Journal of Environmental Economics and Management 60 (2010) 145–160

Contents lists available at ScienceDirect

Journal of Environmental Economics and Management

journal homepage: www.elsevier.com/locate/jeem

Which hedonic models can we trust to recover the marginal willingness to pay for environmental amenities? ☆

Nicolai V. Kuminoff [a,*], Christopher F. Parmeter [b], Jaren C. Pope [c]

[a] Department of Economics, Arizona State University, Main Campus, PO Box 873806, Tempe, AZ 85287-3806, USA
[b] Department of Economics, University of Miami, Jenkins Building, Coral Gables, FL 33146-6500, USA
[c] Department of Economics, Brigham Young University, 127 Faculty Office Building, Provo, UT 84602, USA

ARTICLE INFO

Article history:
Received 14 September 2009
Available online 27 July 2010

Keywords:
Hedonic
Functional form
Omitted variable
Quasi-experiment

ABSTRACT

The hedonic property value model is among our foremost tools for evaluating the economic consequences of policies that target the supply of local public goods, environmental services, and urban amenities. We design a theoretically consistent and empirically realistic Monte Carlo study of whether omitted variables seriously undermine the method's ability to accurately identify economic values. Our results suggest that large gains in accuracy can be realized by moving from the standard linear specifications for the price function to a more flexible framework that uses a combination of spatial fixed effects, quasi-experimental identification, and temporal controls for housing market adjustment.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

The hedonic property value model is among our foremost tools for evaluating the economic consequences of policies that target the supply of environmental amenities. In theory, Rosen's [28] hedonic model of market equilibrium can be used to assess the welfare implications of changes in goods and services that are not explicitly traded in formal markets, but are conveyed through the choice of a home. Economists have relied on this model to measure the willingness-to-pay for improved air quality [30], to estimate the value of a statistical case of pediatric leukemia [11], to assess property value externalities from bacterial contamination of waterways [21], and to evaluate the benefits from numerous policies enacted at the local, state, and federal levels. Examples include the Clean Air Act [8], the Superfund program for cleanup of hazardous waste sites [23], and disclosure laws for airport noise [27]. Widespread use of the hedonic model for policy evaluation makes it especially important to understand the method's strengths and limitations.

Over the past few years, there has been increasing concern that misspecification of the hedonic price function can seriously undermine its ability to accurately estimate economic values. The main concern is omitted variable bias. Neighborhood characteristics that matter to households, but are not observed by the econometrician, are often expected to be correlated with the amenity of interest or other independent variables [6,8,11,20,22,27]. While the omitted variable problem is quite general, it is challenging to evaluate its implications for hedonic estimates of nonmarket values because housing prices and consumer welfare are simultaneously determined as equilibrium outcomes of the market clearing process [28]. Cropper et al. [9] were the first to meet this challenge by developing a theoretically consistent framework for

☆ We thank Matt Kahn, Arik Levinson, Dan McMillen, Ray Palmquist, Dan Phaneuf, V. Kerry Smith, Laura Taylor, Chris Timmins, and two anonymous referees for helpful comments and suggestions on this research. We also thank seminar participants at the 2008 AAEA meetings, the 2009 AEA meetings, and the Triangle Resource and Environmental Economics Seminar. Kuminoff and Pope gratefully acknowledge research support from the Lincoln Institute of Land Policy.

* Corresponding author. Fax: +1 480 965 0748.
E-mail addresses: kuminoff@asu.edu (N.V. Kuminoff), cparmeter@miami.edu (C.F. Parmeter), jaren_pope@byu.edu (J.C. Pope).

0095-0696/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jeem.2010.06.001

computing hedonic equilibria. One of their key findings was that flexible econometric specifications for the equilibrium price function performed best when all variables were included in the model, but simpler functional forms (linear, log–linear, log–log, and linear Box–Cox) performed best in the presence of omitted variables. This result was widely applied in the subsequent empirical literature. The vast majority of hedonic studies published over the last two decades have used linear, log–linear, or log–log models in an effort to hedge against the risk of omitted variable bias.

In the 20 years since Cropper, Deck, and McConnell's innovative study (henceforth CDM) econometric methodology in the hedonic literature has evolved to address omitted variables directly. In studies with large cross-section data sets, it has become standard practice to add spatial fixed effects to the hedonic price function (e.g. fixed effects for school districts, subdivisions, or census tracts). The intention is to absorb the price effect of spatially clustered omitted variables. Likewise, the increasing availability of repeated cross-section and panel data sets has allowed researchers to develop quasi-experimental designs that offer the potential to purge the price effect of time-constant omitted variables [11,22,27]. Surprisingly there has been no systematic effort to evaluate the relative performance of these techniques in dealing with omitted variables and in recovering nonmarket values. We seek to fill this void.

This paper describes a theoretically consistent Monte Carlo evaluation of the accuracy of modern hedonic methods in recovering nonmarket values when key variables are omitted from the model. Five features of our analysis add realism and allow us to inform current hedonic practice: (i) we define the urban landscape using a comprehensive set of data on 104,000 housing sales that occurred in Wake County, NC during the late 1990s, (ii) these data are used to compute hedonic equilibria that reflect the real pattern of spatial correlation between housing characteristics and amenities, (iii) each Monte Carlo replication uses a sample size of 2000 homes, which is consistent with modern empirical studies and allows the use of spatial fixed effects, (iv) our evaluation of alternative hedonic methods focuses on the consequences of omitting a subset of neighborhood characteristics from the regression and we test the robustness of our results by randomly selecting which variables to omit, and (v) we simulate how housing markets adjust to the types of unexpected events that form the basis for quasi-experiments. For realism, our analysis is based on an actual change in commuting patterns that occurred in Wake County during our study period. This allows us to simulate a quasi-experimental design to evaluate pooled cross-section, first difference, and difference-in-difference approaches to estimation.

Our analysis yields several important results. First, adding spatial fixed effects to the estimated price function substantially reduces the bias from omitted variables in cross-section data. This reduction is much larger than the reduction that comes from using simpler functional forms. Moreover, the reduction is sufficiently large for "flexible" Box–Cox specifications that they outperform the linear, log–linear, and log–log models that currently dominate empirical practice. Second, difference-in-difference and first differenced estimation further reduces the bias from *time-constant* omitted variables, beyond the gains from spatial fixed effects. Third, we find that the scope for econometric misspecification expands substantially when using time-series data that include market adjustment to unexpected events. Our simulations characterize the movement between equilibria following a moderate change in commuting patterns. Simply pooling the data from the two equilibria leads to a bias in estimated MWTP that is similar in magnitude to the bias from omitted variables. This bias can be averted by using a generalized difference-in-differences estimator to control for changes in implicit prices.

2. Empirical specification and omitted variables in the hedonic literature

In his seminal 1974 paper, Sherwin Rosen demonstrated that the functional relationship between the price of a differentiated product and its attributes can be interpreted as an equilibrium outcome from the interactions between all the buyers and sellers in a market. Under the assumptions of his model, regressing product prices on their attributes can reveal consumers' marginal willingness-to-pay (MWTP) for individual attributes of a differentiated product. Importantly, the form of the equilibrium price function depends on the underlying distributions of preferences and technology. Under specific parametric assumptions about these latent distributions, such as Tinbergen's [33] linear–normal model, the equilibrium price function can take a convenient closed form. In general however, it is nonlinear without a closed-form solution. Moreover, Ekeland et al. [12] demonstrate that nonlinearity is a generic property of the price function. This means a linear functional form would be a special case in the sense that small perturbations to the underlying distributions of preferences and technology can produce large deviations from linearity.

While theory suggests the equilibrium price function is generally nonlinear, most empirical studies take linearity as a maintained assumption.[1] This practice is often justified by citing CDM's Monte Carlo analysis of how the accuracy in predicting MWTP varies across competing functional form assumptions. The distinguishing feature of their study is that it is theoretically consistent. They solve for an equilibrium vector of housing prices under specific assumptions about the parametric form of utility, the distribution of preferences, and the supply of housing. This allows them to compare the "true" MWTP for each housing characteristic (e.g. # bedrooms, square feet) with the econometric predictions made by each of six functional forms: *linear*, *semi-log*, *double-log*, *quadratic*, *linear Box–Cox*, and *quadratic Box–Cox*. When every housing characteristic that enters the utility function is included as an explanatory variable in the hedonic regression, the quadratic Box–Cox model produces the lowest mean percentage error in estimating MWTP. This result changes when one of the

[1] Due to space constraints, we do not work through the standard analytics of the hedonic property value model. Readers seeking an analytical review of the literature are directed to [25].

variables is omitted or replaced by a proxy. In this case, the more parsimonious functional forms – linear, semi-log, double-log, and linear Box–Cox – are the ones that perform the best.

CDM has served to guide the subsequent empirical literature. Their results are featured prominently in textbooks on environmental economics [14], reviews of the hedonic literature [25], and practitioners' guides to hedonic estimation [32]. Moreover, it has become common practice in the empirical literature to adopt one of the more parsimonious functional forms from CDM's study to represent the equilibrium hedonic price function with little or no specification testing. This is somewhat surprising because modern hedonic studies use data and econometrics that have little in common with the simulations designed by CDM.

