition is that variation from Equation 4 is directly informative about household's trade offs between consumption and leisure, and thus the labor supply decision, whereas Equation 5 is directly informative about household's trade offs between housing and leisure, and thus the home ownership decision. Cross-sectional variation in workers' equilibrium sorting decisions from Equations 3 and 4 provides a way of estimating preference parameters associated with environmental quality. Because econometricians infer willingness to pay for public goods based on the identifying assumption that the quality of a location-specific amenity is increasing in the price ("value") of its location-specific private goods (e.g., wages or housing rents), then understanding the feedback among the traded and non-market goods is crucial. Individuals choose their consumption of non-market goods only through their consumption of market goods; nonseparability is the only lens to facilitate such an analysis.

There are two reasons that this simpler specification is preferred to the "true" non-linear solution. First, cleanly identifying the micro-elasticities is impossible in the non-linear model since it is not yet computationally feasible to add such large fixed effects specifications with instrumental variables. Second, because of the imputation of consumption and leisure, measurement error may enter the non-linear model in unexpected ways and bias the coefficients of interest. On top of these two substantive reasons, the results are much more transparent with the linear model. Nonetheless, I recognize that mis-specification may be correlated with unobserved shocks to the outcome variable and introduce a separate source of bias. The Appendix contains two robustness checks that obviate the concern. First, I show that the log-linearized version of the model—including the extra non-linear terms—produces similar coefficients. Second, I apply a Kmenta (1967) approximation around the points $\omega = 0$ for $\omega \in \{C, L\}$ and show that the higher-order terms are not significant after controlling for the direct effects.

While hedonic wage regressions have (to my knowledge) only been used to decompose the riskiness of different jobs using partial equilibrium methods (Viscusi and Aldy, 2003; Aldy and Viscusi, 2008; Viscusi, 1979), Equation 4 shows that a relationship between wages and air quality can be derived from a theoretically consistent structural general equilibrium model. The slope of the hedonic wage on air quality has the interpretation as a marginal change in air quality for the individual found at the baseline level (Bockstael and McConnell, 2007). The traditional Roback (1982) framework will take the difference between the estimated coefficient on pollution obtained from a regression of housing values on pollution and labor income on pollution. However, Equations 6 and 7 reveal the potential for bias since these conventional hedonic wage and housing regressions conflate the aggregate price times quantity with simply the price. That is, from the lens of the Roback model, the willingness to pay for air quality is given by the elasticity of air quality in Equation 7. There are other empirical concerns that traditional hedonic regressions face concerning the bundling of unobserved consumption/leisure into the error and the reliance on county, rather than household, -level data; these will be discussed in detail later. The value of a statistical life is the "population"s aggregate willingness to pay for an increase in one expected life saved (Bockstael and McConnell, 2007 p.219). Nevertheless, the interpretation of $\mathrm{dlog}(p/w)/\mathrm{dlog}S$ is: a 1% change in air quality induces a $(\psi-\zeta)\%$ decline in the ratio of the relative price of housing versus labor. For example, if a policy intervention increases air quality by 1% for N people, then the aggregate WTP would be $N \times (\psi-\zeta)\%$ and the value of a statistical life would be $(\psi-\zeta)\% \times$ risk of death. Getting the VSL right matters: of the $2 trillion in estimated net benefits of the CAAA, a staggering $1.7 trillion are attributed to reductions in mortality.

3. Micro-data on Consumption, Leisure, and the Environment

3.1. Descriptive Statistics

Through the American Community Survey, the Census Bureau recovers detailed household-level information across the entire United States at disaggregated geographical levels. Unfortunately, the Census does not contain information on aggregate consumption or leisure. To deal with this limitation, the final data set exploits auxiliary data from the Consumption Expenditure Survey and the American Time Use Survey to impute consumption and leisure in the Census, matched with county-level data on air quality and weather, discussed later in the section. The Census's wide coverage is vital since my identification strategy requires variation in regional housing and labor markets. That is, observing observationally equivalent workers in different locations reveals information about their valuation of environmental amenities. The summary statistics below provide a characterization of the cross-sectional and intertemporal variation between 2005-2010.

All of the traditional anecdotal facts about the Great Recession are evident in the established data set: consumption declined by 10%, both the mean and standard deviation of leisure rose by about 3%, housing prices declined by 10%, unemployment doubled, and pollution—whether measured as total suspended particulates or PM10—declined by about 20%. While the rise in consumption inequality (Aguiar and Bils, 2009) and time use (Aguiar et al., 2013) over the Great Recession are not new, the decline in pollution is stark. However, what remains unknown is how counties with higher air quality differ from those with lower air quality. Figure 2 plots the raw correlations in the data between county-level air quality and market "goods", such as consumption, leisure, wages, and housing values. Counties with greater air quality tend to have lower consumption and wages, but higher leisure and housing values, consistent with the hypothesis that environmental amenities are capitalized into both human and physical assets (e.g., time and houses).

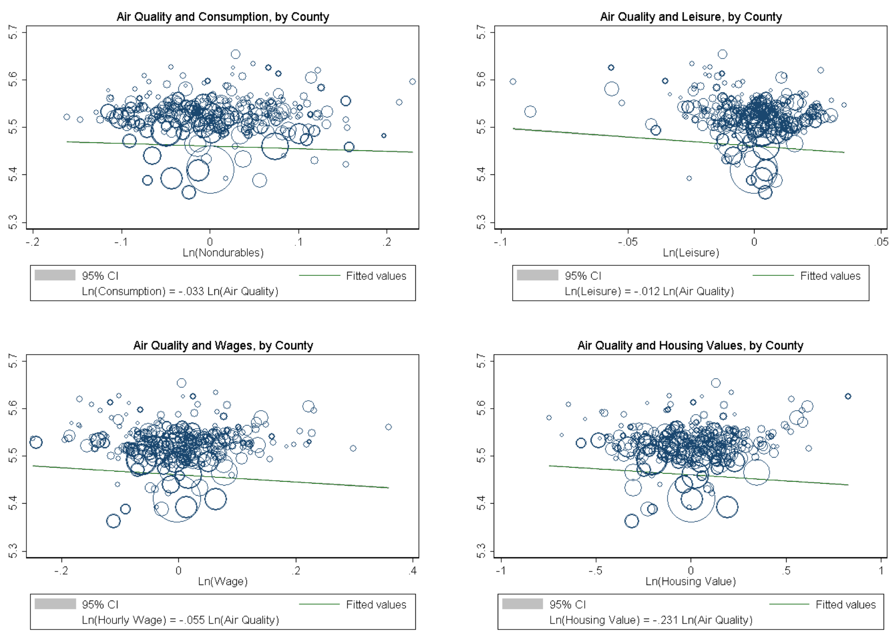

Figure 2: The Relationship between Air Quality and Market Goods

Notes.–Sources: CEX, ATUS, ACS, EPA. The figure plots the residualized log variable consumption, leisure, wages, and housing values) with log air quality (normalized index so higher is better) and subsequently averaged across counties. Each observation is a county-level average between 2000-2014 and the circles are weights based on the population of the county. Controls used for the partial correlation include: educational attainment, number of children, indicator for disability, age, gender, number of bedrooms, five bins for the year the house was built, household tenure, race fixed effects, and population. Economic controls include the state labor force, employment, payroll expenditures, and the number of establishments within 9 different bins of firm size (1-4 employees, 5-0, 10-19, 20-49, 50-99, 100-249, 250-499, 500-999, 1000+).

While not exogenous variation, the Great Recession is a suitable natural experiment for recovering preferences over environmental amenities because of the vast amount of reallocation and migration dynamics. As counties experienced different magnitudes and durations of labor and housing market shocks, individuals moved locations. For example, many households moved from

Table 1: Summary statistics, 2005-2010

	2005		2006		2007		2008		2009		2010	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Individual												
Consumption	24299	7270	24434	7154	25071	7201	23387	6679	24346	6914	23804	6844
Utilities	3073	2577	3370	2744	3428	2819	3572	2901	3560	2911	3576	2917
Leisure	1854	215	1867	210	1792	227	1866	216	1860	216	1815	216
Labor earnings	40293	37487	41088	38012	43537	41071	44116	40934	44378	40855	44586	42682
Education	13.75	2.73	13.74	2.72	13.77	2.74	13.77	2.77	13.80	2.79	13.84	2.83
Family income	88582	75353	91170	78330	96662	84957	97941	87425	96884	82950	96284	81594
White	0.78	0.41	0.77	0.42	0.77	0.42	0.78	0.42	0.77	0.42	0.77	0.42
Black	0.09	0.29	0.09	0.29	0.09	0.29	0.10	0.30	0.10	0.29	0.10	0.29
Age	43.08	12.67	43.24	12.75	43.37	12.88	43.42	13.07	43.65	13.13	43.66	13.20
Male	0.53	0.50	0.53	0.50	0.53	0.50	0.52	0.50	0.52	0.50	0.51	0.50
Bedroom	4.01	0.98	4.02	0.98	4.04	0.98	4.06	1.18	4.05	1.13	4.04	1.17
Married	0.61	0.49	0.61	0.49	0.61	0.49	0.59	0.49	0.59	0.49	0.58	0.49
Housing/Economy												
Housing Tenure	2.02	2.19	2.05	2.20	2.04	2.20	2.06	2.20	2.09	2.23	2.09	2.25
Selected owner costs	13984	13433	15188	14817	16042	15826	15938	16084	15610	15649	15204	15371
Housing Value	231023	237789	248659	251542	253930	252417	248735	359452	225089	317915	217464	299507
Year home built	4.91	2.34	5.04	2.42	5.16	2.50	5.23	2.63	5.33	2.78	5.33	2.86
Unemployment Rate	5.11	0.84	4.57	0.83	4.64	0.85	5.92	1.11	9.50	1.71	9.96	1.83
No. Employed, 000s	1794	2392	1896	2524	1875	2501	1869	2521	1773	2388	1736	2314
Payroll Exp., 000s	76212	105504	83715	116882	86241	120914	87554	124064	83498	118219	85240	118427
No. Establishments, 000s	113	150	118	157	119	159	117	157	115	154	115	154
Population, 000s	1367	2098	1422	2157	1417	2145	1408	2132	1423	2151	1439	2164
Environment												
Air Quality (Index)	242.62	12.44	244.50	12.25	243.79	12.87	247.15	11.22	249.22	11.21	249.68	10.58
Total Suspended Particulates	57.81	15.06	54.55	14.29	60.57	18.17	53.69	13.77	50.73	13.43	46.25	8.06
PM10, micrograms/cubic meter	25.02	6.53	24.93	7.47	26.17	8.13	23.44	6.19	21.72	6.36	20.51	4.88
Fastest 2 Min. Wind	17.13	3.67	17.62	3.44	17.23	3.80	17.79	3.68	17.28	3.66	17.38	3.42
Fastest 5 Sec. Wind	21.09	4.90	21.87	4.69	22.02	5.33	23.04	5.20	22.56	5.17	22.88	4.98
Mean Resultant Wind	6.10	2.06	6.32	1.96	6.11	1.88	6.34	1.91	6.14	1.94	6.27	1.80
Wind, miles/hour	7.83	1.95	7.97	1.75	7.62	1.79	7.88	1.83	7.59	1.82	7.69	1.69
Precipitation (in. to 100ths)	8.93	70.20	7.37	53.51	8.26	58.46	9.53	59.93	10.67	67.58	10.87	65.53
Snow (in. to 10ths)	0.16	1.55	0.08	0.68	0.12	0.93	0.41	4.08	0.41	3.44	0.50	3.89
Max Temp.	68.36	11.98	69.43	12.50	68.62	13.22	67.78	14.10	67.28	14.32	67.54	13.57
Min Temp	47.88	9.62	48.45	9.31	47.85	9.95	46.86	10.77	46.84	11.13	47.45	10.14
Observations	78538		78072		78242		83008		80872		82774	

Notes.–Sources: Environmental Protection Agency AirData's air quality index (AQI), the Census Bureau's American Community Survey (ACS), Consumption Expenditure Survey (CEX), and American Time Use Survey (ATUS). The table contains the means and standard deviations of the most relevant variables contained in the econometric specifications. Using the definition of non-durables from Attanasio and Weber (1995), it is the sum of food (home and away), alcoholic beverages, tobacco, services (e.g., repairs), heating, gasoline, transportation, electricity, water, fuel, personal care, clothing, footware, and rents. Using the definition of leisure from Aguiar and Hurst (2007), specifically their "Measure 1", it is the sum of socializing, passive and active leisure, volunteering, pet care, gardening. Housing values are self-reported from the ACS and are upwards biased to the extent homeowners are overly optimistic about the sale price of their property. Air quality, call it S, is transformed from the Environmental Protection Agency's (EPA) "AirNow" air quality index, call it $\tilde{S}$, by taking $S = 300 - \tilde{S}$, where 300 is a hazardous measure of air quality. With the transformation, higher values enter positively into utility. Total suspended particulates and particulate matter are measured in micrograms per cubic meter.

one county to another for new employment opportunities and/or housing decisions (i.e., to move into a less expensive house). The demographic reallocation and resorting coincides with a massive decline in pollution (rise in air quality) that affected household's locational choices. To the extent that households value environmental amenities, holding all else constant, a homeowner with tastes for non-market goods would choose to locate in an area with better amenities. The density and time series properties of air quality are illustrated below in Figure 3. All additional data details are located in the Appendix.

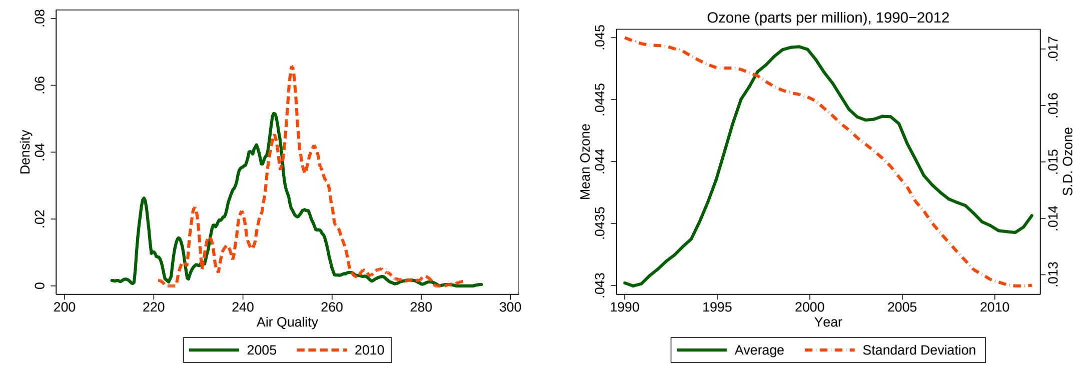

Figure 3: Distribution of Air Quality (2005/2010) and Ozone Pollution (1990-2014)
Notes.–Source: Environmental Protection Agency AirData's air quality index (AQI) and Annual Summary files. Using the same air quality indices as before, the left figure plots the distribution of counties' air quality levels (using the same transformed measure using a smoothed kernel density estimator and, using the annual summary file, the right figure plots the average total suspended particulate mean and standard deviation over the past two decades. The smoothed polynomial weights observations according to county-level population averages. The TSP measure is computed as the 24-hour average level, but is also robust to the 1-hour average level. See https://aqs.epa.gov/aqsweb/codes/data/SampleDurationCodes.html for details.

3.2. Imputing Nondurables Consumption and Leisure

While the Census data contains rich household-level demographic details and comprehensive coverage, it lacks measures of consumption and leisure. A naive way—that fails to provide sufficient variation to identify the parameters of interest—would involve proxying non-durables consumption with electricity expenditures and leisure with non-work hours (e.g., 5100 minus hours worked in a year).[15] There is a detailed theoretical literature on corrections for measurement error and im-

[15] Just as food expenditures were used to proxy for non-durables in the early literature on estimating the intertemporal elasticity of substitution (Hall and Mishkin, 1982), and later found to be weak proxies because of their relative low covariance with the overall consumption bundle (Attanasio and Weber, 1995), electricity expenditures fall into the same trap. Similarly, taking leisure as the difference of total hours and hours worked would attenuate the variation since many allocations of leisure time (e.g., sleeping) are unresponsive to environmental amenities. An entirely separate strategy would involve creating grids of representative households (e.g., households in a certain income bracket, education level, race, etc) and match among the datasets; this approach is guaranteed to provide strictly worse and more inaccurate results since the household-level behavioral responses to county-level air quality is vital for identification—and preferences for environmental amenities differ immensely across the population.

putation, but a sparse empirical literature.[16] Using a sieve estimator and auxiliary data from the Consumption Expenditure Survey (CEX) and the American Time Use Survey (ATUS), I predict the distributions of non-durables and leisure. The semi-parametric estimator is able to capture the rich nonlinearities of the demand for non-durables and leisure based on data observed in all of the datasets and follows thematically in line with Blundell et al. (2008) who estimate a demand system to impute non-durables in the Panel Study of Income Dynamics using data from the CEX.[17] Aside from some flexible semi-parametric restrictions, the only limitation is the requirement that the variables used in the imputation are common across all datasets such that only the to-be-imputed variable is missing in the using (primary) data set. To further enhance the quality of the imputation, I estimate a propensity score associated with the probability of an individual being in the American Community Survey and consider scores within the range of .1 and .9. Crump et al. (2009a) show that this helps ensure overlap in the distributions.

Sieve estimators are relatively easy to implement and have desirable properties when the approximating functions are unknown (Chen, 2007). Its accuracy depends on the extent to which the target (to-be-imputed) variable is a smooth function of its explanatory variables. The sieve estimator achieves asymptotic consistency by allowing these explanatory variables to enter nonlinearly and with many higher order terms. An alternative approach would be to specify a structural model; see the footnote for details.[18] However, because it relies on the assumption that the model is properly specified, mis-specification in the imputation can give rise to unknown forms of non-classical measurement error in the actual estimation of air quality elasticities. Robustness checks indicate that my sieve estimator yields the most accurate estimates.

Denote $Y_{1i} = (Y_{1i}1, ..., Y_{1iM})' \in \mathcal{R}^M$ as the target M variables (e.g., non-durables and leisure),

[16] One vein of the literature (e.g., Robins et al. (1994)) focuses on consistent estimation when some of the variables are missing for a subset of the sample series, but not all; these methods have tended to involve inverse probability weights associated with the missing variables. Another vein of the literature (e.g., Chen et al. (2008)) focuses on consistent estimation when variable(s) might be missing for a large subset of the sample series or all of it; these methods have tended to emphasize auxiliary datasets and semi-parametric method of moments and propensity score estimators for out-of-sample estimation. See Chen et al. (2011) for a detailed survey.

[17] Although this was the first method that I attempted, it did not succeed in allowing me to impute consumption and leisure, potentially because there is less time series variation. Campos and Reggio (2014) emphasize that instrumenting does not address the asymptotic bias resulting from the covariance between their control variables and the error, which is non-zero especially if demographic and other household-level variables are reported with error differentially in the two datasets. For example, Gibson (2002) finds that household size is correlated with measurement error since a single respondent asked to remember expenditures for an entire household is likely to make more mistakes the greater the number of people he must remember for.

[18] For example, let $u(c,l) = [ac^\alpha + (1-a)l^\alpha]^{1/\alpha}$ denote a constant elasticity of substitution function between consumption and labor. If households maximize utility subject to a budget constraint with no savings (just equal to the wage times labor services), then the intratemporal Euler implies that $(1-a)l^{\alpha-1}/(aw) = c^{\alpha-1}$. Taking the log of both sides yields $\log c = \log l - (\alpha-1)^{-1}\log w + \epsilon$, where $\epsilon = (1-a)/a$. Running some version of least squares with instruments could yield an unbiased measure of $\hat{\alpha}$. Using this estimate, and letting a be the corresponding value share taken from the data, then consumption could be imputed as $\tilde{c} = \left[(1-a)l^{\hat{\alpha}-1}/(aw)\right]^{1/(\hat{\alpha}-1)}$.

Y_{2i} as the proxy for Y_{1im}, $X_{1i}^{'} \in \mathcal{X}$ as the "equivalence scale" common across households, with $\dim(X_1) \geq 1$, $\dim(X_2) \geq 1$, and $Z_i = (Y_{1i}^{'}, Y_{2i}, X_i)'$. Denote $\alpha = (\theta, h_1, ..., h_M)$ as all the unknown parameters of interest and $\mathcal{A} \equiv \Theta \times \mathcal{H}_1 \times \cdots \times \mathcal{H}_M$ as the parameter space where $\theta \equiv (\theta_1^{'}, \theta_{2,1}^{'}, ..., \theta_{2,M}^{'})'$ denotes the vector of finite dimensional parameters given by $\theta \in \Theta$, a compact subset of $\mathcal{R}^{d_\theta}$ with $d_\theta \equiv (1+M)\dim(X_1)$. The terms $h_m \in \mathcal{H}_m$ will denote the unknown demand curves associated with subsets of the target good m, $m = 1, ..., M$, where $\mathcal{H}_m$ is a subset of a space of functions that are square integrable against the probability measure of Y_{2i}. Letting $\rho \equiv (\rho_1, ..., \rho_M)' \in \mathcal{H}^M$ represent

$$\rho_m(Z_i, \alpha) \equiv Y_{1im} - h_m(Y_{2i} - \chi(X_{1i}^{'}, \theta_1)) - X_{1i}^{'}\theta_{2,m} \tag{8}$$

for a specified functional form $\chi(\cdot)$, then individual i facing the same prices for goods $m = 1, ..., M$ will have a demand curve that satisfies $E\left[\rho(Z_i, \alpha_o)|X_i\right] = 0$ where $\alpha_o \equiv (\theta_o, h_{o1}, ..., h_{oM}) \in \mathcal{A}$ is the true (unknown) parameter. The objective is to estimate θ_o and the demand functions h_{om} in order to recover parameters that fully characterize the mapping between the target and input variables of interest. Under very general regularity conditions, Blundell et al. (2007) show that these demand restrictions will be satisfied. Following Ai and Chen (2003) who establish a framework for estimating moment conditions of unknown functional forms, consider approximating functions $h_m \in \mathcal{H}_m$ by $h_{m,n} \in \mathcal{H}_{m,n}$ for $m = 1, ..., M$ where $\mathcal{H}_{m,n}$ is a sieve space for $\mathcal{H}_m$ (e.g., Fourier series, splines, and so on) so that $\mathcal{H}_{m,n}$ becomes dense in $\mathcal{H}_m$ as $n \to \infty$. Arbitrarily fixing a value of $\alpha = (\theta, h_{1,n}, ..., h_{M,n})$ in the sieve parameter space, then the population conditional moment function, characterized by $g(x, \alpha) \equiv (g_1(x, \alpha), ..., g_M(x, \alpha))'$, can be estimated using sieve generalized least squares

$$\min_{\alpha \in \mathcal{A}_n} \frac{1}{n} \sum_{i=1}^{N} \rho(Z_i, \alpha) \left[\hat{W}(X_i)\right]^{-1} \rho(Z_i, \alpha)$$

where $\hat{W}$ is a consistent estimator of a positive definite weighting matrix.[19],[20] Using flexible semi-parametric functions (e.g., splines and polynomials) over electricity and hours worked for the cases of non-durables and leisure, respectively, as well as entropy balancing based on the first, second, and third moments of hours worked and log wage income, I estimate Equation 8 to obtain relationships on demand.[21],[22] While I consider a variety of specifications (see the Appendix), the annual means/variances are plotted below in Figures 4 and 5 for imputed consumption and leisure.

[19]Blundell et al. (2007) provides an excellent methodology for implementing this procedure when the target variables are endogenous and an instrument is needed for Y_2; here, it is taken as exogenous.

[20]Hellerstein and Imbens (1999) show that orthogonality conditions can be created by using moments from axillary data (e.g., Census and ATUS).

[21]Splines provide a better fit of the data within local regions of the space of points and are well behaved on the endpoints, whereas polynomials can experience unexpected fluctuations ("Runge's phenomenon").

[22]See Hainmueller and Xu (2013) for a procedure on entropy balancing. The treated (ATUS) sample is made to resemble the control sample (Census) based on the first three moments of hours worked and log wage income.

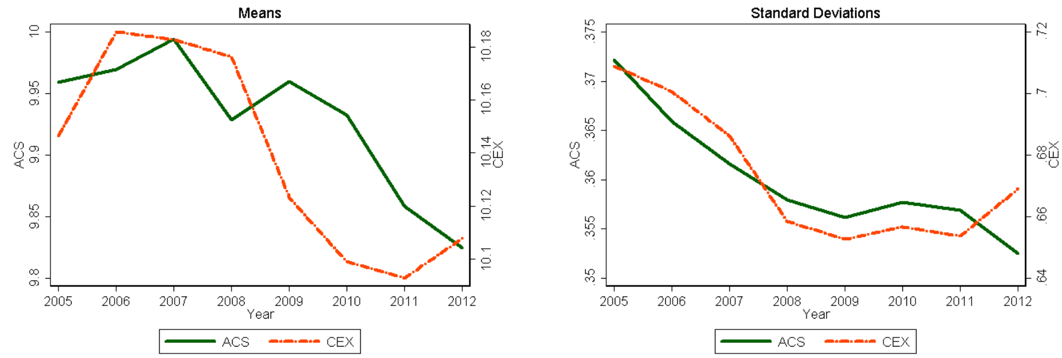

Figure 4: Non-durables Mean and Variance, Actual (CEX) and Imputed (ACS)
Notes.–Source: CEX and ACS. These plot the mean and variance of consumption from the CEX with the mean and variance of the ACS-imputed consumption measure. Using the definition of non-durables from Attanasio and Weber (1995), non-durables is the sum of food (home and away), alcoholic beverages, tobacco, services (e.g., repairs), heating, gasoline, transportation, electricity, water, fuel, personal care, clothing, footware, and rents.

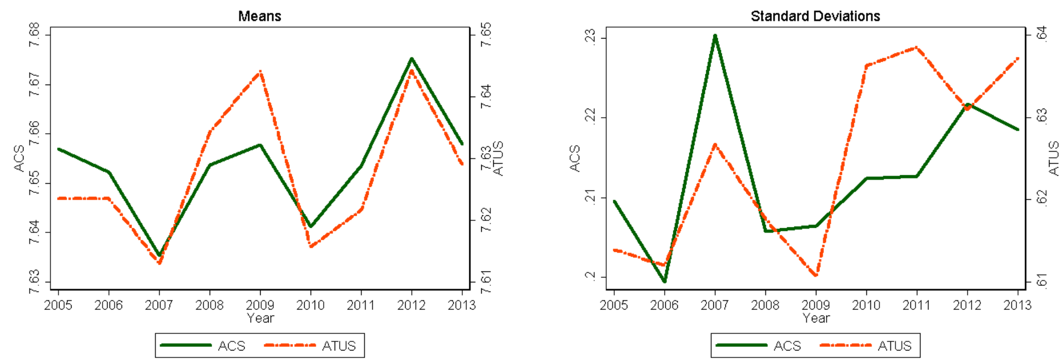

Figure 5: Leisure Mean and Variance, Actual (ATUS) and Imputed (ACS)
Notes.–Source: CEX and ACS. These plot the mean and variance of leisure from the ATUS with the mean and variance of the ACS-imputed leisure measure. Using the definition of leisure from Aguiar and Hurst (2007), their "Measure 1" definition includes: socializing, passive and active leisure, volunteering, pet care, gardening.

Figures 4 and 5 both characterize the mean log non-durables and leisure in the actual (CEX and ATUS) datasets and the imputed values (in the CEX and ATUS) datasets. To the extent that there is a heavy overlap in the distribution of covariates (see the Appendix for tables), then the coefficients used to impute non-durables and leisure in the CEX and ATUS datasets will be externally valid for application in the Census. For brevity, all my robustness checks are relegated to the Appendix with the summary that the imputed values match the key features of the actual distributions, as defined by the CEX and ATUS. Differences in levels are controlled in my second stage regressions using fixed effects on year. Importantly, remember that the objective here is prediction, not causality, so endogeneity is not germane. While I also adopted an instrumental variables strategy (similar to those in Blundell et al. (2008)), these results proved to be the most predictive.

4. Identification Problems in the Hedonic Method

Hedonic methods are used to infer prices on environmental amenities by exploiting individuals' observed choices. The canonical framework that followed from Rosen (1974) considers an environment where households have different tastes and will choose a location that maximizes their utility based on the amenities of the location, including both market and non-market goods. Suppose that housing can be decomposed into its various components, differentiating between air quality and all other taste shifters or housing attributes

$$\log H_{icts} = X_{icts}\beta + \log S_{icts}\zeta + \epsilon_{icts} \tag{9}$$

where there exists endogeneity arising from locational sorting based on air quality and omitted variables that co-vary with air quality. These hedonic regressions are confounded by omitted variables and non-random sorting problems. While panel data can be used to remove time-invariant unobservables, first-differencing and/or instrumental variable approaches using the Clean Air Act attainment status identify the marginal willingness to pay interpretation only if the hedonic price function is constant over the sample period. In other words, while quasi-experimental methodologies successfully address omitted variables, they come at the expense of eliminating key cross-sectional variation needed to identify willingness to pay in the presence of non-marginal changes or heterogeneous treatment effects (Kuminoff and Pope, 2014).[23] My model provides a tractable framework for estimating structural parameters in light of possible omitted variables (e.g., consumption and leisure) and their endogeneity.[24]

While these concerns are not new, the model in Section 2 illustrates additional identification problems. Depending on the strength of nonseparabilities, determining the willingness to pay and/or conducting welfare analysis requires knowledge of the income elasticities of demand for all other rationed goods (Flores and Carson, 1997). My model highlights three additional sources of upwards bias in traditional approaches. First, unobserved shocks to housing values are correlated with changes in the marginal utilities of consumption and leisure (Mian et al., 2013; Mian and Sufi, 2014). Environmental policies affect housing prices in many ways other than through air quality, such as labor demand and asset prices. Second, the variance of these unobserved *household-*

[23] Bockstael and McConnell (2007) remark that "overcoming or avoiding this type of omitted variable bias is much more difficult... virtually all applied papers ignore this source of bias, and little is known about whether the bias thus generated is of a significant magnitude" (p. 177). Both Ekeland et al. (2004) and Heckman et al. (2010) clarify that this is not a failure of the original setup by Rosen (1974), but rather a convenience that has taken root in much of the literature.