In the 20 years that have passed since CDM's study, advances in microeconometric methods, together with the increasing availability of microdata, have changed the way hedonic models are estimated. To document these changes, we reviewed the 123 studies published between November 1988 and November 2008 that cite CDM according to the *Social Science Citation Index* (SSCI). In addition to empirical property value studies, this set of papers includes theoretical work and applications to markets for labor, breakfast cereal, fruit, automobiles, herbicides, knitted garments, ecolabeled paper towels, appliances, collectable coins, television, fish, forestry, and agricultural land. Narrowing the focus to residential property value studies decreased the size of our sample to 69 papers published in 35 different journals. While these 69 articles are a small fraction of a vast literature, they serve to illustrate two things: (i) how econometric methodology in the literature has evolved over time and (ii) how CDM's experimental design differs from the studies that invoke their results.

Table 1 compares the features of CDM to the papers in our SSCI sample.[2] 80% of the SSCI studies rely on one of the three linear functional forms: linear, semi-log, and double-log. Most of the others use a linear Box–Cox. Meanwhile, compared to CDM, the typical hedonic study uses a larger sample size, a broader definition for the housing market, and explicitly controls for variation in *unobserved* attributes across space and time.

As data on individual housing transactions have become increasingly available, sample sizes have increased. The median number of observations in the SSCI studies published during the past ten years (2459) more than quadrupled from the previous ten year period (593) which was nearly triple the number of observations used by CDM (200). As sample sizes have grown, so have the geographic and temporal boundaries used to define the market. CDM used data on homes in Baltimore City/County in 1977–1978. In comparison, 51% of the SSCI papers use data from multiple cities or counties and 71% use sales data over more than two years.

In their original analysis, CDM omitted two structural housing characteristics—*lot size* and the *number of rooms*. While concern about omitted variables has intensified since their study, the focus has shifted to omitted neighborhood characteristics. One reason is that data on structural characteristics have become readily available.[3] Perhaps more importantly, 84% of the studies in Table 1 focus on estimating the MWTP for a spatially delineated amenity such as air quality, open space, or proximity to hazardous waste sites. We often expect these amenities to be correlated with unobserved neighborhood characteristics [6,8,11,20,22,27].[4] Patterns of spatial correlation between included and omitted neighborhood characteristics may be very different from the spatial correlation between structural characteristics.

Over the past 20 years, the literature has also evolved to address omitted variables directly. More than half the studies in the SSCI sample use spatial fixed effects to absorb the effect of unobserved amenities that vary between cities or between neighborhoods within a city. A smaller share (14%) use spatial econometrics to impose more structure on the relationship between unobserved variables and housing prices. Finally, researchers are increasingly adopting quasi-experimental designs. While these studies only account for 20% of the SSCI sample, most were published recently.[5] These studies typically use fixed effects, first difference, or difference-in-difference estimators to identify the rates at which unexpected changes in amenities are capitalized into property values [8,11,27]. In some cases, repeat sales data are used to purge time-constant omitted variables. None of these strategies for addressing omitted variables were considered by CDM.

The bottom line is that the empirical hedonic literature that relies on the results from CDM has evolved to the point where it bears little resemblance to their original study. This evolution of the literature suggests it is time to revisit the question of empirical specification.

3. A theoretically consistent Monte Carlo experiment, calibrated to real world data

We design a Monte Carlo experiment to compute hedonic equilibria and investigate how the accuracy of predictions for MWTP varies with the econometric strategy for mitigating omitted variable bias. This section outlines our theoretical framework, model calibration strategy, and experimental design.

[2] Many of these studies report the results from multiple econometric models. We focus on the model that the authors identify as their main specification. If the authors do not identify a main specification, we focus on the model that produces the results that enter their discussion of policy implications and/or conclusions.

[3] County assessors are often required to keep detailed records of the structural characteristics and transaction price of every home sold in the county for tax purposes. This public information is collected and sold to researchers by commercial vendors such as *DataQuick*.

[4] For example, suppose we want to estimate the MWTP for air quality. If homeowners also care about crime rates, and industrial areas with low air quality tend to be located in high-crime areas, failing to control for crime rates will place an upward bias on the hedonic coefficient for air quality.

[5] We define a hedonic study as being quasi-experimental if there is an exogenous discontinuity (in time or space) that generates clear treatment group(s) and clear control group(s) which are used to estimate treatment effect(s). Many older studies satisfy these criteria without using the modern quasi-experimental terminology.

Table 1
Features of empirical hedonic studies: 1988–2008.[a]

	CDM	SSCI (69)
Functional form		
Main specification: lin–lin, log–lin, log–log		80%
Main specification: Box–Cox		17%
Sample size		
Median # observations	200	1917
Published in 1989–1998		593
Published in 1999–2008		2459
Distribution of studies by # observations		
0–200	•	6%
201–500		10%
501–1000		23%
1001–10,000		39%
More than 10,000		22%
Market dimensions		
Geography (#)		
Smaller than a city		7%
City or county	•	42%
Multiple cities or counties		42%
Nation		9%
Time period (#)		
0–1 year		7%
1–2 years	•	22%
2–5 years		26%
5–10 years		28%
More than 10 years		17%
Space and time		
Spatial error or spatial lag model	No	14%
Quasi-experimental identification	No	20%
Spatial variable of interest	No	84%
Includes temporal fixed effects	No	39%
Day, month, or quarter		14%
Year		25%
Includes spatial fixed effects	No	58%
Neighborhood		23%
City or county		29%
Region		6%

[a] SSCI (69) is the set of 69 property value hedonic studies that cite Cropper et al. [9] according to the Social Science Citation Index and were published between November 1988 and November 2008.

3.1. Theoretical framework

In the conventional hedonic property value model, housing services and amenities vary across the urban landscape. Let $j=1, \ldots, J$ homes be defined over a vector of characteristics, x_j, that includes structural characteristics of the home as well as spatially delineated amenities. Household i's utility depends on the characteristics of its home (x_j) and on its consumption of a composite numeraire good (c), given its idiosyncratic income (y_i) and preferences (α_i). Each household is assumed to choose a specific home and a quantity of c that maximizes its utility subject to a budget constraint:

$$\max_{j,c} U(x_j, c; \alpha_i) \quad \text{subject to} \quad y_i = c + p_j. \tag{1}$$

In the budget constraint, the price of the numeraire is normalized to one, and p_j represents annualized expenditures on house j.

A hedonic equilibrium is achieved when every household occupies a utility-maximizing location, given prices and characteristics, and the demand for housing equals supply. To formalize this assignment problem let b_{ij} denote household i's bid for the jth home and let A_{ij} be an indicator where $A_{ij}=1$ if and only if household i occupies that home. Then equilibrium can be defined as follows:

$$b_{ij} = \max_i \{b_{ij}\} \quad \text{iff} \quad A_{ij} = 1, \tag{2}$$

$$\sum_i A_{ij} = \sum_j A_{ij} = 1. \tag{3}$$

Each household occupies exactly one home, for which it has the maximum bid. Ties for the maximum bid are assumed to occur with zero probability.[6]

Under the assumption that utility is monotonic in the numeraire, bids can be expressed as a function of housing characteristics and preferences [28]. To see this, let $\tilde{u}$ be a reference level of utility and then invert the utility function to solve for $c = U^{-1}(x,\tilde{u};\alpha)$. Inserting $U^{-1}(\cdot)$ into the budget constraint and rearranging terms provides an expression for the household's maximum willingness to pay for each home:

$$b = y - U^{-1}(\tilde{u},x;\alpha). \tag{4}$$

Given a specification for utility and data on housing characteristics, the bid function can be used to compute a vector of prices and an assignment of people to homes that jointly satisfy the conditions for a hedonic equilibrium. Because there is generally no closed-form solution, the problem is solved numerically by iterating over adjustments to the price vector and the assignment of people to homes until (2)–(3) are simultaneously satisfied [19].

In general, there is no guarantee that hedonic equilibria will exist or be unique. Previous studies have not reported any difficulty in finding equilibria, but the issue of uniqueness has received little attention. CDM and other assignment studies have used algorithms that solve for the highest price equilibrium [2,9,17,19]. However, there may be additional equilibria at lower prices with the same assignment of people to homes, or different assignments. Kuminoff and Jarrah [19] recently developed an "iterative bidding algorithm" that can recover multiple equilibria from an assignment model, if they exist. Their application to the housing market in central California provided the first demonstration that assignment models generally support more than one equilibrium.

We would like to avoid having our conclusions about the relative performance of different estimators depend on an arbitrary choice for the equilibrium. This concern is reflected in three features of our experimental design. Firstly, we calibrate the assignment model to approximately reproduce actual prices in Wake County as the highest price equilibrium. Calibration helps to ensure that the relationships between housing prices, housing characteristics, and household income in our simulated equilibria will be similar to the actual relationships between those variables in our data. Secondly, we solve for a large number of different equilibria using randomly chosen subsets of our data. Using these data to calculate average measures of model performance should help to smooth over extreme results that may be tied to particular equilibria. Finally, for each randomly chosen subset of the data, we use the iterative bidding algorithm to solve for additional equilibria at lower prices. This allows us to test the sensitivity of our results on econometric performance to the choice among equilibria.

3.2. Calibrating the model to Wake County, North Carolina

The stock of housing is defined from data on approximately 104,000 single-family housing sales in Wake County, North Carolina between 1992 and 2000.[7] The data contain the sale price of each home, its structural characteristics, the characteristics of its neighborhood, and its proximity to open space and shopping centers. Table 2 provides summary statistics for these variables. The average home sells for approximately $201,000, it has 2.5 baths, is located on a half acre lot, has a fireplace but lacks a garage, has 1900 square feet of heated living space, is about 10 years old, and is located 4 miles from the nearest park and 8 miles from the nearest shopping center. The neighborhood attributes assigned to each home are defined by the census block in which the home is located. For the average home, median household income in the neighborhood is $68,000, median commute time is 23 min, and 27 percent of people living in the neighborhood are under the age of 18. The values for these attributes are based on the 1990 Census of Population and Housing.[8]

We use the Wake County data to compute separate hedonic equilibria for Cobb–Douglas, Translog, and Diewert utility functions.[9]

$$U_{ij} = \ln(c) + \sum_j \alpha_{ij} g(X_j) + \frac{1}{2}\sum_j \sum_k \beta_{jk} h(X_j) h(X_k),$$

[6] This implies houses and households must be unique. No two households can be identical because they would then have identical bids. Likewise, because identical houses would have identical equilibrium prices, a household assigned to one house would obtain the same utility from its twin and the occupants of both houses would have identical bids for each option. The requirement that households are unique is consistent with Rosen's [28] continuity assumption. The requirement that houses are unique is consistent with a landscape in which amenities are spatially delineated.

[7] These data have been used in several previous studies [26,27]. Fulcher [15] provides additional background on the data and how they were originally developed.

[8] Many variables are highly correlated. For example, distance to the nearest shopping center is highly correlated with commute time (ρ=0.77) and distance to the nearest park (ρ=0.73). Likewise, main heated living area is highly correlated with the presence of a garage (ρ=0.67) and the number of bathrooms (ρ=0.65). A full table of correlation coefficients is provided in the supplemental appendix together with maps of the region.

[9] In the Translog and Cobb–Douglas specifications, only the continuously varying characteristics are transformed. In other words, for the fireplace and garage dummy variables, we set $g(1)=1$ and $g(0)=0$.

Table 2
Summary statistics for Wake County data.[a]

Variable	Units	Mean	Std.	Min	Max
Price	$1,000	201	105	16	2976
Bathrooms	#	2.50	0.76	1.00	10.50
Acreage	#	0.50	0.93	0.01	97.52
Garage	Dummy	0.29	0.26	0.00	1.00
Fireplace	Dummy	0.91	0.36	0.00	1.00
Main heated living area	sqft (1000)	1.93	0.73	0.40	9.08
Age	Years	10.38	15.05	1.00	99.00
Median household income	$1,000	67.87	21.30	8.32	146.76
Median commute time	Minutes	22.71	4.49	7.00	37.00
% Under 18	%	26.77	5.18	2.15	49.84
Nearest park	Miles	4.34	2.84	0.41	18.59
Nearest shopping center	Miles	7.86	4.76	0.39	26.07

[a] *Median household income*, *median commute time*, and *% under 18* are defined for the Census block in which the home is located. *Nearest park* and *nearest shopping center* are uniquely defined for each home.

$$\begin{aligned}&\text{Diewert}: g(x)=h(x)=\sqrt{x},\\&\text{Translog}: g(x)=h(x)=\ln(x),\\&\text{Cobb-Douglas}: g(x)=\ln(x),\ h(x)=0.\end{aligned} \tag{5}$$

The Cobb–Douglas specification allows people to differ in their relative preferences for different housing characteristics (α_{ij}). This preference heterogeneity is maintained in the Diewert and Translog specifications, which also introduce interaction terms with constant preference coefficients (i.e. $\beta_{ijk}=\beta_{jk}$ for all i). In principle, the interaction coefficients could be allowed to vary across individuals. We held them constant in order to keep the model calibration tractable.

Each utility function was calibrated to the Wake County data by solving an optimization problem. The objective was to choose the distribution of preferences to minimize the distance between actual housing prices and the equilibrium hedonic prices predicted by the model. The heterogeneous α_{ij} parameters were assumed to follow independent gamma distributions, the β_{jk} interaction coefficients were all drawn from a common normal distribution, and income, which was assumed to be independent of preferences, was drawn from the distribution for Wake County reported in the 1990 Census data.[10] Given these distributional assumptions, the Nelder–Mead algorithm was used to choose values for the mean, variance, and gamma shape and scale parameters that allowed each utility function to approximately reproduce the Wake County data as a hedonic equilibrium.[11] The resulting parameter values were used to define the distribution of preferences for the population of households.

Fig. 1 contrasts the difference between the predicted and observed distributions of housing prices for 2000 randomly drawn households and homes based on the calibrated Diewert utility function. The solid line in panel A represents the empirical cumulative distribution function of actual prices in Wake County. The dashed line represents the equilibrium prices assigned to those homes in our simulation. Differences between actual prices and simulated prices reflect our inability to exactly reproduce the location choice process for homeowners in Wake County. This imperfection arises from the ways in which we abstract from reality, including: (i) the rigidity of the Diewert specification for preferences; (ii) our lack of data on school quality, crime rates, and other neighborhood characteristics that may influence prices; and (iii) measurement error arising from the fact that we assign 1990 Census block demographics to homes that actually sold between 1992 and 2000.[12] Despite these limitations, panel A of Fig. 1 illustrates that our model clearly reproduces the general price trends in Wake County. This is reinforced by the close match between the corresponding simulated and empirical probability density functions in panel B.[13]

[10] Independent normal and uniform distributions were also considered for α_{ij}. We chose the gamma over these alternatives because it provided the best fit of our model to the Wake County data. Our sampling procedure also assumes that people are uniformly distributed within each of the 16 income bins in the Census data. An interesting possibility for future research would be to generalize our preference specification to allow for correlation between preferences and income; correlation between preferences for different housing characteristics; or interactions between housing characteristics and the numeraire.

[11] This approach to model calibration differs from CDM. They defined intuitively plausible functional relationships between preference parameters and household demographics. While we could certainly allow preferences to vary systematically with demographic characteristics from the Census data, this would complicate our analysis without necessarily improving the fit of our model to the Wake County property sales data.

[12] Of course, imperfections in the calibration of our model should not affect our ability to evaluate the relative performance of hedonic estimators. When we use the calibrated model to compute a hedonic equilibrium, our preference specification becomes the "truth". Likewise, we maintain full control over the number of omitted variables and the degree of measurement error in the data used for econometric estimation.

[13] The calibrated Translog and Cobb–Douglas models perform similarly.

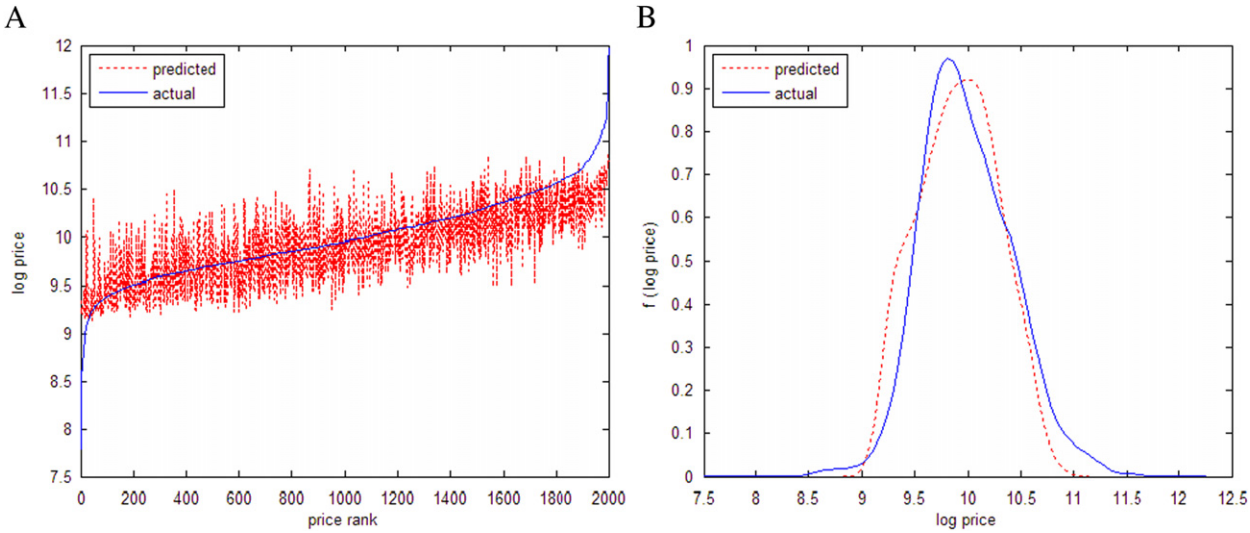

Fig. 1. Reproducing the empirical distribution of housing prices in Wake County. Panels A and B superimpose the equilibrium prices predicted by the calibrated Diewert utility function on the actual CDF and PDF for 2000 randomly selected housing sales. To facilitate comparison, we have flipped the axes on the empirical CDF and formatted the horizontal axis to display the ranking of homes by price.