[24] See Bajari et al. (2012) for an alternative solution that exploits assumptions about expectations and unobservables in the hedonic price function. However, during the Great Recession and/or environmental policy interventions, these assumptions may be fragile.

level shocks is time varying and orders of magnitude larger than those in the *county-level* data. Aggregating to the county-level tends to ignore important sources of heterogeneous treatment effects that unobserved shocks have on the population. Third, using property values or total labor earnings as the dependent variable, rather than the relative price of housing services or time, creates a correlation between the aggregate expenditures on a resource and other endogenous variables, such as consumption and leisure (Flores and Graves, 2008).

4.1. Unobserved Shocks

Unobserved shocks to housing values, such as local labor market and transitory income shocks, affect housing values shift the hedonic price function, driving a wedge between the average capitalization effect and household's willingness to pay for amenities (Kuminoff and Pope, 2014). In the absence of an ideal instrument that randomizes individuals into counties with different air quality levels, identifying assumptions may be sensitive to whether environmental policy affects housing values through mechanisms other than pollution. The most likely threats to identification emerge from the ways in which shocks to the hedonic pricing function affect consumption and leisure among households. While the direct effects of environmental policy on housing values are examined carefully in companion work (Cai and Makridis, 2015), I turn to reduced-form evidence to suggest mechanisms through which environmental regulations may affect housing values other than via pollution levels. Using variation in attainment status under the Clean Air Acts over the 2005-2010 period, arising from various changes to environmental regulations from the Nitrogen Oxide Budget Trading Program (NBP, see Curtis (2014) for details), I regress log property values on a county's attainment status, conditional on controls and fixed effects. Finding a statistically significant effect on attainment status, conditional on air quality, is evidence of a violation to the orthogonality restriction that environmental regulations affect housing values only through their effects on pollution.

To evaluate the plausibility of the identifying assumption that attainment status affects property values only through air quality, consider two concrete examples of the "coefficient comparison test," an analogue of the "balancing test" in regression discontinuity designs (Pischke and Schwandt, 2015). Table 2 documents these results. First, property values and attainment status are relatively correlated even after controlling for air quality (e.g., see columns 1, 3, and 5). Even though fixed effects help mitigate this problem, counties in attainment of PM10 regulation tend to have about 20% higher property values even after differencing out time-invariant observables across counties. Related to an argument by Curtis (2014), this suggests that there is some value for caution associated with using attainment status as an instrument for pollution since counties in and out of attainment may differ in systematic ways, prompting the need for a matching estimator. Second, while failure to control for consumption or leisure does not dramatically affect the point estimates on air quality when fixed effects are included (e.g., see columns 4 and 6), they do when county fixed effects are not included. As long as quasinatural experimental strategies control for these systematic differences across counties, the main insight from Chay and Greenstone (2005) remains valid.

4.2. Time Varying Standard Deviations, Counties versus Households

The variance of unobserved shocks to households might change over time and differ from the variance of unobserved shocks at the county-level. Within-county reallocation correlated with air quality will bias the elasticity between air quality and housing prices. Such reallocation is likely given that the Great Recession—and any period featuring a large-scale intervention—had such heterogeneous impacts across the United States. For example, Yagan (2014) documents that there were large spatial adjustment frictions in the U.S. labor market during the Great Recession. Households, rather than counties, should be the unit of analysis.[25]

Undertaking a similar decomposition exercise as Lemieux (2006) and Chay and Lee (2000), suppose the error in Equation 4 (and analogously for Equation 5) consists of two components, $\epsilon_{it} = q_t e_{it} + v_{it}$, where q denotes the price of unobserved skills and v denotes idiosyncratic noise. Since $Var(\epsilon_{it}) = q_t^2 Var(e_{it}) + Var(v_{it})$, then it becomes possible to decompose the residual by the level of aggregation. If the extent of unobserved variation at the county versus household level is varying over time, then merely conditioning on demographics with year fixed effects will not provide consistent estimates of marginal willingness to pay when relying on county-level aggregations. To illustrate this phenomenon of the data, Figure 6 plots the county and household average log consumption, leisure, housing, and earnings over the relevant sample period with correlations displayed at the bottom of the figures. Interestingly, the level of aggregation matters a great deal for all except leisure, indicating heterogeneous treatment effects in the shocks individuals received over the Great Recession (consistent with skill biased technical change). The extent of the reallocation within counties implies that relying on county-level, rather than individual-level, can produce biased parameter estimates.

[25] The importance of examining dynamics at sufficiently disaggregated levels has been articulated by Davis and Haltiwanger (1992) in the context of firm dynamics, Auffhammer et al. (2009) in the context of the Clean Air Acts, and Banzhaf and Walsh (2008) in the context of Tiebout sorting.

Table 2: Housing Values and Clean Air Act Attainment Status

	(1)	(2)	(3)	(4)	(5)	(6)
Ozone						
Attainment Status	.237***	-.025	.273***	-.026	.176***	-.028
	[.061]	[.031]	[.064]	[.031]	[.059]	[.032]
Ln(Air Quality)			.842	-.135	.460	-.105
			[.820]	[.363]	[.912]	[.342]
Ln(Consumption)					1.125***	1.116***
					[.072]	[.083]
Ln(Leisure)					-.549***	-.317***
					[.085]	[.070]
R-squared	.38	.42	.38	.42	.42	.46
Sample Size	479641	479641	479641	479641	479641	479641
Year FE	No	Yes	No	Yes	No	Yes
County FE	No	Yes	No	Yes	No	Yes
PM10						
Attainment Status	.231*	.186**	.257*	.185**	.366***	.171*
	[.120]	[.093]	[.148]	[.094]	[.141]	[.095]
Ln(Air Quality)			.339	-.095	.923	-.064
			[.746]	[.367]	[.699]	[.347]
Ln(Consumption)					1.183***	1.116***
					[.074]	[.083]
Ln(Leisure)					-.479***	-.317***
					[.085]	[.070]
R-squared	.37	.42	.37	.42	.42	.46
Sample Size	479641	479641	479641	479641	479641	479641
Year FE	No	Yes	No	Yes	No	Yes
County FE	No	Yes	No	Yes	No	Yes
Carbon Monoxide						
Attainment Status	.205*	.054	.193*	.052	.251**	.049
	[.116]	[.051]	[.117]	[.051]	[.115]	[.049]
Ln(Air Quality)			-.337	.446	-.133	.498
			[.712]	[.392]	[.788]	[.379]
Ln(Consumption)					1.129***	1.083***
					[.068]	[.083]
Ln(Leisure)					-.619***	-.322***
					[.091]	[.075]
R-squared	.38	.43	.38	.43	.42	.46
Sample Size	393960	393960	393960	393960	393960	393960
Year FE	No	Yes	No	Yes	No	Yes
County FE	No	Yes	No	Yes	No	Yes

Notes.–Sources: Census, EPA, NOAA. Columns 1 and 2 regress log property (household level) values on a county's attainment status under different criteria pollutants between 2005-2010; columns 3 and 4 add log air quality as a control; columns 5 and 6 add log non-durables consumption (2010 dollars) and leisure (hours) as controls. Controls on the household include educational attainment, number of children, indicator for disability, age, gender, number of bedrooms, year the house was built, household tenure, race fixed effects, and population. Economic controls include the percent of the (state) population that is civilian and in the labor force, employment, payroll expenditures, and the number of establishments within 9 different bins of firm size (1-4 employees, 5-0, 10-19, 20-49, 50-99, 100-249, 250-499, 500-999, 1000+). All standard errors are clustered at the county-level.

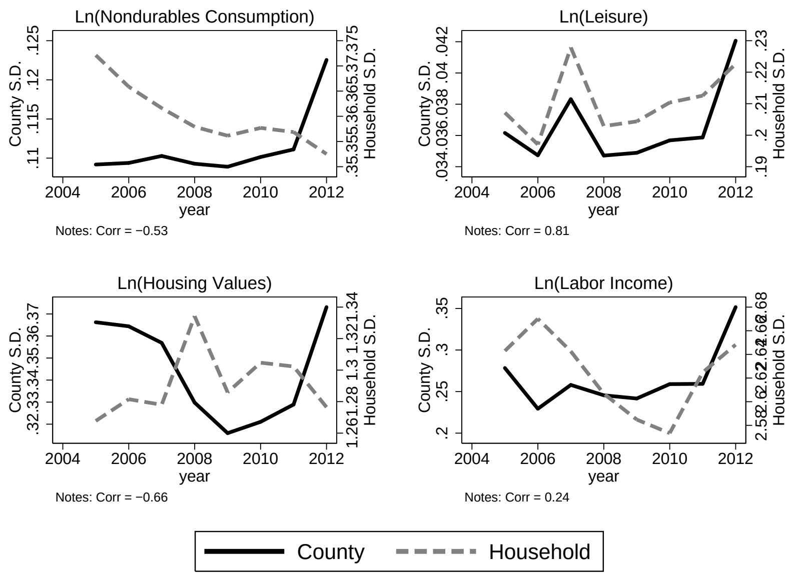

Figure 6: County versus Household Trends, 2005-2010

Notes.–Source: Census. The plots consist of the standard deviation of the ratio of log residual property values and earnings at the county to household level. After regressing log property values (and separately, log labor earnings) on controls, then the county measure is generated by averaging across individuals, and finally averaging across all individuals and county averages. Controls include: age, educational attainment, family size, number of children, age of the house, household tenure, state employment, establishments, and population. The aggregated statistics are weighted by Census county population weights.

4.3. Prices and Quantities

Applications of hedonic models frequently use property values, rather than the price of housing services, in order to estimate willingness to pay. Consistently estimating the marginal willingness to pay for air quality requires the assumption that the included controls remove all heterogeneity that is correlated with the measurement error between property values and the price of housing services (e.g., see Equations 4 and 5). If this does not hold, then using an aggregate expenditure as the dependent variable, like total labor earnings, introduces an additional premium that will tend to be negatively correlated with the right hand side variables (e.g., since higher time on the market mechanically implies lower leisure). To make the exercise as comparable with prior applications of hedonic models, I use (log) pollution, rather than the air quality index, and estimate variants of Equations 4 and 5, denoting "aggregate" and "price ratio" as the instances where the dependent variable is equal to total labor earnings (ratio of housing value to earnings) versus the hourly wage (ratio of the price of housing to the hourly wage).

Table 3 shows that the associations are dramatically different depending on the specification of the dependent variable. First, the coefficients on pollution are lower in magnitude (absolute value) when the dependent variable is in terms of the price ratio, rather than the full aggregate. Even after controlling for year, county, and industry fixed effects, the coefficient on pollution is 67%

lower in magnitude in column 4 of Panel A since hours worked is both positively correlated with pollution and earnings, thereby generating upwards bias. Similarly, the coefficient on pollution is 283% *higher* in magnitude in column 4 of Panel B since housing services will tend to be more strongly (negatively) correlated with pollution than labor services through the complementarity of leisure and air quality. (That is, time spent out work is typically in doors, so the correlation between housing and pollution is stronger than it is for hours and pollution.) Since the price ratio is increasing in housing services, this generates downwards bias.

Second, the signs on the market goods, including consumption, leisure, and housing are all qualitatively and quantitatively different depending on whether the dependent variable is an aggregate expenditure versus price ratio. Looking first at Panel A, the coefficient on leisure in column 2 is negative versus positive in column 4, whereas it is positive in column 2 and negative in column 4 for Panel B. These qualitative differences emerge from the fact that the quantity of labor services is necessarily and highly negatively correlated with leisure since the more hours an individual spends working, the more it trades off with leisure time. The bias for non-durables and housing services is less stark, but still very present. These qualitative and quantitative differences based on the choice of the dependent variable highlight the fact granular fixed effects and controls do not remove the endogenous feedback mechanisms that cause the hedonic price function to adjust and clear the market.

5. Using Intratemporal Variation to Estimate the Demand for Environmental Quality

5.1. Identification Problems

A naive application of OLS on Equation 6 will lead to biased estimates for four reasons. The first two are standard empirical challenges and can be resolved easily; the latter two require novel approaches. First, the constant terms—bundled in T_C and T_L—will induce a covariance between the error and (C, L, S) because of time trends. Second, because there is a large mass of individuals who work zero hours within any given year—due to, for example, voluntary or involuntary unemployment—my estimate of $\hat{\psi}$ will be downwards biased.

Third, the right hand side variables of interest household-level consumption-leisure and county-level air quality are endogenous objects that vary with transitory and permanent income shocks (Mian et al., 2013; Mian and Sufi, 2009, 2011, 2014). The Great Recession features a significant amount of reallocation across and within counties, causing wages, consumption, housing values, and leisure to fluctuate. Since unobserved productivity shocks to wages are negatively correlated

Table 3: Using Aggregate Expenditures versus Price Ratios in Hedonic Regressions

	Dep. var. = Aggregate		Dep. var. = Price Ratio	
	(1)	(2)	(3)	(4)
Panel A: Labor				
Ln(Ozone Pollution)	4.555***	1.754	-.788	1.145*
	[1.414]	[1.782]	[.716]	[.656]
Ln(Leisure)	-1.192***	-.421***	.986***	1.097***
	[.108]	[.100]	[.036]	[.036]
Ln(Nondurables)	2.962***	2.820***	1.903***	2.017***
	[.066]	[.066]	[.030]	[.033]
R-squared	.66	.71	.34	.35
Sample Size	577406	577406	415394	415394
Panel B: Housing/Labor				
Ln(Ozone Pollution)	-4.890***	.689	-.546	1.769*
	[1.413]	[2.269]	[.686]	[1.013]
Ln(Leisure)	2.433***	1.687***	-.913***	-1.096***
	[.115]	[.110]	[.039]	[.038]
Ln(Housing)	.931***	.980***	.260***	.264***
	[.009]	[.008]	[.005]	[.004]
R-squared	.64	.69	.12	.18
Sample Size	344642	344642	247760	247760
Year FE	No	Yes	No	Yes
Industry FE	No	Yes	No	Yes
County FE	No	Yes	No	Yes

Notes.–Sources: Census, EPA, NOAA. Columns 1 and 2 use log total labor earnings (top panel) and the ratio of log property values to total labor earnings (bottom panel), whereas columns 3 and 4 use log hourly wages and the ratio of log housing prices to hourly wages. The housing price is defined according to the household's selected annual owner costs and the hourly wage is defined as total earnings divided by total hours worked (both within a year). Controls on the household include educational attainment, number of children, indicator for disability, age, gender, number of bedrooms, year the house was built, household tenure, race fixed effects, and population. Economic controls include the percent of the (state) population that is civilian and in the labor force, employment, payroll expenditures, and the number of establishments within 9 different bins of firm size (1-4 employees, 5-0, 10-19, 20-49, 50-99, 100-249, 250-499, 500-999, 1000+). All standard errors are clustered at the county-level.

with leisure and air quality, but positively correlated with consumption and housing values, the coefficients on leisure and air quality will be downwards biased and on consumption and housing will be upwards biased. Merely using year fixed effects still leaves a great deal of time-varying endogenous variation reallocation at the county-level for the simple reason that the effects of the Great Recession were highly heterogeneous across labor and housing markets.

Fourth, households are not randomly assigned to locations with different levels of market and non-market goods. Rather, individuals with preferences for air quality will sort into areas with higher environmental amenities. While longitudinal data would enable me to incorporate individual fixed effects to purge variation driven solely by time-invariant productivity differences, risk attitudes and preferences are unambiguously changing over the Great Recession due to county-specific labor and housing shocks. Unobserved heterogeneity in workers' productivity can create a downward bias since more productive workers will receive systematically different wage offers and be correlated with leisure and risk preferences (Epple, 1987). Similarly, since people sort into residential communities and local labor markets based on their preferences for public goods (Rhode and Strumpf, 2003), including air quality (Banzhaf and Walsh, 2008), naive applications of least squares will attribute variation in wages to preferences for air quality when it is really driven by unobserved heterogeneity in tastes. Upward bias emerges since hotter climates are negatively correlated with air quality, and wealthier people favor temperate areas (Albouy et al., 2013).

Fifth, because of non-random sorting across households, the naive least squares estimators bundle both individual and group effects (Bayer and Ross, 2006; Bayer et al., 2007).[26] Although group fixed effects are often viewed as a solution (e.g., county fixed effects here), selective assortment induces a correlation between the group effects and other individual-level attributes. To make this idea concrete within the current application, re-write Equation 4 to illustrate the bundling of sorting across groups and individual behaviors.

$$\log w_{ij} = (\psi - 1)\log L_{ijt} + (1-\phi)\log C_{ijt} + (\phi - \psi)\log S_{jt} + \beta_1 X_{it} + \beta_2 G_{jt} + \eta_j + \varepsilon_i + \epsilon_{it}$$

where X contains all individual-level preference shifters, G contains all group-level preference

[26] For example, since hours are more volatile than wages over the business cycle (Heathcote et al., 2010a) because of sticky bargaining arrangements (Hall and Milgrom, 2008), controlling for unobserved heterogeneity in match quality is essential. In the absence of such controls, unobserved shocks to hourly wages will load onto the coefficient on leisure and cause it to be downwards biased: job-specific match quality is positively correlated with wages and negatively correlated with leisure since more productive workers and/or matches will face higher returns and rewards to working. Similarly, since both non-durables and housing consumption patterns are relatively clustered based on geographical location (Handbury and Weinstein, 2014), controlling for unobserved heterogeneity in group consumption behavior is equally as important.

shifters, ε_i is the individual-specific idiosyncratic error, and η_j is the j-group idiosyncratic error. Non-random sorting due to unobserved tastes implies that

$$Cov\left[(\beta_1 X_{it} + \varepsilon_{it})(\beta_2 G_{jt} + \eta_j)|X_{it}, \varepsilon_{it}\right] = \mathbb{E}\left[(\beta_1 X'_{it})\eta_j|G_{jt}, \varepsilon_{it}\right] > 0$$

The equation means that there is a positive correlation between unobserved group effects and observed individual-level covariates; the correlation is positive under the assumption that unobserved locational quality enters utility positively and satisfies the single crossing property.

5.2. Identification Solutions

First, to address the presence of growth rates and utility weights, I include year fixed effects to focus on year-to-year changes, (i.e., removing bias arising from the time-invariant values in Equations 4 and 5). Second, to address selection effects arising from an individual being observed with zero earnings when he is unemployed, I estimate a probit regression for employment with heteroskedastic robust standard errors (see the Appendix). Estimating the probit equation allows me to construct an estimate of the inverse Mill's ratio to characterize offered wages to those who remain in unemployment.

The remaining identification challenges—unobserved shocks to housing and labor markets, non-random sorting and unobserved tastes for air quality, and disentangling individual and neighborhood heterogeneity—all have two primary solution strategies.

A. Fixed effects and controls: The granularity of the Census micro-data allows me to include state-by-year, year-by-industry, and county fixed effects to focus purely on how within-county changes in air quality are related with changes in the marginal rate of substitution between air quality and market goods. These fixed effects are included on top of already parsimonious controls for county-level employment, establishments, payroll expenditures, and individual-level demographic / housing information (age, family size, marital status, gender, race, educational attainment, housing tenure, age of the home, and number of bedrooms). I also show that the results are robust to a semi-parametric measure of establishment closures (six bins containing the number of closures within a particular size category of establishments) in case the relationship between economic activity and factor prices is highly non-linear. These fixed effects also behave as a proxy for removing time-invariant sources of heterogeneity in individuals' tastes for air quality due to locational sorting of individuals into counties that have their optimal provision of public goods (Banzhaf and Walsh, 2008).

B. Instrumental variables: As discussed above, each of the independent variables of interest are potentially endogenous. I instrument air quality with a quadratic in wind speeds and their direction; non-durables consumption with the interaction between electricity expenditures and the age of the home; interactions between mean and median maximum and average temperature and age brackets (four bins); and, housing supply elasticity. The intuition behind the wind instruments is that, conditional on county unobservables, those with greater wind speeds will tend to be less polluted because the wind carries the pollution out towards a neighboring county. Since these wind speeds are correlated with other climatic attributes, I also control for precipitation. The quadratic terms for the climate variables capture potential nonlinearities (Schlenker and Walker, 2012; Hanna and Oliva, 2015).[27]

The intuition behind the electricity (a subset of non-durables consumption) and age of the home instruments is that the age of the home is related with its energy efficiency and aggregate utility payments of the owner. While the direct effects are certainly not exogenous, since the age of the home is correlated with the individual's tastes for housing and electricity could face a simultaneity problem, the interaction captures the plausibly exogenous shock to the cross-section of houses within a county.[28],[29] To the extent than an individual takes the aggregate changes in the market as exogenous, these shocks to the cross-section affect the set of houses with different degrees of energy efficiency and access to utility plans. The intuition behind the temperature interactions with age brackets arises from the impact of temperature on overall physical energy and alertness, whose effect varies significantly by age (Graff Zivin and Neidell, 2014). The intuition behind the housing elasticity follows along the lines of Saiz (2010) in that national shocks to the housing market will not affect housing prices as much in relatively supply constrained counties versus those that are unconstrained. Land constraints tend to decrease the housing elasticity, meaning that, to the extent that these land endowments are exogenous after conditioning on county fixed effects, the county-year-specific housing elasticity should be highly correlated with housing consumption. The Appendix provides detailed graphical and regression evidence for each of the first-stage relationships.

C. Group effects: Although the preceding instruments and fixed effects address the group effects identification problem—that time-varying unobserved factors can explain the same observed individual location decisions—I turn to a control function approach along the lines of Bayer and

[27] Aufhammer et al. (2013) caution the naive application of weather data. To summarize: (1) deviations around the mean temperature might be inaccurate, even if average temperatures are accurately constructed; (2) averaging across non-missing weather station data induces measurement error; (3) the correlation between weather variables varies across space significantly in sign and magnitude; (3) weather indicators are often spatially correlated because of the extrapolation methods used to create the measures, inducing collinearity.

[28] In practice, the elasticity of electricity with respect to income is quite low (Espey and Espey, 2004; Kilian, 2008).

[29] Landvoigt et al. (2015) use an assignment model with micro-data from San Diego county to characterize how housing prices adjust to assign houses that differ by quality to movers to differ by age, income and wealth.

Ross (2006) (inspired by Epple and Platt (1998) and Epple and Sieg (1999)). I generate county-level leave-one-out averages of housing values. After conditioning on observables, counties with relatively higher prices must be attractive for unobserved reasons, whereas those with low prices must be explained through the individual's unobservable (e.g., low taste for environmental amenities). Using a non-linear function of county housing prices allows me to non-parametrically control for the quality of county markets. While it introduces another potential source of endogeneity, since local housing prices are invariably correlated with consumption and labor shocks, in practice the bias seems to be small, if any. Robustness checks instrumented county property values using averages across similar groups of individuals as suggested by Bayer and Ross (2006).

D. Exclusion Restrictions: The exclusion restriction for the air quality instrument requires that unobserved shocks to wind speeds are mean independent of hourly wages—that is, wind speeds affect hourly wages only through the effect on air quality, which induces households to select into one county over another. The exclusion restriction for the consumption instrument requires that unobserved shocks to hourly wages are uncorrelated with electricity consumption within the set of houses built within the same year—that is, electricity consumption within houses of similar ages affects hourly wages only through non-durables consumption. Since these regressions will condition on county-level fixed effects, together with measures of topography, state-by-year, and industry fixed effects, there is no risk of the climate instruments absorbing endogenous variation correlated with economic activity through geography mechanisms. By conditioning on both county, industry, and state-by-year fixed effects, the elasticity of air quality is identified from county-industry specific deviations in weather around the county-industry averages and after controlling for all shocks common to counties within a state-year. Similarly, since previous literature finds that household's demand for electricity is inelastic (Reiss and White, 2005)—often not responding to price signals at all (Shin, 1985; Bushnell and Mansur, 2005)—unobserved shocks to wages are likely to cause households to cut back other forms of consumption, rather than electricity (or, water and gas).

D. Comparison to Other Methods: While the estimated elasticities are entirely novel contributions, another major advantage of my approach is that they can be used either independently or together with other external information (e.g., as sufficient statistics) in order to understand the welfare effects of policy intervention over public goods. In contrast, in the hedonic model, failing to account for adjustment in the hedonic price function biases the coefficients of interest since variation in the relationship between amenities and assets used to infer willingness to pay (e.g., housing markets) is loaded onto air quality (Kuminoff et al., 2010). The identification problem arises from the fact that an exogenous shock to the spatial distribution of a public good, like air quality, changes the gradient of the hedonic price function in order to clear the housing market, introducing a wedge between the average capitalization effect and household's willingness to pay. Rather than conflating willingness to pay for air quality with changes in the shadow price of air quality—represented through the price of consumption or labor—my structural model explicitly controls and instruments for them. Nevertheless, I also implement robustness checks where air quality and wind speeds are interacted with a linear time trend. The hedonic price function is linked with the valuation of environmental amenities, requiring some intertemporal link.[30]

E. Interpretation of Treatment Effects: Most quasi-natural identification strategies are challenged with the extent to which their estimates can be interpreted as average treatment effects (ATE) over local average treatment effects (LATE) (see Deaton (2009) and Heckman and Urzua (2010) for some critiques and Imbens (2010) for a defense). In particular, instrumental variables techniques tend to assume constant treatment effects in that the effect is identical for every observation. In the context of this paper, instruments are used to address the fact that counties with different air qualities are not assigned randomly. To the extent that the instrument introduces exogenous variation into the compliers group—counties (e.g., households residing in them) whose treatment status can be manipulated through the wind speeds instrument—these results may only identify the LATE. However, given the great deal of cross-sectional variation within and across counties over time, it is hard to imagine a scenario where the recovered coefficients are not average treatments. I also study the presence of heterogeneous treatment effects by income bracket, age, educational attainment, and exposure.

5.3. Main Results

To identify the elasticities on consumption and leisure, I estimate a system of simultaneous equations following from the first-order equilibrium conditions (Equations 4 and 5).[31] The price of housing services is proxied using a household's selected annual owner costs and housing consumption is imputed using the definition in Prescott (1997).[32] Given that the identifying assumption that

[30] Including a cubic time trend simply controls for unobserved shocks to hourly earnings—a similar candidate "fix"—it does not control for changes in the slope of the hedonic price function. Interacting air quality and the instruments with year dummies also failed; the estimated coefficients were too imprecise to have any meaning.

[31] Although the division bias (Borjas, 1980) is a common problem in labor studies, and is best addressed using another measurement of wages as an instrument, there are no alternative measurements of hourly wages available in the Census. Fortunately, given the sample size, any attenuation resulting from division bias does not have a quantitatively strong effect on my estimates given the sample size and quality of variation. Furthermore, measurement error in the dependent variable (wages are typically an independent variable) will just raise the standard error at worst.

[32]

$$H_{icts} = pr_{icts}\left[r_t^f + \tau_{icts}^p - \tau_t^m(m + \tau_{icts}^p) + \delta - \Delta p_{icts} + r_t^e\right]$$

where pr is the self reported property value, r^f is the risk-free rate, τ^p the imputed property tax rate, τ^m the marginal tax rate, m the mortgage rate (10-year average of 30-year fixed rate mortgage rate), δ is the depreciation rate, Δp is the capital gain (change in property value), and r^e the (equity) risk premium. Since the mortgage rate is unobserved shocks to individuals' marginal rates of substitution among consumption, leisure, and housing will, by construction, be correlated with the price ratios since they are an equilibrium outcome, identification comes from plausibly exogenous instrumental variables with fixed effects.

Table **??** documents these results. The elasticities of substitution are identified from within county-industry deviations in consumption, leisure, and air quality after adjusting for all the state-year specific shocks that are common to counties within the same state. Under the preferred specifications (columns 5-7), the results suggest that a 1% rise in leisure is associated with a 2-3% decline in hourly wages, a 1% rise in non-durables is associated with a .14-.2% rise in hourly wages, a 1% rise in air quality is associated with a 1.8-2.7% rise in hourly wages, and a 1% rise in housing services is associated with a precisely estimated .15% rise in the price ratio of housing to labor.[33] The naive least squares results produce highly biased estimates, generating incorrect signs in many cases (e.g., air quality is a substitute, not complement, with leisure). For example, whereas the IV results suggest that local labor markets capitalize non-market goods, reflecting by the positive association between air quality and wages, the OLS results predict a strong negative relationship. Similarly, the elasticities on consumption are quite large and outside the normal span of consumption dynamics for a household.

To my knowledge, the only other paper that has estimated a potentially comparable elasticity is West and Williams III (2007) between gasoline and leisure, finding evidence of complementarity between the two, although they do not advocate a particular mechanism behind the results. Since gasoline produces emissions, and thus negatively correlated with air quality, the result is evidence of substitutability between air quality and leisure. There are a variety of reasons that motivate the contrasting result. First, they exploit cross-sectional variation in labor supply and cross-state variation in gasoline prices. Because unobserved shocks to gasoline prices are time varying, and they have heterogeneous effects on workers' transitory incomes and local labor demands, they introduce

set according to a 10-year average, I use the 10-year return on a treasury bill from CRSP to obtain r^f. r^e is simply (.02 for now) the difference between the real return on stocks (equity) and the risk free rate.In particular, I use the S&P 500 index composite (monthly close value) as a measure of the real return to equity averaged out over the year. τ^p is imputed by taking the annual property taxes paid by the household divided by the self reported property value. From Harding et al. (2007), I set $m = .055$.While the ACS provides data on the mortgage payments, the actual rate is much more complicated than taking the ratio between the payment and the house. I use TAXSIM's average marginal tax rates to compute τ^m (Feenberg and Coutts, 1993). I set the depreciation rate according to best practices for annual data: $\delta = .025$. Finally, Δp is simply $(p_t - p_{t-1})/p_{t-1}$ for every household. (See Bieri et al. (2014) for spatial variation in the user cost of housing.) This measure of housing prices also is robust to constructing a price index defined by the coefficients on county fixed effects from a regression of log housing values (on controls); see Sieg et al. (2002).

[33]The implied elasticity on labor supply is above the traditional micro-elasticity (around .6 for full time males; Keane (2011)), but is explained by the convexification of labor supply generated through the inclusion of air quality. Just as in Imai and Keane (2004) who find an elasticity near four, since local labor markets also include amenities, then the selection of a location to work is tantamount to a choice on public goods. Since air quality is negatively correlated with hours worked, but positively correlated with the hourly wage, then the traditional elasticity is downwards biased.