3.3. *Experimental design: hedonic equilibria before and after changes in amenities*

We use the calibrated utility functions to simulate hedonic equilibria for our Monte Carlo experiments. Pseudo-data were generated by computing 100 pairs of equilibria for each of the three utility functions. Consider a single replication. We begin by drawing a random sample of households and homes. Then we solve for the assignment of households to homes and the vector of prices that jointly clear the market for housing. This is the "ex ante" equilibrium. Next, we shock a subset of neighborhood amenities and solve for the "ex post" equilibrium prices and housing assignments. The pseudo-data on housing prices and characteristics from the ex ante equilibrium are used to evaluate the alternative cross-section estimators. Data from both the ex ante and ex post equilibria are used to evaluate the panel data estimators.

Throughout the simulation process, the urban landscape is spatially explicit. When we draw a random sample of homes and solve for the equilibrium prices, the resulting pseudo-data reflect the actual spatial correlation between prices, neighborhood attributes, and amenities in Wake County.[14] This realism is important for assessing the consequences of omitting spatially delineated variables. We also use the latitude and longitude of each home to develop spatial fixed effects for census tracts and to generate weights matrices for estimating spatial lag and spatial error models [1].

The changes in amenities are also designed to be realistic. We focus on an actual event that occurred during our study period. In the late 1990s, the construction of a new beltline highway reduced commute times in selected areas of Wake County. With help from the Wake County planning department, we determined which census tracts would have experienced the largest reductions. Then we divided the tracts into four categories, each of which we associate with a constant percentage reduction in commute time: large reductions (50%), moderate reductions (30%), small reductions (15%), and very small reductions (5%).[15] Fig. 2 provides a map of these tracts in relation to the new highway. We evaluate the panel data estimators based on their ability to use these changes to identify MWTP for a further reduction in commute time.

One of the challenges for panel-data estimation is that a change in the amenity of interest may be correlated with changes in other amenities, some of which may be omitted from the model. The correlation can arise from a functional relationship between them (e.g. increased roadway congestion causes air quality to decline [18]). It can stem from the sorting mechanism that underlies the concept of a hedonic equilibrium (e.g. when polluting facilities move out of a neighborhood, wealthier households move in [4]). Or the amenities may be simultaneously generated by a deeper structural process (e.g. population growth increases congestion while decreasing air quality and open space [24]). To address these possibilities we simulate changes in three other neighborhood characteristics at the same time as we

[14] We do not use weights based on population density. Each sample is randomly drawn from the population of houses that were actually sold in Wake County during our study period. This procedure generates repeated samples of the type of "assessor data" that are routinely used to measure the MWTP for environmental amenities.

[15] These simulated changes are roughly in line with the changes implied by the 1990 and 2000 Censuses. Based on Census data, median commute times declined or were unchanged in 65 out of 105 tracts in Wake County. The largest decline reported in the Census was 59% compared to 50% in our simulation.

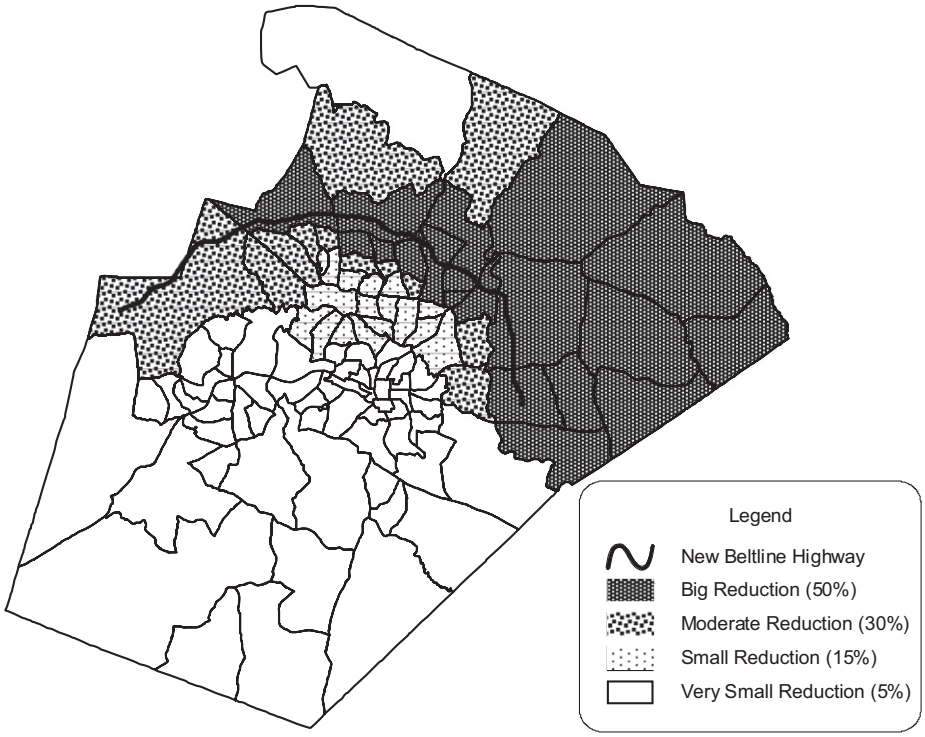

Fig. 2. New beltline highway and census tracts with reductions in commute time.

simulate the change in commute time. First we replace the 1990 Census values for *median household income* and *% under 18* in each census block with their year 2000 counterparts. Then we simulate a small change in the *nearest park* variable. Specifically, we used a GIS map of Wake County to choose several wooded areas to convert into parks and recalculated the distance from each home to the nearest park. On average, distance decreases by 22%. This type of land conversion could arise from a public program that raises money to preserve open space and build parks.[16] From an econometric perspective, the key implication of simulating park construction is that it induces a spatially consistent pattern of correlation between the change in commute time and the change in another continuously varying amenity that may be included or excluded from the econometric model. The correlation coefficient for the two sets of changes is 0.11.

In reality, neighborhood amenities change gradually. Some changes are easily anticipated, others are not. It is important to keep in mind that our Monte Carlo experiments abstract from dynamic decision making under uncertainty. The simulated changes in commute time, neighborhood demographics, and distance to parks are all defined to be strictly exogenous.[17] There are no moving costs, feedback effects, or other sources of friction.[18] We simply calculate market equilibria before and after an interval over which there were several changes to the spatial distribution of amenities. This setup mirrors the conceptual design of capitalization studies such as Chay and Greenstone [8] in which markets are observed at 10 year intervals, before and after nonmarginal changes to the amenity of interest. Thus, our evaluation of panel data estimators is conditioned by the assumption that the econometrician has identified a valid quasi-experiment. This is analogous to assuming that the econometrician working with cross-section data has correctly defined the extent of the market.

Using the pseudo data generated by the pre-shock and post-shock equilibria, we ran 54,000 regressions to evaluate 540 different hedonic models. The models varied according to the shape of the utility function (*Cobb–Douglas*, *Translog*, *Diewert*); the form of the price function (*linear*, *log–linear*, *log–log*, *Box–Cox linear*, *quadratic*, *Box–Cox quadratic*); the spatial controls for omitted variables (*no controls*, *fixed effects for census tracts*, *spatial error model*, *spatial lag model*);

[16] Wake County is one of several metropolitan areas to use bond initiatives to raise money to preserve open space [34].

[17] Development of Wake County's beltline highway took more than 20 years from designation to completion. While the eventual opening of the highway was not unexpected, some details of the planning and construction phases are consistent with a quasi-experimental design [31]. For example, we suspect it would have been difficult for homebuyers to predict the actual changes in commute times because these changes were endogenous to the behavior of other commuters and to the ongoing urban development process.

[18] While these assumptions are standard in the empirical literature, recent work has suggested that accounting for moving costs and imperfect information can affect hedonic estimates for nonmarket values [6,27]. In principle, one could extend our experimental design to investigate how these sources of friction affect estimates for MWTP. A related possibility for future research would be to evaluate the robustness of quasi-experimental approaches to hedonic estimation in the presence of weak instruments, endogenous sorting, and measurement error.

the approach to panel data estimation (*pooled cross-section*, *differences-in-differences*, *first differences*); the variables omitted from the regression (*none*, *three neighborhood amenities, three randomly chosen variables*); and the number of homes in the simulated market (*200*, *2000*).[19] Between 2 and 11 variables entered the price function in each regression, generating more than half a billion estimates for individual MWTP. With this multiplicity of models and wealth of estimates, finding a coherent way to synthesize the results is challenging. The next three sections provide what we think are sensible approaches.

4. Evidence on the performance of hedonic estimators for cross-section data

To quantify the performance of an individual estimator, we first calculate the difference between every household's MWTP for each housing attribute and the corresponding partial derivative of the estimated price function, $\hat{P}(x)$, on each Monte Carlo replication. The following equation defines this error, e_{ikr}, in estimating household i's valuation of characteristic k on replication r

$$e_{ikr} = \partial\hat{P}_r(x_i)/\partial x_k - MWTP_{ikr}. \tag{6}$$

We follow CDM by using this expression to construct summary statistics for the distribution of errors in estimating MWTP for the population of households in the simulated market. The following equation defines the normalized mean (β_{kr}) and standard deviation (S_{kr}) of the errors for each hedonic attribute on a given replication

$$\beta_{kr} = \frac{\overline{e}_{kr}}{N^{-1}\sum_i MWTP_{ikr}}, \quad S_{kr} = \frac{\overline{s}_{kr}}{N^{-1}\sum_i MWTP_{ikr}}, \quad k = 1,\ldots,K. \tag{7}$$

The normalized mean and standard deviation are simply the mean ($\overline{e}_{kr}$) and standard deviation ($\overline{s}_{kr}$) of the error from (6), divided by the average MWTP for characteristic k. Finally, averaging β_{kr} and S_{kr} over all 100 Monte Carlo replications summarizes the overall success of the model in estimating the MWTP for characteristic k.[20]

Table 3 reports the bias in our estimates for MWTP from each of CDM's six functional forms. The first two rows summarize the overall performance of each functional form in markets where equilibrium prices are generated by 2000 households with Diewert preferences and all variables are observed. Row 1 reports the mean absolute bias calculated over the 11 housing characteristics in Table 2. In the simple linear model, for example, the mean error is 40% of MWTP for the average Monte Carlo replication. Its relatively small standard deviation (6%) indicates that this statistic varies little as we draw different sets of houses and households over the 100 replications. As expected, the more flexible models perform better. The mean absolute bias for the quadratic Box–Cox model (13%) is one-third the bias of the simple linear model.