Table 4: Structural Elasticity Estimates

	Least Squares			Instrumental Variables			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ln(Leisure)	1.01***	1.33***	1.59***	-.46***	-2.03***	-2.59***	-2.88***
	[.02]	[.02]	[.02]	[.13]	[.31]	[.66]	[.70]
Ln(Nondurables)	2.07***	2.22***	2.15***	.30***	.20***	.15**	.14**
	[.01]	[.01]	[.01]	[.03]	[.05]	[.07]	[.07]
Ln(Air Quality)	-3.08***	-3.55***	-3.74***	.16	1.83***	2.44***	2.74***
	[.02]	[.02]	[.02]	[.14]	[.34]	[.71]	[.75]
Ln(Housing)	.25***	.24***	.26***	.12***	.15***	.15***	.16***
	[.00]	[.00]	[.00]	[.02]	[.02]	[.02]	[.02]
R-squared	.35	.36	.40	.24	.19	.16	.15
Sample Size	274267	274267	274267	263925	165364	165364	165364
Year FE	No	Yes	Yes	Yes	Yes	Yes	Yes
State FE	No	Yes	No	No	Yes	No	Yes
Industry FE	No	No	No	Yes	Yes	Yes	Yes
County FE	No	No	Yes	Yes	No	Yes	Yes

Notes.–Sources: Census, EPA, NOAA. The table reports the coefficients associated with a simultaneous equations regression where the first regresses log hourly wages (earnings/hours worked) on log non-durables consumption, leisure, air quality, with controls / fixed effects, and the second regresses the log price ratio between housing and labor (selected annual owner costs / hourly wage) on log housing services, leisure, and air quality (with the same fixed effects). Housing services are defined according to Prescott (1997), non-durables are defined according to Attanasio and Weber (1995), and leisure is defined according to the first definition in Aguiar and Hurst (2007). Column 1 implements the naive OLS; column 2 adds state and year fixed effects; column 3 replaces state with county fixed effects. Columns 4-7 instrument the endogenous regressors. Column 4 instruments for consumption, leisure, and housing with fixed effects on year, industry, and county; column 5 also instruments for air quality, but only with fixed effects on year, industry, and state; column 6 replaces the state fixed effects with county; column 7 adds state by year fixed effects. Leisure is instrumented using interactions between mean/median max and average temperatures and age brackets; housing consumption is instrumented using interactions between a leave-one-out average of the county property tax rate and the housing supply elasticity; non-durables is instrumented using interactions between electricity expenditures and five bins for the year an individual's house was built; and, air quality is instrumented using wind speeds. The coefficients correspond to $(\psi - 1)$ for leisure, $(1 - \phi)$ for consumption, $(\zeta - 1)$ for housing services, $(\phi - \psi)$ for air quality. Controls on the household include educational attainment, number of children, indicator for disability, age, gender, number of bedrooms, five bins for the year the house was built, household tenure, race fixed effects, and population. Economic controls include the state labor force, employment, payroll expenditures, and the number of establishments within 9 different bins of firm size (1-4 employees, 5-0, 10-19, 20-49, 50-99, 100-249, 250-499, 500-999, 1000+). Consumption and leisure are imputed from the CEX and ATUS. All standard errors are clustered at the county-level.

endogeneity. The negative correlation with both consumption and labor supply/earnings induces upwards bias and makes the elasticity between air quality and leisure appear more substitutable. Second, they instrument for earnings using occupation, state, and gender -specific means, real income using an alternative price index, and gasoline prices using national-average gas prices and gasoline refinery outages. However, the variation in these instruments does not match the variation needed to identify household-level valuations of environmental amenities (or gasoline).[34]

These conclusions remain relatively robust to additional checks. First, non-parametrically controlling for county-level time varying economic shocks through the inclusion of different bins indicating the number of establishment closures of a particular size (e.g., 1-10 employees, 10-20, and so on) should imply similar estimates. Second, I test for a pre-trend across counties using a flexible estimator operationalized by Mora and Reggio (Forthcoming).[35] After conditioning on county-level measures of economic activity and county fixed effects, there is no significant pre-trend. Third, and very interestingly, to the extent that employment at the county-level is a broad indicator of a county's economic trajectory, regressing it on the number of establishments and payroll expenditures yields an R^2 of .9927. Evidently, the R^2 cannot rise much further. Adding in county-specific year trends only raises the R^2 to .9999, meaning that these proxies are successfully absorbing the relevant county-specific time varying economic shocks.

5.4. Validity of the Exclusion Restriction

Many economic forces were present during the Great Recession. While it is highly unlikely that counties with different climate variabilities in wind speeds and temperature were on differential trends, such a threat to identification would induce bias. Tests of overidentifying restrictions provide information under the assumption of a null hypothesis that the model is correctly specified and overidentifying restrictions are valid. Because the test is implemented by only excluding some of the instruments, the remainder are assumed to be exogenous to test the subset of excluded instruments. To assure that my elasticities are not driven by unobserved variation in weather outcomes, I examine the plausibility of the exclusion restriction directly and, second, introduce another set of instruments that generate similar results.

While these are explicitly considered in the Appendix, I summarize the basic checks here. First, I show that proxies for economic shocks (establishment closures, employment, pay, and hours at the county-level) are uncorrelated with each of the instruments. Since the only threatening violation to

[34] They also use a shorter time series (1996-1998) with much lower sample size and fewer controls.

[35] Mora and Reggio (2012) show that recovering causal effects in difference in difference estimators actually requires more than just a parallel path condition, but also similar time trends. They introduce a flexible form for allowing for various types of trends—not just linear—and show that results in many prior papers are not robust to this threat to identifying assumptions.

the identifying assumptions of the model is that time varying economic shocks to economic activity are in some way correlated with time varying unobservables. Each of the four proxies of economic factors are uncorrelated with the instruments. Second, estimating Equation 4 and introducing the instruments as additional controls yields statistically insignificant coefficients associated with each of the instruments. The lack of statistical significance implies that—even with a large sample size—the instruments are unlikely to be correlated with omitted variables since they do not have any meaningful relationship with hourly wages (or the price ratio of housing to labor). Third, the distribution of hourly wages is very similar above and below the median level of the given instrument of interest, meaning that the distribution of unobservables is likely to be uncorrelated with the instruments. Taken even further, conditioning on controls and plotting the distribution of residuals implies a nearly identical overlap. Fourth, there is no correlation between two proxies for locational sorting (housing values and commuting time from home to work) with the instruments. To the extent that time-varying unobservables, these results imply that they are unlikely to be endogenous time-varying unobservables since they are uncorrelated after conditioning on fixed effects and observable controls.

5.5. Interpretation of Treatment Effects

Quasi-natural identification strategies are ultimately susceptible to concerns articulated by Heckman and Urzua (2010) and Deaton (2009) that the estimated parameters are only local average treatment effects. Importantly, there is also a literature specific to hedonic methods that cautions against the interpretation of results when linear approximations to the hedonic price function are used (Ekeland et al., 2004; Heckman et al., 2010). The standard approach in the literature—which they show has severe drawbacks—involves computing linear approximations to the first order conditions implied by a utility-maximizing consumer subject to a budget constraint that nests the hedonic price function. These approaches are typically justified through the rationale of clean instrumental variable or quasi-natural experiment strategies. Kuminoff et al. (2010) implement a series of monte carlo experiments and show that many of the conventional results are not robust to functional form, and tend to induce significant bias when the linear approximation to the hedonic price function cannot capture the movement to a new market equilibrium.[36]

These theoretical insights prompt concern about the interpretation of the results thus far and the extent to which heterogeneous treatment effects may bias towards the interpretation of local average treatment effects (LATE). The purpose of this section is to provide evidence of identification

[36] After a policy intervention, the market will obtain a new equilibrium that shifts the hedonic price function (Heckman and Vytlacil, 2007). Kuminoff and Pope (2014) shows that this feature affects implied capitalization effects in hedonic models.

that allow results to have an average treatment effects (ATE) interpretation.[37]

First, as Ekeland et al. (2004) emphasize, multimarket data (e.g., different regional housing markets) is essential. To the extent that my results provide an unbiased estimate of the underlying preference parameters, which are common across agents across markets, my model still allows the distribution of individual heterogeneity to vary across markets since the identifying source of variation arises from cross-market differences in prices and locational choice. Furthermore, when interacting the quadratic in wind speeds with a household-level indicator for whether the house uses electricity (rather than gas) for heating, the estimated parameters are nearly the same. The equivalence between these two cases reflects the fact that there is sufficient cross-sectional variation for identifying an ATE.

Second, the typical assumption required to obtain an ATE with instrumental variables is that treatment effects are constant across sub-populations ("identification at infinity").[38] The condition required for extrapolating from subpopulations to achieve an ATE interpretation is that compliers, never-takers, and always-takers are not found to differ substantially in levels with respect to the outcome variable. In the Appendix, I provide plots of the wage distribution for counties above and below the median value of the corresponding instrument (e.g., wind speeds); as expected, the distributions do not meaningfully differ, suggesting that an ATE interpretation suffices.

Third, Section 7 is devoted to understanding the degree to which subgroups have different preferences over environmental amenities across different brackets of income, age, schooling, and exposure to air quality levels. To understand the extent to which these results are informative about aggregate elasticities of air quality, I test for the extent to which heterogeneous treatment effects change my results using a procedure developed by Xie et al. (2012).[39] While certain groups unsurprisingly have more or less tastes for air quality, the qualitative results remain.

5.6. Robustness without Structural Assumptions

While the exclusion restrictions of the instruments withstand various robustness exercises, I now turn to reduced-form tests of non-separability and elasticities that relaxes two main assumptions

[37]Bayer et al. (2007) introduce an identification strategy for obtaining causal estimates of local amenities through property values when households are heterogeneous by exploiting discontinuities along the boundaries of geographies. To the extent that housing and geographical attributes are more or less continuous throughout metropolitan areas, variation in consumption and leisure within counties with discontinuously different wages or housing values—based on their distance from a specified county or state boundary—would identify the capitalization of air quality.

[38]For an interpretation of LATE, the already discussed identification arguments are sufficient, namely independence (wind speeds is as good as randomly assigned and do not directly affect hourly wages, the outcome), random assignment, the exclusion restriction (wind speeds are uncorrelated with unobserved shocks to wages), and monotonicity ("no defies"). The monotonicity condition requires that wind speeds affect counties uniformly, which is achieved by conditioning on topography, other weather related variables, and county fixed effects.

[39]Crump et al. (2009b) develop a fully nonparametric test for treatment effects of heterogeneity.

and follows more closely in the tradition of the quasi-experimental literature (Chay and Greenstone, 2003, 2005; Greenstone, 2002; Greenstone et al., 2013). First, I abstract from the constant elasticity of substitution assumption that was present in the utility function. Second, I avoid having to instrument for leisure and consumption by exploiting only the exogenous variation in year-to-year wind speeds and their correlation with air quality. Specifically, I consider regressions of some outcome variable (consumption, leisure, housing) on air quality with the following instrumental variables specification

$$\begin{aligned} \log S_{cst} &= \alpha_1 + \theta f(Z_{cst}) + \beta_1 X_{icst} + \nu_{icst} \\ \log y_{icts} &= \alpha_2 + \varepsilon_X \widehat{\log S_{cst}} + \beta_2 X_{icst} + \epsilon_{icst}, \quad \forall y \in \{C, L, H\} \end{aligned} \tag{10}$$

where X denotes the usual vector of controls and $f(Z)$ denotes a quadratic vector of instruments consisting of wind speed measurements. The elasticities, ε_X, can be transformed to map into the model parameters as follows by letting $\omega = 1 - (1/\varepsilon_X)$ $(\varepsilon = 1/(1-\omega))$ where $\omega \in \{\phi, \psi, \zeta\}$ under the assumption that the error now includes unobserved variation in leisure, consumption, and/or housing depending on the outcome variable in Equation 10 and the coefficient on air quality is scaled by a constant. The identifying assumption is that changes in unobserved shocks to the outcome variable are uncorrelated with wind speeds after controlling for differences across counties and industries.

Table 5 documents these results. The results are qualitatively aligned with the elasticities estimated from the structural model. Under the preferred specification (column 8), a 1% increase in air quality is associated with a 2.43% decline in non-durables, .09% increase in leisure, and 3% increase in housing services. These translate into values of $\phi = .86$, $\psi = -3.8$, and $\zeta = -.84$, which are quite close to the estimated elasticities in the structural model with the only exception that the coefficient on leisure is imprecisely estimated. The reason arises from the fact that identification is driven off of year-to-year changes in the impact of temperature on different age brackets, which is only changing exogenously based on the age composition of counties. To the extent that this distribution is only shifting slowly, the year-to-year changes are small and generate a somewhat weak first-stage. The interesting observation is that the naive least squares estimate produces quite inconsistent elasticities. In particular, regressing log housing consumption on air quality produces a negative coefficient, suggesting that better air quality actually reduces property values. The bias emerges from omitted variables—a major theme emphasized by Chay and Greenstone (2005). Adding fixed effects does not solve the identification problem—in fact, they merely attenuate the estimates, producing statistically and economically insignificant coefficients (columns 3 and 4).

Table 5: Reduced-Form Elasticity Results

	Dep. var. = Ln(Nondurables Consumption)							
	OLS	OLS	OLS	OLS	IV	IV	IV	IV
Panel A								
Ln(Air Quality)	-.35***	.29***	-.02	-.04	-.39	1.07***	-1.64***	-2.43***
	[.08]	[.09]	[.04]	[.04]	[.34]	[.33]	[.56]	[.51]
R-squared	.30	.33	.35	.41	.30	.37	.32	.36
Sample Size	636878	636878	636878	636878	338220	338220	338220	338220
	Dep. var. = Ln(Leisure)							
Panel B								
Ln(Air Quality)	-.11***	-.05***	.00	.01	-.05	-.08	.16	.09
	[.02]	[.02]	[.02]	[.01]	[.09]	[.06]	[.18]	[.17]
R-squared	.44	.44	.45	.62	.43	.61	.61	.62
Sample Size	636878	636878	636878	636878	338220	338220	338220	338220
	Dep. var. = Ln(Housing Services)							
Panel C								
Ln(Air Quality)	-.50	1.50***	.09	.09	-2.66	5.52***	3.11*	3.11*
	[.74]	[.24]	[.13]	[.13]	[1.69]	[1.38]	[1.84]	[1.84]
R-squared	.34	.42	.44	.45	.33	.42	.44	.44
Sample Size	381757	381757	381757	381757	231108	231108	231108	231108
Year FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
State FE	No	Yes	No	No	No	Yes	No	Yes
Industry FE	No	No	No	Yes	No	Yes	No	No
County FE	No	No	Yes	Yes	No	No	Yes	Yes
Semiparametric	No	No	No	Yes	No	No	No	Yes

Notes.–Sources: Census, EPA, NOAA, CEX, ATUS. The table reports the coefficients associated with reduced-form regressions of log non-durables consumption, leisure, and housing (separately) on log air quality and a vector of controls. The columns add various fixed effect specifications and controls. Columns 4-8 instrument for air quality using a quadratic in wind speeds. The "full controls" specifications contain a quadratic in employment, payroll expenditures, wage income, and a five-way bin of establishment closures (1-10, 10-49, 50-249, 500-999, 1000+). All standard errors are clustered at the county-puma level. Observations are weighted by the ACS person-level weights, and the puma-to-county allocation factor discussed in the Appendix.

5.6.1. Other Issues

A. Alternative Weather Instruments: The main orthogonality condition required to achieve unbiased estimates of air quality in the reduced form regressions is that household's consumption and leisure bundles are influenced by wind speeds only through air quality, conditional on controls and fixed effects. A reasonable alternative to using wind speeds as instruments might be precipitation. However, many alternative climate variables *do* influence household's locational choice and non-pecuniary utility. For example, areas that are subject to a lot of rain tend to be less desireable locations, holding all else constant. Indeed, temperature and snowfall, for example, make good instruments in the structural regressions that include leisure and housing.

B. Measurement of Pollution: The benchmark model has used a transformation over the EPA's recently created air quality index, whereas all of the literature thus far has exploited variation in different pollution criteria (e.g., TSPs). To the extent that households react to air quality and air quality alerts, rather than the mere presence of TSPs, identifying a credible elasticity might be even more of a complex task. To answer this question, replace S^{ω} with $P^{-\omega}$ ($\omega \in \{\phi, \psi, \zeta\}$) in Equation 1, where P denotes pollution (TSP). Because pollution is an economic bad, it is raised to the power of $-\omega$ for $\omega \in \{\phi, \psi, \zeta\}$ in order to characterize the fact that $1/P$ is decreasing in P—that is, higher pollution reduces the quality of consumption, leisure, and housing. The fact that the results are nearly identical is assuring that not only the transformation is reasonable, but also that the instruments are successful at identifying the causal effect.

C. Capitalization Effects: Kuminoff and Pope (2014) show that capitalization effects may not identify willingness to pay since policy interventions affect the shadow price of the public good. To the extent that the variation during the Great Recession changed the shadow price of air quality—which can be studied through a test they provide for the gradient of the equilibrium price function before and after that defines the capitalization effect—then the WTP still is insufficient to identify welfare consequences. There are two reasons this concern is not applicable here. First, my instrumental variables approach provides exogenous variation in air quality, meaning that the "policy change" is uncorrelated with the amenity of interest. Second, my model identifies structural elasticities, rather than a MWTP, which together can be used to produce a MWTP estimate. Third, all of my estimates are exploiting cross-sectional variation in choices among individuals within the same county. I also show in the Appendix that the elasticity on air quality is not statistically different when estimated separately by year.

D. Willingness to Pay v. Willingness to Accept: Amiran and Hagen (2014) caution that willingness to pay and willingness to accept—which are crucial for identifying the value of a non-market good—may differ, especially when exploiting cross-sectional variation since aggregation problems may arise. Determining the empirical relevance of the concern is likely to vary with the quality of the data (e.g., variation and richness of controls) and is the topic of a companion paper. Nonetheless, this approach provides serious advantages to the traditional hedonic approach. First, since matching site-specific amenities with housing unit and demographic records often occurs at different levels of aggregation, my approach avoids an errors-in-variables problem that would traditionally introduce noise and/or bias depending on the correlation of the errors with other independent variables. Second, while a standard challenge is addressing omitted variables concerns arising from unobserved heterogeneity in housing prices that are correlated with the amenity, my household level data is able to provide a much finer set of controls on local labor market outcomes and housing values.[40]

F. Long versus Short Run Elasticities: A recurring tension in macroeconomics is the identification of Frisch versus Hicksian elasticities (e.g. see Chetty (2012) for the former and Keane and Rogerson (2012) and Keane (2011) for the latter). While one concern with my results is that the use of county fixed effects removes most of the identifying variation, which would imply that my estimates are informative only for short-run elasticities, I pursue two complementary exercises. First, I compare the estimates with and without the inclusion of county fixed effects; the specification without county FEs includes state and year-by-industry FEs. The difference between the two characterizes the contribution of the across-county variation within a given state. Second, I introduce an additional instrument, a wage-housing elasticity, that characterizes the capitalization of asset prices into local wages by regressing log wages on log housing prices, conditional on controls. The intuition arises from the fact that individuals with tastes for air quality will sort into counties where the pass-through rate of asset prices to the labor market is relatively lower since housing prices are more reflective of non-market goods that are unrelated to the labor share. In the Appendix, I show that this instrument predicts air quality well, but is also surprisingly uncorrelated with consumption and leisure, providing a separate, and potentially more cross-sectional, source of variation to identify the elasticities and test for over-identifying restrictions.

5.7. Heterogeneous Preferences: Aristotle's Hierarchy of Needs

While there is considerable heterogeneity across households with respect to their willingness to pay for environmental quality, surprisingly little quantitative evidence exists. To characterize these differences and uncover potential mechanisms, consider the benchmark model—fixed effects on county, year, and the usual controls—and partition the population into different groups. Denoting $g \in G = \{g_1, g_2, ...\}$ as a vector of groupings (e.g., income categories), I estimate Equation 4 sepa-

[40] See Bajari et al. (2012) for an alternative approach that exploits rational expectations over the unobserved error.

rately for every different group, g, with fixed effects on county, year, and industry. Understanding the presence of heterogeneity is important for characterizing the sources of heterogeneity that are relatively more important for structural macroeconomic models.

Table 6 suggests that there are considerable differences in elasticities among households. First, there is greater complementarity between air quality and leisure among those individuals with less than $60,000 in annual household income. It is possible that poorer individuals also live in areas with lower air quality, thereby raising their marginal utility of leisure when air quality is good. There is some support for this story (e.g., regressing log annual earnings on log ozone pollution produces a coefficient of 10 with a *p*value equal to 0.00). Second, there are major state-dependent effects. For example, the signs on the coefficients for leisure and non-durables are effectively switched when estimating the regressions separately under good and bad air quality states, indicating the different sources of identifying variation for recovering stable preference parameters. Third, leisure is more of a substitute for air quality among more educated and younger workers, but consumption is more of a complement. This could be consistent with an Environmental Kuznets story where wealthy individuals treat consumption as a complement to non-market goods and there are non-homothetic effects. The rationale for lower complementarity with leisure among the more educated workers is that their opportunity cost of time is greater, so substitution effects might encourage working more over leisure.

6. Macroeconomic Implications

Using these estimated elasticities and willingness to pay for air quality, my results allow me to implement a welfare evaluation of changes in air quality, offering a point of comparison with the EPA results on the Clean Air Act Amendments. Higher levels of air quality are valued, but come at the cost of lower household real incomes since firms pass-through the cost of emissions in the form of higher prices (Fabra and Reguant, 2014). A common approach is to assume that preferences are homogeneous and linear with respect to air quality, making the marginal willingness to pay for air quality constant (Freeman, 1974). Unfortunately, the gradient of the hedonic price function provides only an average *marginal* willingness to pay for a one-unit change in air quality or pollution, rather than (total) willingness to pay through the identification of the entire hedonic price function. What's worse is the fact that the hedonic price function may shift over time based on unobserved shocks, undermining the mapping of MWTP estimates into WTP estimates both due to bias (mis-specification) and lack of identification of the new (shifted) hedonic price function (Kuminoff et al., 2010; Kuminoff and Pope, 2014). These purely reduced form methods are not

Table 6: Heterogeneity in Preferences for Air Quality

	by household income				by air exposure	
	Q1	Q2	Q3	Q4	Bad	Good
Panel A						
Ln(Leisure)	2.87***	1.88*	.91	-2.38	-1.13**	-4.64***
	[.94]	[1.02]	[.88]	[1.61]	[.58]	[1.10]
Ln(Nondurables)	-.23***	-.52***	-.63***	-.17	.26***	.02
	[.09]	[.11]	[.12]	[.18]	[.07]	[.12]
Ln(Air Quality)	-2.64***	-1.35	-.28	2.55	.87	4.63***
	[.99]	[1.07]	[.94]	[1.72]	[.62]	[1.19]
Ln(Housing)	.28***	.31***	.33***	.38***	.21***	.09***
	[.02]	[.02]	[.02]	[.04]	[.02]	[.03]
R-squared	.11	.09	.06	.11	.22	.04
Sample Size	19081	37931	52272	56080	99464	65900
	by years of schooling			by age		
	<HS	HS	College+	20-40	41-55	55+
Panel B						
Ln(Leisure)	1.37	.34	-3.41**	-4.29***	-2.76**	-.71
	[1.21]	[.74]	[1.34]	[1.03]	[1.09]	[1.35]
Ln(Nondurables)	.62***	.13*	.15	-.13	.35***	.35***
	[.13]	[.07]	[.13]	[.09]	[.13]	[.12]
Ln(Air Quality)	-1.99	-.47	3.27**	4.43***	2.41**	.37
	[1.25]	[.79]	[1.44]	[1.09]	[1.18]	[1.45]
Ln(Housing)	.29***	.23***	.25***	.12***	.11***	.15***
	[.02]	[.02]	[.03]	[.03]	[.03]	[.03]
R-squared	.25	.21	.08	.11	.16	.23
Sample Size	10799	91779	62786	55916	66807	42641

Notes.–Sources: Census, EPA, NOAA, CEX, ATUS. The table shows the coefficients associated with the estimation of the structural model separately by group. There are four (household) income quartiles: borrowers with $-29,997-36,000, middle earners with 36,001-64,800, upper class with 64,804-105,000, and wealth with 105,001-3,563,100. There are three age quantiles: 20-40, 40-55, and 55+ years old. There are three educational attainment quantiles: those without a high school degree, those with, and those with a college degree. There are two air quality quantiles: good and bad air quality consisting of index levels between 248-298 and 203-248, respectively. The same controls are includes as the benchmark model, as well as county, year, and industry fixed effects. All standard errors are bootstrapped with 50 replications.

helpful for policy analysis because they do not allow for changes in behavioral responses as a result of shifts in the hedonic price function.

6.1. Estimating Counterfactual Densities

Given the estimated micro-elasticities, I now compute how much households would be willing to pay to keep air quality at its 2005 level in 2010, e.g., solving for the Δ such that $u(C(1+\Delta), L, H, S)$ equals $u(\tilde{C}, \tilde{L}, \tilde{H}, S)$ where $\tilde{D}$ denotes the counterfactual distribution of variable $D \in \{C, L, H\}$ in 2010. While all these variables are observed in the 2005 and 2010 data, the problem with computing Δ directly (without computing counterfactuals) is that Δ will be driven by changes in C, L, and H, rather than just S. Following Dinardo et al. (1996), I compute the 2010 counterfactual densities and weighting the observed 2010 values appropriately.[41] These counterfactuals have the interpretation of the density that would have prevailed if attributes (X) stayed the same at their 2005 level and the relationship between the outcome variable of interest, $D \in \{C, L, H\}$, and economic returns are as observed in 2010.

Much like the treatment effects literature has emphasized (Rosenbaum and Rubin, 1983; Heckman and Robb, 1985), typically these counterfactual exercises imply a causal interpretation only under the assumption of conditional exogeneity, or selection on observables. For example, these densities assume that changes in the distribution of attributes did not affect the underlying structure of market goods and services. General equilibrium changes in the distribution of attributes on the outcome variable are not considered. However, at least as a first-order analysis, these counterfactual densities allow me to provide a structural interpretation on the welfare implications of changes in air quality. The inclusion of housing services allows me to include the health effects of changes in air quality since property values capitalize local health risks (Davis, 2004). In order to

[41] Dinardo et al. (1996) developed an initial approach for computing these counterfactuals when separating between two groups (e.g., union status). Bound et al. (2010) has recently applied a version of their counterfactual simulation in the context of the economics of education. Heckman and Vytlacil (2005) studies it in the context of estimating treatment effects. The counterfactual density is given by $f^{2010}_{2005}(\ln D) =$

$$\int f^{2010}(\ln D|X)h(X|t = 2005)dx = \int \left(\rho^{2005}(X)/(1 - \rho^{2005}(X))\right) \left(P^{2010}/P^{2005}\right) f^{2010}(\ln D|X)h(X|t = 2010)dx$$

where P denotes the proportion of observations in the corresponding years, ρ denotes the predicted probability of a county being a 2005 observation, given the observed distribution of characteristics X. ρ^{2005} can be estimated by running a probit model with X controls on the distribution of attributes within a county, including both household-level attributes and business patterns (e.g., employment and establishments). Controls include an individual's property value, labor earnings, bins for the number of closures of plants within a particular size (1-10, 10-49, 50-249, 250-499, 500-999, and 1000 or more employees), establishments, payroll, and employment. The counterfactual density is weighted using

$$\Phi(X) = \left(\rho^{2005}(X)/(1 - \rho^{2005}(X))\right) \left(P^{2010}/P^{2005}\right)$$

map these counterfactual distributions into a policy exercise comparable to the EPA's evaluations of the CAAA, I will scale the WTP appropriately.

6.2. Calibration

Letting $\psi = -3.8$, $\phi = .86$, and $\zeta = -.84$ from the benchmark specification, the three remaining parameters to calibrate are the α's. Since $\alpha = .48$ typically in log-log preferences of the form $\alpha \log C + (1-\alpha)\log L$ in business cycle models, I keep $\alpha_L = .52$ and set α_C and α_H to match the share of expenditures on non-durables and housing durables, respectively, between 2005-2010. The Bureau of Economic Analysis (BEA) national accounts implies that 60% of private expenditures on non-durables and housing consumption is spent on housing, meaning that $\alpha_C = .48 \times .4 = .192$ and $\alpha_H = .288$.[42] While $\mu = .4608$ and $\gamma = .5136$ are calibrated from Makridis (2014) to match features of the U.S. economy between 1970-2010, setting a reasonable value for π is trickier. However, since the implicit argument in the hedonics literature is that air quality is capitalized into housing values through its effect on land, then I defer to Davis and Heathcote (2005) who calibrate land's share in new housing to .106, meaning that $\pi = 1 - .106 = .894$.[43] Under this calibration, then the willingness to pay can be computed closed form.[44]

6.3. Results

How can we use these elasticities and densities to understand the costs and benefits of changes in non-market and market goods? Specifically, what do these estimates imply about the EPA's evaluation of the Clean Air Acts? Average annual TSP were 29.2 μ/m^3 (micrograms per cubic meter) in 1990, 22.15 μ/m^3 in 2005, and 17.78 μ/m^3 in 2010 (Smith, 2012), amounting to declines of 24% and 19.7%, respectively. This is equivalent to a 39% decline in TSP and a comparable 4% rise in the air quality index from 1990 to 2010. Solving for Δ under the preferred elasticity estimates $(\psi, \phi, \zeta) = (-3.8, .84, -.86)$ under the counterfactual densities for consumption, leisure, and housing, together with a 4% rise in air quality from 2005, implies that $\Delta = .18$. In other words, households would be willing to accept 18% of current consumption in order to be made

[42] I am not including other durables, such as health and financial services. From the BEA national accounts, I include all units in the non-durable goods category (food/beverage, clothing/footwear, gasoline/energy, other) and transportation, recreation, and food services as a composite of non-durables.

[43] Letting H^{adj} denote the air quality adjusted housing services from the CES aggregation with air quality, an alternative strategy is to realize that the share of income (utility) from housing, call it κ, is equal to: $\kappa = (\partial H^{adj}/\partial H)(H/H^{adj})$, or $\kappa = \pi(H/H^{adj})^{\zeta}$. Given data on rental rates on κ, then π can be estimated.

[44] Letting $F = F(\tilde{C}, \tilde{L}, \tilde{H}, \tilde{S}, L, H, S)$ denote a constant consisting of the 2010 counterfactual (and, for air quality, actual) and 2005 leisure/housing/air distributions

$$\Delta = \left\{ \left[\exp(F/\alpha_C)^{\phi} - (1-\mu)S^{\phi} \right] / \mu \right\}^{\frac{1}{\phi}} / C - 1$$

indifferent between their 2005 bundle of market and non-market goods and counterfactual 2010 market goods with the level of air quality that prevailed. Since aggregate consumption expenditures totaled approximately \$8,790 billion in 2005, this means that the comparable net benefits of the Clean Air Act Amendments amounted to -\$1,582 billion. These estimates would imply that the CAAA would fail a cost-benefit test, contrasting heavily with their estimated \$83 billion in net benefits. To be clear, these results are very preliminary and undergoing additional robustness to the empirical strategy of simulating counterfactual densities.[45] Nonetheless, it is important to understand that these results do not imply a negative WTP, but rather that the CAAA led to too large of reductions in pollution at the expense of additional consumption and labor distortions. Figure 7 characterizes the sensitivity of willingness to pay based on different assumptions about the leisure and consumption elasticities with respect to air quality.