Comparing rows 1 and 2 reveals a bias-variance tradeoff. Increasing parametric flexibility in the price function tends to decrease the average bias in estimating MWTP, while simultaneously increasing its standard deviation *within* a given market. The linear model has the highest average bias and the lowest average standard deviation (1.55), whereas the quadratic Box–Cox model has the lowest average bias and the highest standard deviation (1.72). The standard deviation is important for "second-stage" hedonic studies that rely on variation in MWTP within a market to identify the demand for an amenity. However, for the typical hedonic study that aims to recover average MWTP, the mean absolute bias is the more relevant metric. Based on this criterion, the quadratic Box–Cox model outperforms all other functional forms, as in CDM.

To investigate the consequences of omitted variables, we begin by omitting three spatially delineated amenities: (i) median income in the census block, (ii) share of the population under 18 in the census block, and (iii) distance from the home to the nearest shopping center. Since these amenities are correlated with many of the variables that remain in the model, we would expect all of the estimators to perform poorly.[21] This is indeed the case. Mean absolute bias approximately doubles for the linear, semi-log, log–log, and linear Box–Cox functional forms. This can be seen by comparing rows 1 and 3 in Table 3. The increase in bias is even more extreme for the two quadratic specifications. In the Box–Cox quadratic case, the mean absolute bias increases fivefold! The normalized standard deviations on the error also increase for every functional form (moving from rows 2 to 4).

While the overall pattern of results in the first four rows is generally consistent with CDM, there is one important difference. The relatively large increases in mean absolute bias for the two quadratic specifications are not large enough to diminish their relative performance. They still perform about as well, or slightly better, than the simpler linear specifications in row 3. One explanation for this divergence from CDM is that we have increased the sample size by an

[19] The sample size of 2000 approximates the median number of observations (1917) in the 69 articles we reviewed. We also repeated the experiment for a sample size of 200. This is discussed in the context of our meta-analysis in Section 6.

[20] Other common measures of econometric performance, such as the mean square error and rejection rates for null hypotheses about model coefficients, will be highly correlated with the normalized mean and standard deviation. While these measures of overall performance are consistent with the usual objective of estimating average MWTP, they may be less applicable in special cases where the researcher aims to recover the MWTP for a subset of households clustered at an extreme point on the price function. In this case, one could use the data to calibrate an assignment model, like ours, and then use it to predict the relative performance of candidate estimators. This would be consistent with Klaiber and Smith's [17] proposal for the application-specific use of assignment models.

[21] A full matrix of correlation coefficients is provided in an appendix available at JEEM's online archive of supplementary material, which can be accessed at http://aere.org/journals/. Individual coefficients range from −0.46 for (*% under 18*, *age*) to 0.77 for (*nearest shopping center*, *time to work*). There are 14 pairings of (observed, unobserved) variables with correlation coefficients greater than 0.2 or less than −0.2.

Table 3
Summary statistics for the bias in cross-section hedonic estimates for MWTP[a] (means and standard deviations from 100 Monte Carlo replications, N=2000).

Row	Equilibrium	Utility function	Omitted variable scenario	Fixed effects	Summary statistic	Specification for hedonic price function					
						Linear	Semi-log	Log–Log	Box–Cox linear	Quadratic	Box–Cox quadratic
[1]	Wake calibration	Diewert	None	No	Average $\lvert\beta_j\rvert$	0.40 (0.06)	0.34 (0.05)	0.37 (0.03)	0.35 (0.03)	0.26 (0.02)	0.13 (0.02)
[2]	Wake calibration	Diewert	None	No	Average $\lvert S_j\rvert$	1.55 (0.03)	1.55 (0.03)	1.59 (0.04)	1.61 (0.04)	1.62 (0.05)	1.72 (0.06)
[3]	Wake calibration	Diewert	Spatial	No	Average $\lvert\beta_j\rvert$	0.81 (0.07)	0.71 (0.07)	0.78 (0.06)	0.76 (0.06)	0.64 (0.06)	0.71 (0.08)
[4]	Wake calibration	Diewert	Spatial	No	Average $\lvert S_j\rvert$	1.59 (0.04)	1.64 (0.04)	1.68 (0.04)	1.78 (0.05)	1.70 (0.09)	2.24 (0.13)
[5]	Wake calibration	Diewert	Spatial	Yes	Average $\lvert\beta_j\rvert$	0.43 (0.06)	0.38 (0.06)	0.40 (0.05)	0.41 (0.05)	0.28 (0.04)	0.26 (0.06)
[6]	Wake calibration	Diewert	Spatial	Yes	Average $\lvert S_j\rvert$	1.59 (0.04)	1.60 (0.04)	1.64 (0.04)	1.67 (0.04)	1.65 (0.07)	1.87 (0.11)
[7]	Wake calibration	Diewert	Random	Yes	Average $\lvert\beta_j\rvert$	0.50 (0.09)	0.46 (0.08)	0.42 (0.08)	0.38 (0.08)	0.35 (0.08)	0.31 (0.13)
[8]	Wake calibration	Translog	Random	Yes	Average $\lvert\beta_j\rvert$	0.56 (0.09)	0.52 (0.09)	0.25 (0.08)	0.23 (0.06)	0.42 (0.17)	0.31 (0.22)
[9]	Wake calibration	Cobb–Douglas	Random	Yes	Average $\lvert\beta_j\rvert$	0.66 (0.12)	0.61 (0.12)	0.47 (0.08)	0.42 (0.08)	0.49 (0.13)	0.35 (0.12)
[10]	Wake calibration	Diewert	Random	Yes	Max $\lvert\beta_j\rvert$	0.86 (0.20)	0.83 (0.13)	0.86 (0.54)	0.59 (0.40)	0.71 (0.16)	0.57 (0.48)
[11]	Wake calibration	Translog	Random	Yes	Max $\lvert\beta_j\rvert$	0.94 (0.16)	0.93 (0.15)	0.70 (0.14)	0.69 (0.15)	0.86 (0.41)	0.69 (0.73)
[12]	Wake calibration	Cobb–Douglas	Random	Yes	Max $\lvert\beta_j\rvert$	1.15 (0.26)	0.96 (0.20)	1.08 (0.27)	0.88 (0.29)	0.91 (0.27)	0.74 (0.31)
[13]	Lower price	Diewert	Spatial	Yes	Average $\lvert\beta_j\rvert$	0.42 (0.06)	0.37 (0.06)	0.41 (0.05)	0.36 (0.04)	0.29 (0.04)	0.27 (0.07)
[14]	Lower price	Diewert	Random	Yes	Average $\lvert\beta_j\rvert$	0.50 (0.09)	0.47 (0.08)	0.46 (0.09)	0.41 (0.07)	0.40 (0.08)	0.35 (0.14)
[15]	Lower price	Translog	Random	Yes	Average $\lvert\beta_j\rvert$	0.59 (0.09)	0.53 (0.09)	0.33 (0.08)	0.33 (0.06)	0.46 (0.11)	0.35 (0.07)
[16]	Lower price	Cobb–Douglas	Random	Yes	Average $\lvert\beta_j\rvert$	0.67 (0.12)	0.61 (0.12)	0.49 (0.08)	0.46 (0.08)	0.52 (0.13)	0.39 (0.12)

[a] Average $\lvert\beta\rvert$ is the absolute percentage bias in estimates for MWTP averaged over the 11 housing characteristics in Table 2. Max $\lvert\beta\rvert$ is the maximum percentage bias among the 11 characteristics. Average $\lvert S\rvert$ is the standard deviation of the bias averaged over the 11 characteristics. On each replication of the "spatial" omitted variable scenario three neighborhood amenities are omitted: (i) median household income, (ii)% under 18, (iii) distance to the nearest shopping center. On each replication of the "random" scenario, three variables are randomly chosen to be omitted. The lower price equilibria were obtained by multiplying the share of income spent on housing in the Wake County calibration by 0.5 in each household's reference consumption bundle.

order of magnitude relative to their study (from 200 to 2000). The increase in sample size should help to pin down the curvature of the nonlinear models.[22]

Rows 5 and 6 report the results from repeating the omitted variable regressions after adding spatial fixed effects. These fixed effects are simply dummy variables for census tracts. A random sample of 2000 homes on the typical Monte Carlo replication contained approximately 100 census tract dummies. These dummies are not perfect controls. They do not control for variation in omitted variables within census blocks. Nor do they control for variation between blocks within a tract. Nevertheless, adding them to the regression produces striking results. The spatial fixed effects nearly eliminate the bias in our estimated MWTP for every functional form. As a result, the quadratic and Box–Cox quadratic models outperform the simpler specifications by a considerable margin. Comparing rows 2, 4, and 6 reveals that the fixed effects also help to mitigate the effect of omitted variables on the standard deviation.