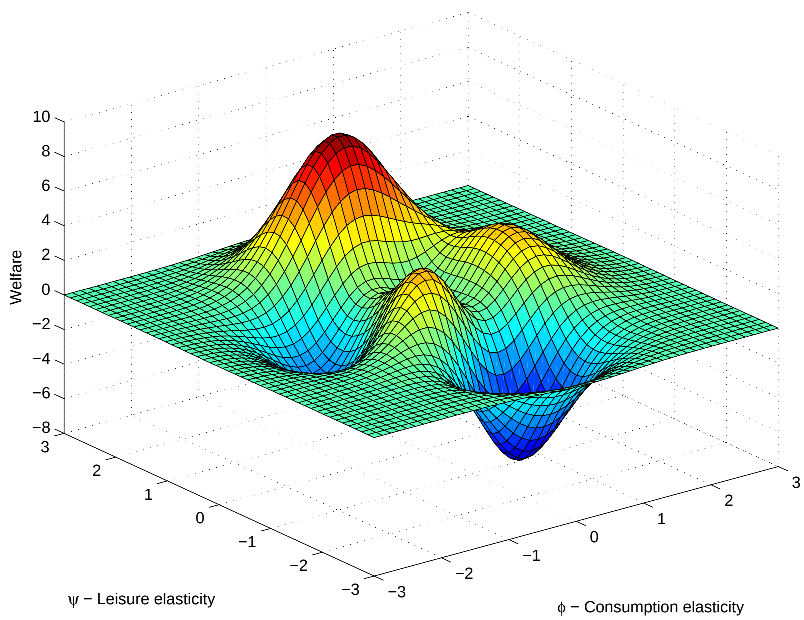

Figure 7: Implied Willingness to Pay between 2005 (actual) and 2010 (counterfactual)
Notes.–Sources: Census, EPA, NOAA, CEX, ATUS. The table plots the implied consumption equivalent that would make the average individual indifferent between the actual 2005 observed levels of consumption, leisure, housing, and air quality, and the counterfactual levels in 2010 had the distribution of observables stayed the same since 2005. The consumption equivalent amounts to computing Δ such that $u(C(1+\Delta), L, H, S)$ equals $u(\tilde{C}, \tilde{L}, \tilde{H}, S)$ where $\tilde{D}$ denotes the counterfactual distribution of variable $D \in \{C, L, H\}$ in 2010.

[45]While the same identification strategy could be used to obtain comparable measures of the value of a statistical life, doing so would require additional data on deaths at an industry-county level. However, doing so is important since nearly 90% of the EPA's estimated benefits of the CAAA are driven by a crude measure of the VSL—in particular, a measure that overestimates the preferred \$2-3 million range in the literature (Mrozek and Taylor, 2002) with approximately \$6 million. On top of the contrast from the best practices in the literature, Smith et al. (2003) emphasize that current measures of the VSL do not take into account endogenous labor supply, which biases the underlying wage/job risk combinations on a few orders of magnitude.

7. Conclusion

This paper has developed the most comprehensive database to date on individual-level consumption, leisure, housing, and demographic outcomes matched with weather and pollution outcomes at a county-level in order to structurally estimate elasticities between air quality and market goods and services (consumption, leisure, housing). Building on the canonical framework of Rosen (1974) and Roback (1982), I estimate a model where individuals choose where to live and where to work based on their preferences for both market and non-market goods. Importantly, I allow non-market goods, specifically air quality, to affect the marginal value of market goods. The estimated elasticities between air quality and consumption, housing, and leisure are 7.1, .54, and .2 ($\psi = -3.9$, $\phi = .86$, and $\zeta = -.84$), implying complementarity between housing and leisure with air quality, but substitutability with consumption. Since the parameters are derived directly from the equilibrium conditions of a utility-maximizing general equilibrium model, they naturally inform parameterizations of macroeconomic models used for policy evaluation, like those in Makridis (2014), Carbone and Smith (2008), and Carbone and Smith (2013). Aside from subjecting my results to various robustness exercises, I also find similar elasticities from reduced-form regressions that exploit only the plausibly exogenous variation in year-to-year wind speeds.

These elasticities jointly identify a willingness to pay for air quality through the computation of a Lucas consumption equivalent. Specifically, I ask: how much would an individual be willing to pay (or accept) in order to be made indifferent between their 2005 bundle of consumption, leisure, housing, and air quality and their counterfactual 2010 bundle, taking air quality in 2010 as given and simulating counterfactual distributions for consumption, leisure, and housing according based on their 2005 covariates? By scaling 2010 air quality to make the rise in air quality between 2005-2010 comparable to the decline in TSPs between 1990-2010 (the time span of the Clean Air Act Amendment evaluation), I find that households would be willing to accept 18% of current consumption in order to be made indifferent between their 2005 bundle and the counterfactual 2010 bundle, holding the distribution of covariates fixed to their 2005 levels. In other words, the CAAA amounted to net benefits of -$1,582 billion.

There are three main reasons these results different from more common estimates of willingness to pay in the literature. First, using county-level aggregates of earnings and/or housing values washes away much of the heterogeneity that is inherent in the data. Depending on compositional changes and the nature of reallocation within-counties, it is possible to overstate willingness to pay. Second, omitting consumption and/or leisure effectively produces an omitted variables bias problem since market goods are implicitly chosen whenever an individual makes a choice about market goods. Although controlling for income would help address this problem, there is no suitable instrument for income satisfying the relevant exclusion restrictions within this context. Third, using aggregate labor earnings (housing values_ as the dependent variable, rather than the price ratio, introduces a correlation between the quantity of labor (housing services) that is necessarily correlated with right hand side variables. It is important to use the dependent variable that is implied from the equilibrium conditions in order to recover structural parameter estimates.

This paper provides many avenues for future research. First and foremost, disentangling heterogeneity in preferences over private goods/services and amenities matters and cannot be captured purely by examining county-level data. Refinements should incorporate Roy sorting in a more structured way (e.g., as suggested by DeLeire et al. (2013)) in order to confront the inherent endogeneity in standard hedonic wage-risk regressions.[46] Second, there are a variety of distributional issues that can be investigated using accessible software routines (e.g., Frolich and Melly (2010)). For example, the comprehensive database can be used to understand how the spatial distribution of air quality inequality has changed over time—that is, differential access to non-market and market goods. Third, and most importantly, these elasticities can be used in macroeconomic models used to understand the welfare and general equilibrium effects of environmental policy intervention (see my companion work (Makridis, 2014; Cai and Makridis, 2015)).

References

Aguiar, M. and M. Bils (2009): "Has Consumption Inequality Mirrored Income Inequality?" University of Rochester mimeo.

Aguiar, M. and E. Hurst (2007): "Measuring Trends in Leisure: The Allocation of Time over Five Decades," *Quarterly Journal of Economics*, 122, 969–1006.

Aguiar, M., E. Hurst, and L. Karabarbounis (2013): "Time use during the Great Recession," *American Economic Review*, 103, 1664–1696.

Ai, C. and X. Chen (2003): "Efficient estimation of models with conditional moment restrictions containing unknown functions," *Econometrica*, 71, 1795–1844.

Albouy, D., R. Kellogg, W. Graf, and H. Wolff (2013): "Climate amenities, climate change, and the American quality of life," *Journal of the Association of Environmental and Resource Economists, RR*.

Aldy, J. and W. K. Viscusi (2008): "Adjusting the value of a statistical life for age and cohort effects," *Review of Economics and Statistics*, 90, 573–581.

Altonji, J. G. (1986): "Intertemporal substitution in labor supply: Evidence from the micro data," *Journal of Political Economy*, 94, S176–S215.

Amiran, E. Y. and D. A. Hagen (2014): "Tests of the discrepency between willingness to accept and willingness to pay: A theoretical analysis," *Working paper*.

[46] Kniesner et al. (2012) take an alternate route with panel data and using first-differences to remove time invariant heterogeneity. However, as many have documented (Bound et al., 2001, 1994; Bound and Krueger, 1991), there are massive measurement error problems in the PSID and researchers need to weigh the costs and benefits of using panel data over small samples with cross-sectional data over large samples.

Attanasio, O. P. and G. Weber (1995): "Is Consumption Growth Consistent with Intertemporal Optimization? Evidence from the Consumer Expenditure Survey," *Journal of Political Economy*, 103, 1121–1157.

Auffhammer, M., A. M. Bento, and S. E. Lowe (2009): "Measuring the effects of the Clean Air Act Amendments on ambient PM10 concentrations: The critical importance of disaggregated analysis," *Journal of Environmental Economics and Management*, 58, 15–26.

Aufhammer, M., S. M. Hsiang, W. Schlenker, and A. Sobel (2013): "Using weather data and climate model output in economic analyses of climate change," *Review of Environmental Economics and Policy*, 7, 181–198.

Bajari, P., J. C. Fruehwirth, K. I. Kim, and C. Timmins (2012): "A rational expectations approach to hedonic price reressions with time varying unobserved product attributes: The price of pollution," *American Economic Review*, 102, 1898–1926.

Banzhaf, S. H. and R. P. Walsh (2008): "Do people vote with their feet? An empirical test of Tiebouts mechanism," *American Economic Review*, 93, 843–863.

Bayer, P., F. Ferreira, and R. McMillan (2007): "A unified framework for measuring preferences for schools and neighborhoods," *Journal of Political Economy*, 115, 588–638.

Bayer, P. and S. L. Ross (2006): "Identifying individual and group effects in the presence of sorting: A neighborhood effects application," *NBER working paper*.

Berry, S., A. Khwaya, A. Musalem, K. C. Wilbur, G. Allenby, B. Anand, P. Chintagunta, W. M. Hanemann, P. Jeziorski, and A. Mele (2014): "Structural models of complementary choices," *Marketing Letters*.

Bharadwaj, P., M. Gibson, C. A. Neilson, and J. S. Graff Zivin (2014): "Grey matters: Fetal pollution exposure and human capital formation," *NBER working paper*.

Bieri, D. S., N. V. Kuminoff, and J. C. Pope (2014): "The role of local amenities in the national economy," *Working paper*.

Blundell, R., X. Chen, and D. Kristensen (2007): "Semi-nonparametric IV estimation of shape-invariant Engel curves," *Econometrica*, 75, 1613–1669.

Blundell, R., L. Pistaferri, and I. Preston (2008): "Consumption inequality and partial insurance," *American Economic Review*, 98, 1887–1921.

Bockstael, N. E. and K. E. McConnell (2007): "Environmental and resource valuation with revealed preferences: A theoretical guide to empirical models," *Springer*.

Borjas, G. J. (1980): "The relationship between wages and weekly hours of work: The role of division bias," *Journal of Human Resources*, 15, 409–423.

Bound, J., C. Brown, G. J. Duncan, and W. L. Rodgers (1994): "Evidence on the validity of cross sectional and longitudinal labor market data," *Journal of Labor Economics*, 12, 345–368.

Bound, J., C. Brown, and N. Mathiowetz (2001): "Measurement error in survey data," *Handbook of Econometrics*, 5.

Bound, J. and A. B. Krueger (1991): "The extent of measurement error in longitudinal earnings data: Do two wrongs make a right?" *Journal of Labor Economics*, 9, 1–24.

Bound, J., M. F. Lovenheim, and S. Turner (2010): "Why have college completion rates declined? An analysis of changing student preparation and collegiate resources," *American Economic Journal: Applied Economics*, 2, 129–157.

Bushnell, J. B. and E. T. Mansur (2005): "Consumption under noisy price signals: A study of electricity retail rate deregulation in San Diego," *Journal of Industrial Economics*, 53, 493–513.

Cai, Y. and C. Makridis (2015): "Environmental Policy and Land Values: Evidence from the Housing Market in a Dynamic Stochastic General Equilibrium Model," *Working paper*.

Campos, R. G. and I. Reggio (2014): "Measurement error in imputation procedures," *Economic Letters*, 122, 197–202.

Carbone, J. and V. K. Smith (2008): "Evaluating policy interventions with general equilibrium externalities," *Journal of Public Economics*, 92, 1254–1274.

——— (2013): "Valuing nature in a general equilibrium," *Journal of Environmental Economics and Management*, 66, 72–89.

Chay, K. Y. and M. Greenstone (2003): "The impact of air pollution on infant mortality: Evidence from geographic variation in pollution shocks induced by a recession," *Quarterly Journal of Economics*, 118, 1121–1167.

——— (2005): "Does air quality matter? Evidence from the housing market," *Journal of Political Economy*, 90, 1257–78.

Chay, K. Y. and D. S. Lee (2000): "Changes in relative wages in the 1980s: Returns to observed and unobserved skills and black white wage differentials," *Journal of Econometrics*, 99, 1–38.

Chen, X. (2007): "Large sample sieve estimation of semi-nonparametric models," *Handbook of Econometrics*, 6, 5550–5588.

Chen, X., G. Hong, and A. Tarozzi (2008): "Semiparametric efficiency in GMM models with auxiliary data," *Annals of Statistics*, 36, 808–843.

Chen, X., H. Hong, and D. Nekipelov (2011): "Nonlinear models of measurement errors," *Journal of Economic Literature*, 49, 901–937.

Chetty, R. (2012): "Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply," *Econometrica*, 80, 969–1018.

Costa, D. L. and M. E. Kahn (2003): "The rising price of nonmarket goods," *American Economic Review*, 93, 227–232.

Crump, R. K., V. J. Hotz, G. W. Imbens, and O. A. Mitnik (2009a): "Dealing with limited overlap in estimation of average treatment effects," *Biometrika*, 96, 187–199.

——— (2009b): "Nonparametric tests for treatment effect heterogeneity," *Review of Economics and Statistics*, 90, 389–405.

Currie, J., J. S. Graff Zivin, J. Mullins, and M. Neidell (2014): "What do we know about short and long term effects of early exposure to pollution," *Annual Review of Resource Economics*, 6, 1–31.

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

Curtis, M. (2014): "Who Loses under Power Plant Cap and Trade Programs? Estimating the Impact of the NOx Budget Trading Program on Manufacturing Employment," *Working paper*.

Davis, D. R. and J. I. Dingel (2014): "The comparative advantage of cities," *NBER working paper*.

Davis, L. (2004): "The effect of health risk on housing values: Evidence from a cancer cluster," *American Economic Review*, 94, 1693–1704.

Davis, M. and J. Heathcote (2005): "Housing and the business cycle," *International Economic Review*, 46, 751–784.

Davis, O. A. and A. Whinston (1962): "Externalities, welfare, and the theory of games," *Journal of Political Economy*, LXX, 241–262.

Davis, S. and J. Haltiwanger (1992): "Gross Job Creation, Gross Job Destruction, and Employment Reallocation," *Quarterly Journal of Economics*, 107, 819–863.

Deaton, A. (2009): "Instruments of development: Randomization in the tropics, and the search for the elusive keys to economic development," *Proceedings of the British Academy, 2008 Lecture*, 162, 123–160.