As a robustness check on the results from our "spatial" omitted variable scenario, we repeated the fixed effects regressions for all three utility functions and 100 different combinations of omitted variables. To avoid inadvertent bias in our experimental design, a uniform random number generator was used to select which three variables to omit on each Monte Carlo replication. The omitted variables were drawn from the full set of housing characteristics and amenities. Results for the mean bias are reported in rows 7–9.[23] The Box–Cox quadratic model has the lowest bias in the Diewert and Cobb–Douglas specifications, and the Box–Cox linear model has the lowest bias in the Translog specification. As a final measure of model performance, rows 10–12 report the maximum absolute bias in MWTP among the 11 housing characteristics. Again, the two Box–Cox specifications perform the best.[24]

The results in the first twelve rows of Table 3 are based on the equilibria we calibrated to approximate the distribution of actual housing sales in Wake County. However, we find that additional equilibria exist at lower prices. To test the robustness of our results, we repeated our analysis of model performance using 100 lower-price equilibria. These equilibria were obtained by repeating the iterative bidding algorithm [19] after decreasing the share of income spent on housing in each household's reference consumption bundle by 50% relative to the calibrated Wake County model. All other features of the model were held constant. Rows 13–16 summarize our findings. Measures of bias in the omitted variable scenarios are typically within a standard deviation of our baseline results for the Wake County model (rows 5, 7, 8, and 9). Moreover, there are no changes in our initial findings on the relative performances of different estimators.[25]

Overall, the results in Table 3 suggest that spatial fixed effects can "rehabilitate" flexible functional forms in the presence of omitted variables. This is important because the nonlinear models can substantially improve the accuracy of estimates for nonmarket values. For example, when omitted variables are randomly selected, moving from the popular semi-log specification for the price function to the Box–Cox quadratic model reduces the average bias in estimates for MWTP by 24–43%, depending on the type of equilibrium and the shape of the preference function.[26]

5. Evidence on the performance of hedonic estimators for panel data

Recent hedonic applications have increasingly sought to exploit temporal variation in the supply of environmental amenities. The basic idea is to estimate the rate at which unexpected changes in the spatial distribution of amenities are capitalized into housing prices. This approach requires data on equilibrium prices before and after a shock that "treats" a subset of homes in the market. With data on repeated sales of the same homes at multiple points in time, one can purge the bias that would otherwise arise from time-constant omitted variables.

By purging omitted variables, a panel data model can provide an accurate estimate for the capitalization rate. Yet, hedonic theory does not suggest that capitalization rates will provide accurate estimates for MWTP. Equality between the two concepts requires an additional assumption that the ex ante and ex post equilibria share the same hedonic price function [8,20]. This assumption is likely to be violated by a non-marginal shock to an amenity.[27] A second complication is the possibility of time-varying omitted variables. With both omitted variables and implicit prices potentially varying over

[22] An alternative explanation is advances in computing power and in the numerical algorithms used to solve for Box–Cox parameters. For example, CDM used an early version of Shazam software (ver. 6) that had a weak convergence tolerance (0.01 or 0.001) compared to current versions of Stata and other econometric software.

[23] Results for the normalized standard deviation are very similar to row 6 for the Diewert and Translog utility functions. For the Cobb–Douglas utility function, the Box–Cox quadratic model has the lowest standard deviation.

[24] The relatively large standard deviations on the quadratic forms in the Translog model in row 11 reflect a single replication where the maximum biases are 470% and 700%, respectively. Dropping this observation decreases the standard deviation from 0.41 to 0.13 in the quadratic model and from 0.73 to 0.18 in the Box–Cox quadratic model.

[25] The average home sold for 4% less than in the calibrated Wake model and 14% of households were assigned to different homes. Complete results from this analysis, comparable to rows 1–12 of Table 3, are reported in the supplemental appendix. We also tested income share multipliers of 25% and 75% and found very similar results. These results are also reported in the appendix.

[26] We also tested a nonparametric model. It did not perform nearly as well as the Box–Cox quadratic specification, suggesting that with 11 characteristics more than 2000 observations are required to relax the need for parametric structure. In principle, one could use our experimental design to identify the N–K thresholds at which nonparametric models outperform the simpler parametric specifications. One could also evaluate semi-parametric models as a way to investigate how the "optimal" amount of parametric structure varies with sample size. A thorough investigation of semi and nonparametric estimators would be an interesting and potentially important topic for future research.

[27] See [5,13,20] for illustrations of the comparative statics, and [7,10,16] for evidence that the shape of the price function changes over time in response to changes in environmental amenities.

Table 4
Percentage bias in hedonic estimates of the MWTP for reduced commute time[a] (means and standard deviations from 300 Monte Carlo replications).

Row	Estimator	Data	No omitted variables			Omitted variables		
			Linear	Semi-log	Log–log	Linear	Semi-log	Log–log
[1]	OLS	Cross-section [N=2000]	0.25 (0.15)	0.24 (0.15)	0.42 (0.21)	0.57 (0.24)	0.59 (0.24)	0.96 (0.63)
[2]	OLS with fixed effects	Cross-section [N=2000]	0.26 (0.14)	0.18 (0.13)	0.40 (0.23)	0.17 (0.15)	0.23 (0.20)	0.18 (0.13)
[3]	Pooled OLS with fixed effects	Panel [mean (N)=3643]	0.82 (0.09)	0.76 (0.07)	0.87 (0.06)	0.69 (0.07)	0.66 (0.07)	0.81 (0.05)
[4]	Difference in difference	Panel [mean (N)=3643]	0.21 (0.11)	0.16 (0.10)	0.39 (0.16)	0.41 (0.09)	0.33 (0.09)	0.43 (0.07)
[5]	First difference	Panel [mean (N)=1643]	0.33 (0.09)	0.34 (0.10)	0.17 (0.13)	0.56 (0.05)	0.52 (0.05)	0.49 (0.05)

[a] Results are averaged over 100 replications of hedonic equilibria for each of the Cobb–Douglas, Diewert, and Translog utility functions. The omitted variables are (i) median household income, (ii) % under 18, and (iii) distance to the nearest shopping center. (i) and (ii) vary over time.

time, it is not at all obvious that panel data estimates for capitalization rates will provide more accurate estimates for MWTP than cross-section estimates for implicit prices.

We use the simulated changes in commute times in the calibrated Wake County model to evaluate three approaches to panel data estimation. The first approach naively pools the data from the pre-shock and post-shock equilibria. These regressions include spatial fixed effects but ignore the possibility that the shape of the equilibrium price function may change. The second estimator is a generalized version of difference-in-differences (DID). This approach involves OLS estimation of the linear, semi-log, and log–log price functions using spatial fixed effects, a time dummy, and interactions between the time dummy and each covariate. These interactions allow for the possibility that the marginal implicit prices of the covariates may change as we move between different hedonic equilibria. The final panel data strategy involves estimating first differenced (FD) versions of the linear, semi-log, and log–log models, using data on repeated sales of individual homes.

As a baseline for assessing the panel data estimators, the first two rows of Table 4 report the bias in estimates of the MWTP for commute time from the cross-section simulation. The mean bias and its standard deviation are calculated over all 300 Monte Carlo replications (100 for each utility specification). When no variables are omitted, adding spatial fixed effects to the model has very little impact on the bias.[28] This result changes when we omit the three neighborhood amenities (median income, population under 18, and distance to shopping). Without spatial fixed effects, the OLS estimates in row 1 differ from MWTP by 57–96%. Adding the fixed effects in row 2 essentially absorbs the bias from omitted variables.

Rows 3 through 5 provide results from panel data estimation. Notice that the sample size does not double when we move from OLS to pooled OLS. While the change in commute times affects the value of every home, not every home is resold. We only collect data from the ex post equilibria on the homes that changed ownership. This approach mimics the housing market dynamics that generate the data used by empirical hedonic studies. Put differently, while all property values change over time, it is the subset of movers who negotiate the transactions that ultimately enter the databases we use for econometric analysis.

When all variables are observed, the pooled OLS estimators perform poorly. Their absolute bias in measuring MWTP for commute time ranges from 76% to 87%. These figures are 2–3 times larger than the corresponding cross-section results in row 1. The large increase in bias reflects the inability of pooled estimation to account for changes in implicit prices. When commute times decrease, MWTP changes, and so does the curvature of the equilibrium hedonic price function. Failing to allow for this adjustment introduces a temporal variant of functional form misspecification. The resulting bias is more severe than the effect of omitted variables in cross-section data.

By relaxing the "time-constant price function" restriction, the DID estimators substantially reduce the bias on MWTP. The range of values for the bias (16–39%) is slightly lower than the cross-section estimates without omitted variables. The difference may simply reflect the increase in sample size. First-differenced estimates for the capitalization rate also provide a large improvement over pooled OLS. The bias for the three linear models (17–34%) is essentially the same as for DID. Comparing the FD estimates in the left half of row 5 with the cross-section estimates in row 1 provides an example of how panel data estimation can eliminate the bias from time-constant omitted variables. The average bias in the FD regressions is smaller than the bias from cross-section estimation with omitted variables and similar to the bias from cross-section estimation without omitted variables. In the latter case, the size of the difference between capitalization rates and MWTP is similar to the bias from misspecification of the equilibrium price function.