DeLeire, T., S. Khan, and C. Timmins (2013): "Roy model sorting and nonrandom selection in the valuation of a statistical life," *International Economic Review*, 54, 279–306.

Diamond, P. A. and J. A. Mirrlees (1973): "Aggregate production and consumption externalities," *Quarterly Journal of Economics*, 87, 1–24.

Dinardo, J., N. Fortin, and T. Lemieux (1996): "Labor market institutions and the distribution of wages, 1973-1992: A semi Parametric approach," *Econometrica*, LXIV, 1001–1044.

Ekeland, I., J. J. Heckman, and L. Nesheim (2004): "Identification and estimation of hedonic models," *Journal of Political Economy*, 112, S60–S109.

Epple, D. (1987): "Hedonic prices and implicit markets: Estimating demand and supply functions for differentiated products," *Journal of Political Economy*, 95, 59–80.

Epple, D. and G. J. Platt (1998): "Equilibrium and local redistribution in an urban economy hwen households differ in both preferences and incomes," *Journal of Urban Economics*, 43, 23–51.

Epple, D. and H. Sieg (1999): "Estimating equilibrium models of local jurisdictions," *Journal of Political Economy*, 107, 645–681.

Espey, J. and M. Espey (2004): "Turning on the lights: A metaanalysis of residential electricity demand elasticities," *Journal of Agricultural and Applied Economics*, 36, 65–81.

Fabra, N. and M. Reguant (2014): "Pass through of emissions costs in electricity markets," *American Economic Review*.

Feenberg, D. and E. Coutts (1993): "An introduction to the Taxsim model," *Journal of Policy Analysis and Management*, 12.

Finkelstein, A., E. F. P. Luttmer, and M. J. Notowidigdo (2013): "What good is wealth without health? The effect of health on the marginal utility of consumption," *Journal of the European Economic Association*, 11, 221–258.

Flores, N. E. and R. T. Carson (1997): "The relationship between the income elasticities of demand and willingness to pay," *Journal of Environmental Economics and Management*, 33, 287–295.

Flores, N. E. and P. E. Graves (2008): "Optimal public goods provision: Implications of endogenizing the labor/leisure choice," *Land Economics*, 84, 701–707.

Freeman, A. M. I. (1974): "On estimating air pollution control benefits from land value studies," *Journal of Environmental Economics and Management*, 1, 74–83.

Frolich, M. and B. Melly (2010): "Estimation of quantile treatment effects with Stata," *State Journal*, 10, 423–457.

Gibson, J. (2002): "Why does the Engel method work? Food demand, economies of size and household survey methods," *Oxford Bulletin of Economics and Statistics*, 64, 341–359.

Glaeser, E. L. and M. G. Resseger (2009): "Complementarity between cities and skills," *NBER working paper*.

Goulder, L. H. and R. C. Williams III (2003): "The substantial bias from ignoring general equilibrium effects in estimating excess burden, and a practical solution," *Journal of Political Economy*, 111, 898–927.

Graff Zivin, J. S. and M. Neidell (2012): "The impact of pollution on worker productivity," *American Economic Review*, 102, 3652–3673.

——— (2014): "Temperature and the allocation of time: Implications for climate change," *Journal of Labor Economics*, 32, 1–26.

GREENSTONE, M. (2002): "The impacts of environmental regulations on industrial activity: Evidence from the 1970 and 1977 Clean Air Act Amendments and the Census of Manufacturers," *Journal of Political Economy*, 110, 1175–1219.

——— (2004): "Did the Clean Air Act cause the remarkable decline in sulfur dioxide concentrations?" *Journal of Environmental Economics and Management*, 47, 585–611.

GREENSTONE, M., O. DESCHENES, AND J. S. SHAPIRO (2013): "Defensive investments and the demand for air quality: Evidence from the NOx budget program and ozone reductions," *American Economic Review*.

GREENSTONE, M. AND T. GAYER (2009): "Quasi-experimental and experimental approaches to environmental economics," *Journal of Environmental Economics and Management*, 57, 21–44.

GROSSMAN, G. AND A. KRUEGER (1995): "Economic Growth and the Environment," *Quarterly Journal of Economics*, 110, 353–377.

GUOJUN, H. (2013): "The effect of air pollution on cardiovascular mortality: Evidence from the 2008 Beijing Olympic games," *Berkeley ARE JMP*.

GUVENEN, F. (2012): "Macroeconomics with heterogeneity: A practical guide," *Working paper*.

GUVENEN, F. AND A. SMITH (2014): "Inferring labor income risk and partial insurance from economic choices," *Econometrica*, forthcoming.

HAINMUELLER, J. AND Y. XU (2013): "ebalance: A stata package for entropy balancing," *Journal of Statistical Software*, 54.

HALL, R. E. AND C. I. JONES (2007): "The Value of Life and the Rise in Health Spending," *Quarterly Journal of Economics*, 122, 39–72.

HALL, R. E. AND P. R. MILGROM (2008): "The Limited Influence of Unemployment on the Wage Bargain," *American Economic Review*, 98, 1653–1674.

HALL, R. E. AND F. MISHKIN (1982): "The Sensitivity of Consumption to Transitory Income: Estimates from Panel Data on Households," *Econometrica*, 50, 461–481.

HANDBURY, J. AND D. E. WEINSTEIN (2014): "Goods prices and availability in cities," *Review of Economic Studies*, forthcoming.

HANNA, R. AND P. OLIVA (2015): "The effect of pollution on labor supply: Evidence from a natural experiment in Mexico City," *Journal of Public Economics*.

HARDING, J. P., S. S. ROSENTHAL, AND C. F. SIRMANS (2007): "Depreciation of housing capital, maintenance, and house price inflation: Estimates from a repeat sales model," *Journal of Urban Economics*, 61, 193–217.

HEATHCOTE, J., F. PERRI, AND G. L. VIOLANTE (2010a): "Unequal we stand: An empirical analysis of economic inequality in the United States, 1967-2006," *Review of Economic Dynamics*, 13, 15–51.

HEATHCOTE, J., K. STORESLETTEN, AND G. L. VIOLANTE (2010b): "The macroeconomic implications of rising wage inequality in the United States," *Journal of Political Economy*, 118, 681–722.

——— (2014): "Consumption and labor supply with partial insurance: An analytical framework," *American Economic Review*, 104, 1–52.

HECKMAN, J. J., R. L. MATZKIN, AND L. NESHEIM (2010): "Nonparametric identification and estimation of nonadditive hedonic models," *Econometrica*, 78, 1569–1591.

HECKMAN, J. J. AND R. ROBB (1985): "Alternative methods for evaluating the impact of interventions," *Longitudinal Analysis of Labor Market Data*.

HECKMAN, J. J. AND S. URZUA (2010): "Comparing IV with structural models: What simple IV can and cannot identify," *Journal of Econometrics*, 156, 27–37.

Heckman, J. J. and E. Vytlacil (2005): "Structural equations, treatment effects and econometric policy evaluation," *Econometrica*, 73, 669–738.

Heckman, J. J. and E. J. Vytlacil (2007): "Econometric evaluation of social programs, Part I: Causal models, structural models and econometric policy evaluation," *Handbook of Econometrics*, 6, 4779–4874.

Hellerstein, J. K. and G. W. Imbens (1999): "Imposing moment restrictions from auxilary data by weighting," *Review of Economics and Statistics*, LXXXI, 1–14.

Imai, S. and M. P. Keane (2004): "Intertemporal labor supply and human capital accumulation," *International Economic Review*, 45, 601–641.

Imbens, G. (2010): "Better LATE than nothing: Some comments on Deaton (2009) and Heckman and Urzua (2009)," *Journal of Economic Literature*, 48, 399–423.

Keane, M. and R. Rogerson (2012): "Micro and macro labor supply elasticities: A reassessment of conventional wisdom," *Journal of Economic Literature*, 50, 464–476.

Keane, M. P. (2011): "Labor supply and taxes: A survey," *Journal of Economic Literature*, 49, 961–1075.

Kilian, L. (2008): "The economic effects of energy price shocks," *Journal of Economic Literature*, 46, 871–909.

Klaiber, A. H. and V. K. Smith (2012): "Developing general equilibrium benefit analyses for social programs: An introduction and example," *Journal of Benefit-Cost Analysis*, 3.

Kmenta, J. (1967): "On the estimation of the CES production function," *International Economic Review*, 8, 180–189.

Kniesner, T. J., W. K. Viscusi, C. Woock, and J. P. Ziliak (2012): "The value of a statistical life: Evidence from panel data," *Review of Economics and Statistics*, 94, 74–87.

Kopczuk, W. (2003): "A note on optimal taxation in the presence of externalities," *Economic Letters*, 80, 81–86.

Kuminoff, N. (2012): "Partial identification of preferences from a dual-market sorting equilibrium," *Econometrica, RR.*

Kuminoff, N., C. Timmins, and V. K. Smith (2013): "The new economics of equilibrium sorting and policy evaluation using housing markets," *Journal of Economic Literature*, 51, 1007–1063.

Kuminoff, N. V., C. F. Parmeter, and J. C. Pope (2010): "Which hedonic models can we trust to recover the marginal willingness to pay for environmental amenities?" *Journal of Environmental Economics and Management*, 60, 145–160.

Kuminoff, N. V. and J. C. Pope (2014): "Do "capitalization" effects for public goods reveal the public's willingness to pay?" *International Economic Review*, 55, 1227–1250.

Landvoigt, T., M. Piazzesi, and M. Schneider (2015): "The housing market(s) of San Diego," *American Economic Review*, 105, 1371–1407.

Lemieux, T. (2006): "Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill?" *American Economic Review*, 96, 461–498.

Makridis, M. (2014): "Environmental Policy and Structural Environmental Change," *Working paper.*

Maler, K.-G. (1974): "Environmental economics: A theoretical inquiry," *Baltimore: Resources for the Future.*

Mian, A., K. Rao, and A. Sufi (2013): "Household balance sheets, consumption, and the economic slump," *Quarterly Journal of Economics*, 128, 1687–1726.

Mian, A. and A. Sufi (2009): "The consequences of mortgage credit expansion: Evidence from the U.S. mortgage default crisis," *Quarterly Journal of Economics*, 124, 1449–1496.

——— (2011): "House prices, home equity-based borrowing, and the US household leverage crisis," *American Economic Review*, 101, 2132–56.

——— (2014): "What explains the 2007-2009 drop in employment?" *Econometrica*, 82, 2197–2223.

Mora, R. and I. Reggio (2012): "Treatment effect identification using alternative parallel assumptions," *Working paper.*

——— (Forthcoming): "dqd: A command for treatment effect estimation under alternative assumptions," *Stata Journal.*

Moretti, E. and M. Neidell (2009): "Pollution, health, and avoidance behavior: Evidence from the ports of Los Angeles," *Journal of Human Resources*, 46, 154–175.

Mrozek, J. and L. Taylor (2002): "What determines the value of life? A meta-analysis," *Journal of Policy Analysis and Management*, 21, 253–270.

Neidell, M. (2007): "Information, avoidance behavior, and health: The effects of ozone on asthma hospitalizations," *Journal of Human Resources*, 44, 450–478.

Neidell, M. and J. S. Graff Zivin (2013): "Environment, health, and human capital," *Journal of Economic Literature*, 51, 689–730.

Orazio, A. and L. Pistaferri (2014): "Consumption inequality over the last half century: Some evidence using the new PSID consumption measure," *American Economic Review.*

Perroni, C. (1992): "Homothetic representations of regular non-homothetic preferences," *Economics Letters*, 40, 19–22.

Pischke, J.-S. and H. Schwandt (2015): "Poorly measured confounders are more useful on the left than the right," *Working paper.*

Prescott, E. (1997): "On defining real consumption," *Review of the Federal Reserve Bank of St. Louis*, 47–53.

Reiss, P. C. and M. W. White (2005): "Household electricity demand, revisited," *Review of Economic Studies*, 72, 883.

Rhode, P. W. and K. S. Strumpf (2003): "Assessing the importance of Tiebout sorting: Local heterogeneity from 1850 to 1990," *American Economic Review*, 93, 1648–1677.

Roback, J. (1982): "Wages, rents, and the quality of life," *Journal of Political Economy*, 90.

Robins, J. M., A. Rotnitzky, and L. P. Zhao (1994): "Estimation of regression coefficients when some regressors are not always observed," *Journal of the American Statistical Association*, 89, 846–866.

Rosen, S. (1974): "Hedonic prices and implicit markets: Product differentiation in pure competition," *Journal of Political Economy*, 82, 34–55.

Rosenbaum, P. and D. Rubin (1983): "The central role of the propensity score in observational studies for causal effects," *Biometrika*, 70, 41–55.

Saiz, A. (2010): "Geographic determinants of housing supply," *Quarterly Journal of Economics*, 125, 1253–1296.

Sanders, N. J. (2012): "What doesn't kill you makes you weaker: Prenatal pollution exposure and educational outcomes," *Journal of Human Resources*, 47, 826–850.

Sandmo, A. (1975): "Optimal taxation in the presence of externalities," *Swedish Journal of Economics*, 77, 86–98.

——— (1980): "Anomaly and stability in the theory of externalities," *Quarterly Journal of Economics*, 94, 799–807.

Schlenker, W. and R. Walker (2012): "Airports, air pollution, and contemporaneous health," *Review of Economic Studies, RR.*

Shapiro, J. S. and R. Walker (2014): "Why is U.S. air quality improving? The roles of trade, regulation, productivity, and preferences," *Working paper.*

Shibayama, K. and I. Fraser (2014): “Nonhomothetic growth models for the environmental Kuznets curve,” *International Economic Review*, 55, 919–942.

Shin, J.-S. (1985): “Perception of price when price information is costly: Evidence from residential electricity demand,” *Review of Economics and Statistics*, 67, 591–598.

Sieg, H., V. K. Smith, S. Banzhaf, and R. Walsh (2002): “Interjurisdictional housing prices in locational equilibrium,” *Journal of Urban Economics*, 52, 131–153.

Smith, V. K. (2012): “Reflections–In search of crosswalks between macroeconomics and environmental economics,” *Review of Environmental Economics and Policy*, 6, 298–317.

Smith, V. K. and J. C. Carbone (2008): “Environmental economics and the "curse" of the circular flow,” *Frontiers in Resource and Rural Economics*.

Smith, V. K., S. K. Pattanayak, and G. L. Van Houtven (2003): “VSL reconsidered: What do labor supply estimates reveal about risk preferences?” *Economics Letters*, 80, 147–153.

Van Hee, V. C. and C. A. Pope (2012): “From Olympians to mere mortals: The indiscriminate, global challenges of air pollution,” *American Journal of Respiratory and Critical Care Medicine*.

Viscusi, W. (1979): “Employment hazards: An investigation of market performance,” *Cambridge: Harvard University Press*.

Viscusi, W. K. and J. E. Aldy (2003): “The Value of a Statistical Life: A Critical Review of Market Estimates throughout the World,” *Journal of Risk and Uncertainty*, 27, 5–76.

West, S. E. and R. C. Williams III (2007): “Optimal taxation and cross-price effects on labor supply: Estimates of the optimal gas tax,” *Journal of Public Economics*, 91, 593–617.

Xie, Y., J. E. Brand, and B. Jann (2012): “Estimating heterogeneous treatment effects with observational data,” *Sociological Methodology*, 42, 314–347.

Yagan, D. (2014): “Moving to opportunity? Migratory insurance over the Great Recession,” *Harvard JMP*.