Of the three "spatial" omitted variables, two vary over time: *median income* and *% under 18*. The consequences of omitting these variables can be seen from the last three columns of rows 3–5. Pooled OLS still performs poorly despite the fact that the omitted variable bias slightly offsets the bias from varying implicit prices. DID performs much better, but the

[28] When all variables are observed, adding fixed effects increases the flexibility of the model, which can reduce bias. However, the fixed effects also absorb the variation in commute time between Census tracts. This increases the variance of OLS estimators for the commute time coefficient. As a result, adding fixed effects decreases the accuracy of estimates for MWTP on some Monte Carlo replications. This effect explains the slight (statistically insignificant) increase in mean bias for the linear model.

omitted variables clearly hurt the performance of the linear and semi-log functional forms. The bias in the FD estimates increases substantially, regardless of functional form.

Overall, the generalized DID estimator appears to be the best suited to hedonic estimation in panel data. The interactions between time dummies and housing characteristics control for changes in the shape of the equilibrium price function over time; the spatial fixed effects control for omitted variables in each time period; and no observations are lost. The first-differenced model also performs much better than pooled OLS despite the fact that it drops a large number of observations that were not resold, does not control for time-varying omitted variables, and measures the capitalization rate rather than the implicit price of commute times.

6. A meta-analysis of model performance

While the previous two sections highlighted key results from a subset of our simulations, there are other potentially interesting specifications. For example, we also tested spatial lag models, spatial error models, nonlinear estimators for panel data, and regressions with smaller sample sizes. In order to quickly synthesize our full set of results and allow for broader comparison across estimators, we adapt Banzhaf and Smith's [3] approach to "internal" meta-analysis.

Meta-analysis has traditionally provided a simple and informative way to summarize results across a large number of distinct studies that investigate a common problem. It can also be used in the context of a single study to summarize how subjective modeling decisions influence economic predictions. This internal approach to meta-analysis begins by estimating a large number of candidate specifications for the same underlying model. Each specification is defined by a set of indicator variables describing its econometric features (e.g. functional form, controls for omitted variables) and by a set of variables describing economic outcomes (e.g. MWTP). Regressing economic outcomes on indicators for features of the econometric specification can help to summarize the ways in which the researcher's modeling decisions influence their findings. We adapt this procedure to investigate how the bias in our estimates for MWTP varies with the characteristics of the hedonic specifications we considered.

Our meta-analytic model regresses the value for the normalized welfare bias (β_{kr}) from the calibrated Wake County model on indicator variables for: (a) functional form, (b) the type of spatial control for omitted variables, (c) the type of estimator, (d) the size of the market, (e) selected interactions between the first four categories, (f) the omitted variable scenario, (g) the housing characteristic, (h) the shape of the utility function, and (i) the particular sample of homes.[29] The first five categories describe modeling choices over which the econometrician exerts some control. The last four categories describe features of the simulation; they serve as control variables in the meta-analysis. The estimated coefficient on each indicator variable summarizes how the corresponding feature of the model influences the model's performance. This provides a simple way to detect how specific econometric modeling choices systematically contribute to an empirical hedonic model's ability to identify MWTP.[30]

Table 5 presents the results. The first column uses data from all 11 housing characteristics, the second column uses data on the 5 spatial characteristics (i.e. commute time, nearest park, and census block variables) and the last column uses data on the commute time variable that was the subject of our simulated quasi-experiment. Our interest primarily lies in the second column because 84% of the empirical hedonic studies in our database focus on estimating the MWTP for a single spatially delineated amenity. We include the other two columns because results for the full set of housing characteristics provide the best comparison to CDM and results for commute time are useful in summarizing the performance of alternative estimators when data on the variable of interest are generated by a quasi-experiment. The adjusted R^2 ranges from 0.66 to 0.68 across the three specifications, indicating that hedonic modeling choices explain most of the variation in model performance.[31] These regressions use between 23,537 and 186,946 observations on the bias in measuring MWTP in the presence of omitted variables.

Since all of the regressors are indicator variables, interpretation of results depends on the reference set of modeling choices defined by the excluded indicators. The reference model has a semi-log functional form, is estimated by OLS, uses 200 observations, and has no spatial controls for omitted variables. With this model as the point of reference, the coefficients on the functional form indicators reinforce the main result from CDM. Simpler functional forms perform better in the presence of omitted variables, with the Box–Cox linear and log–log specifications as the top two performers. For example, the Box–Cox linear coefficient of -0.08 in the first column indicates that, all else constant, this model produces estimates for MWTP with 8% less bias than the baseline semi-log model. In comparison, the linear and quadratic models perform the worst.

[29] The metadata do not include results from the tests of alternative equilibria summarized in Section 4. While including these data would decrease standard errors on the coefficients in our meta-regression, we prefer to maintain the realism provided by our calibration of the model to actual transaction prices in Wake County.

[30] The relative magnitudes of the coefficients also provide some guidance on which econometric features are most important to control for in a conventional meta-analysis that seeks to explain why estimates of the MWTP for a particular amenity vary across different hedonic studies. The results in Table 5 suggest it is particularly important to control for functional form, sample size, whether the study uses spatial controls for omitted variables, and whether the study restricts implicit prices to be time-constant.

[31] We dropped 254 observations with biases greater than 500%. These outliers represent approximately one tenth of one percent of our full sample. Dropping these observations leads to a substantial improvement in model fit. An alternative strategy would be to transform the dependent variable to reduce the influence of outliers. This would allow us to avoid dropping observations but it would complicate interpretation of model coefficients.

Table 5
Percentage bias in hedonic estimates for MWTP as a function of modeling choices.[a]

Modeling choice	All characteristics	Spatial characteristics	Commute time
Functional form indicators			
Linear	0.061 (0.004)	0.022 (0.006)	0.018 (0.012)
Log–log	−0.071 (0.004)	−0.133 (0.006)	0.182 (0.012)
Box–Cox linear	−0.083 (0.004)	−0.126 (0.006)	0.253 (0.013)
Quadratic	0.029 (0.005)	−0.017 (0.007)	0.295 (0.014)
Box–Cox-quadratic	−0.050 (0.005)	−0.094 (0.008)	0.310 (0.016)
Market size indicators			
Market size=2000	0.003 (0.002)	−0.045 (0.004)	−0.051 (0.007)
Quadratic, 2000	−0.125 (0.006)	−0.125 (0.009)	−0.170 (0.020)
Box–Cox-quadratic, 2000	−0.120 (0.006)	−0.069 (0.010)	−0.099 (0.021)
Spatial control indicators			
Spatial fixed effects	−0.086 (0.002)	−0.073 (0.004)	−0.148 (0.008)
Spatial error model	0.165 (0.005)	−0.018 (0.008)	0.018 (0.018)
Spatial lag model	0.183 (0.005)	0.030 (0.008)	0.112 (0.018)
Estimation indicators			
Pooled OLS/ML	0.003 (0.003)	0.047 (0.006)	0.112 (0.010)
Difference-in-differences	−0.049 (0.004)	−0.123 (0.008)	−0.133 (0.014)
First difference, FD	0.160 (0.008)	0.100 (0.008)	0.050 (0.015)
Linear, FD	−0.088 (0.012)	−0.054 (0.012)	−0.055 (0.021)
Log–log, FD	−0.180 (0.012)	−0.124 (0.012)	−0.179 (0.021)
Box–Cox linear, FD	0.247 (0.015)	0.304 (0.015)	0.350 (0.027)
N	186,946	68,152	23,537
R^2 adjusted	0.6568	0.6778	0.6708

[a] All regressions include control variables for the particular characteristic being evaluated (e.g. commute time, bedrooms), the functional form for utility (e.g. Diewert, Translog), the omitted variable scenario (e.g. spatial, random), and the particular Monte Carlo replication (e.g. 1, …, 100). Indicators for statistical significance are omitted because almost all variables have p-values below 0.01.

Coefficients on the indicators for market size demonstrate that the accuracy in estimating MWTP for spatial characteristics is improved by increasing the size of the market. This effect comes as no surprise. What is surprising is the magnitude of the improvement for the most flexible functional forms. Increasing the sample size from 200 to 2000 produces a 17% reduction in the bias for the quadratic model (4.5+12.5%) and an 11% reduction in the bias for the Box–Cox quadratic model (4.5+6.9%) in measuring MWTP for spatial characteristics (column 2). Equally important, increasing the sample size enables the use of spatial fixed effects for census tracts without concern for the loss in degrees of freedom.

Spatial fixed effects are clearly the preferable strategy for addressing spatially correlated omitted variables in cross-section data. Adding fixed effects to the model reduces the bias by 7.3%. Interestingly, the spatial error model provides only a marginal improvement over a model with no spatial controls, while the spatial lag model actually performs worse. This relatively poor performance probably reflects the complexity of spatial correlations in our data. In a metropolitan area with a large number of census tracts, spatial fixed effects offer a flexible means of absorbing the irregular concentration of omitted variables. In contrast, the spatial lag and spatial error models impose a rigid structure on the spatial relationships between included and omitted variables. It is natural to expect this rigidity to add to the bias in measuring MWTP. Nevertheless, it is important to point out that our assessment of the two spatial econometric models relies on the "nearest neighbor" specification for the weights matrix. No theory exists for the appropriate construction of this matrix and it is possible that a different specification may lead to materially different results. This would be an interesting topic for future research.

The fourth set of indicator variables describes the relative performance of different cross-section and panel data estimators. Naively pooling panel data increases the average bias in measuring MWTP for spatial characteristics by 4.7% relative to cross-section estimation, despite the increase in sample size. In comparison, the difference-in-differences estimator produces a substantial improvement (12.3%). The stark difference between pooled estimation and DID highlights the importance of controlling for changes in the shape of the equilibrium price function over time. The consequences of failing to do so are the most severe for the variable that changed the most—commute time. In this case, pooled OLS estimates for MWTP are 24% more biased than the DID estimates. The empirical importance of this result is reinforced by recent quasi-experimental studies that demonstrate the public goods we seek to assign implicit prices to often change substantially over time [4,8,20,22,27].

As a final panel data test, we attempted linear Box–Cox estimation of the first-differenced model. This was an utter failure. Our meta-analytic regression predicts a bias on the MWTP for commute time that is 35% higher than the baseline OLS model. It seems likely that differencing the data simply removes the spatial variation needed to pin down the parameters of the nonlinear model. This finding is analogous to the main result from CDM's original study: in the presence

of *time-varying* omitted variables, the more parsimonious *panel data* estimators *with spatial fixed effects* tend to perform the best.

7. Conclusions

Omitted variables have long been viewed as a hidden danger in hedonic modeling and this perception has driven the econometric evolution of the literature. Since Cropper, Deck, and McConnell's innovative study, researchers have been willing to make untested assumptions about the shape of the hedonic price function in order to hedge against the risk of omitted variable bias. While this tradition continues, the literature has also evolved to address the problem directly. It has become routine to use spatial fixed effects to help absorb the confounding influence of omitted variables and recent studies have used quasi-experimental methods to purge time-constant omitted variables from the model. Our simulation results provide the first hard evidence that these new methods work. The evidence also suggests it is time to rethink some of the stylized facts about hedonic estimation.

One stylized fact is that more parsimonious functional forms are less susceptible to omitted variable bias. This is the frequently cited result from Cropper et al. [9]. We find that this result no longer holds when spatial fixed effects are used to control for omitted variables. In this case, the more flexible specifications for the price function, such as the quadratic Box–Cox model, outperform the simpler linear, log–linear, and log–log specifications that have dominated empirical practice for the past two decades. This result is robust to features of our simulation including the specification for utility that underlies hedonic equilibrium and the composition of the set of omitted variables. Thus, it may be time to reconsider the quadratic Box–Cox model and other flexible specifications for empirical hedonic research.

Another stylized fact is that amenities have time-constant implicit prices. About half of the studies we reviewed use more than five years of sales data. They control for changes in amenities over time but, with few exceptions, they restrict the implicit prices of those amenities to be time-constant. This restriction is not suggested by hedonic theory. In general, shocks to preferences, income, information, or amenities can change the shape of the equilibrium price function [13,20,29]. Ignoring this adjustment can bias estimates for MWTP. We find this to be important. Simply pooling data before and after nonmarginal changes in neighborhood amenities conveys a bias that is similar in magnitude to the bias from omitted variables. We find that one can control for time-varying implicit prices by using a generalized difference-in-differences estimator that includes interactions between a time dummy and all of the characteristics of the hedonic price function. Although it is important to acknowledge that our analysis relied on a clean quasi-experiment, these results are encouraging.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jeem.2010.06.001.

References

[1] L. Anselin, Spatial Econometrics: Methods and Models, Kluwer Academic, Dordrecht, 1988.
[2] H.S. Banzhaf, Hedonic pricing in realistic urban structures, or what if tiebout called and nobody sorted? Mimeo, 2003.
[3] S.H. Banzhaf, V.K. Smith, Meta-analysis in model implementation: choice sets and the valuation of air quality improvements, J. Appl. Econometrics 22 (2007) 1013–1031.
[4] H.S. Banzhaf, R.P. Walsh, Do people vote with their feet? An empirical test of Tiebout's mechanism, Amer. Econ. Rev. 98 (2008) 843–863.
[5] T.J. Bartik, Measuring the benefits of amenity improvements in hedonic price models, Land Econ. 64 (1988) 172–183.
[6] P. Bayer, N. Keohane, C. Timmins, Migration and hedonic valuation: the case of air quality, J. Environ. Econ. Manage. 58 (2009) 1–14.
[7] D.S. Brookshire, M.A. Thayer, J. Tschirhart, W.D. Schulze, A test of the expected utility model: evidence from earthquake risks, J. Polit. Economy 93 (1985) 369–389.
[8] K.Y. Chay, M. Greenstone, Does air quality matter? Evidence from the housing market, J. Polit. Economy 113 (2005) 376–424.
[9] M.L. Cropper, L.B. Deck, K.E. McConnell, On the choice of functional form for hedonic price functions, Rev. Econ. Statist. 70 (1988) 668–675.
[10] D.L. Costa, M.E. Kahn, The rising price of nonmarket goods, Amer. Econ. Rev. 93 (2003) 227–232.
[11] L. Davis, The effect of health risk on housing values: evidence from a cancer cluster, Amer. Econ. Rev. 94 (2004) 1693–1704.
[12] I. Ekeland, J.J. Heckman, L. Nesheim, Identification and estimation of hedonic models, J. Polit. Economy 112 (2004) S60–S109.
[13] D. Epple, Hedonic prices and implicit markets: estimating demand and supply functions for differentiated products, J. Polit. Economy 95 (1987) 59–80.
[14] A.M. Freeman, The Measurement of Environmental and Resource Values, Resources for the Future, Washington, DC, 1993.
[15] C.M. Fulcher, Spatial Aggregation and Prediction in the Hedonic Model, Economics, North Carolina State University, Raleigh, 2003.
[16] K. Kiel, K.T. McClain, House prices during siting decision stages: the case of an incinerator from rumor through operation, J. Environ. Econ. Manage. 28 (1995) 241–255.
[17] A.H. Klaiber, V.K. Smith, Evaluating Rubin's causal model for measuring the capitalization of environmental amenities, NBER Working Paper 14957, 2009.
[18] N. Kunzli, R. Kaiser, S. Medina, M. Studnicka, O. Chanel, P. Filliger, M. Herry, F. Horak Jr., V. Puybonnieux-Texier, P. Quenel, J. Schneider, R. Seethaler, J.C. Vergnaud, H. Sommer, Public-health impact of outdoor and traffic-related air pollution: a European assessment, Lancet 356 (2000) 795–801.
[19] N.V. Kuminoff, A.S. Jarrah, A new approach to computing hedonic equilibria and investigating the properties of locational sorting models, J. Urban Econ. 67 (2010) 322–335.
[20] N.V. Kuminoff, J.C. Pope, Hedonic equilibria, land value capitalization, and the willingness to pay for public goods, Mimeo, 2010.
[21] C.G. Leggett, N.E. Bockstael, Evidence of the effects of water quality on residential land prices, J. Environ. Econ. Manage. 39 (2000) 121–144.
[22] L. Linden, J.E. Rockoff, Estimates of the Impact of crime risk on property values from Megan's laws, Amer. Econ. Rev. 98 (2008) 1103–1127.

[23] J.J. McCluskey, G.C. Rausser, Stigmatized asset value: is it temporary or long-term?, Rev. Econ. Statist. 85 (2003) 276–285.
[24] T. Nechyba, J., R.L. Walsh, Urban sprawl, J. Econ. Perspect. 18 (2004) 177–200.
[25] R.B. Palmquist, Property value models, in: K.-G. Mäler, J. Vincent (Eds.), Handbook of Environmental Economics, vol. 2, North-Holland Press, Amsterdam, 2005.
[26] D.J. Phaneuf, V.K. Smith, R.B. Palmquist, J.C. Pope, Integrating property value and local recreation models to value ecosystem services in urban watersheds, Land Econ. 84 (2008) 551–572.
[27] J.C. Pope, Buyer information and the hedonic: the impact of a seller disclosure on the implicit price for airport noise, J. Urban Econ. 63 (2008) 498–516.
[28] S. Rosen, Hedonic prices and implicit markets: product differentiation in pure competition, J. Polit. Economy 82 (1974) 34–55.
[29] S. Rosen, Markets and diversity, Amer. Econ. Rev. 92 (2002) 1–15.
[30] V.K. Smith, J.-C. Huang, Can markets value air quality? A meta-analysis of hedonic property value models, J. Polit. Economy 130 (1995) 209–227.
[31] V.K. Smith, C. Poulos, H. Kim, Treating open space as an urban amenity, Resources Energy Econ. 24 (2002) 107–129.
[32] L.O. Taylor, The hedonic method, in: P.A. Champ, K.J. Boyle, T.C. Brown (Eds.), A Primer on Nonmarket Valuation, Kluwer Academic Publishers, Netherlands, 2003.
[33] J. Tinbergen, On the theory of income distribution, in: L.H. Klaassen, L.M. Koych, H.J. Witteveen (Eds.), Selected Papers of Jan Tinbergen, North-Holland, Amsterdam, 1959.
[34] R.L. Walsh, Endogenous open space amenities in a locational equilibrium, J. Urban Econ. 61 (2007) 319–344.