The Price Elasticity of the Demand for Residential Land: Estimation and Implications of Tax Code-Related Subsidies on Urban Form

Joseph Gyourko and Richard Voith

Lincoln Institute of Land Policy
Working Paper

Lincoln Institute Product Code: WP99JG1

Abstract

The paper presents new estimates of the price and income elasticities for residential land which overcome serious econometric problems associated with earlier analyses. Our estimates are based on a unique data set spanning 28 years of home sales which provides repeated observations on the Montgomery County, PA, housing market. The data allow us to employ the two-stage least squares techniques recommended by Bartik (1987) and Epple (1987) to overcome the central problem associated with estimating demand functions for bundled goods—namely, that consumers simultaneously choose both the price and quantity of the bundled good. Our results suggest that the price elasticity of demand for residential land is around -1 and the income elasticity is near 1.5. Our findings further suggest that ordinary least squares (OLS) estimates are substantially upwardly biased.

The importance of these estimates for understanding the impacts of public policy on the pattern of development is illustrated by examining the potential effects of the U.S. tax treatment of housing on the demand for residential land. Given the price elasticity estimate and tax code-related subsidies to user costs in the 15-percent range, calculations using the assumptions of a traditional monocentric city model suggest that the radius of the occupied area can be as much as seven-percent longer than in the absence of the subsidy.

About the Authors

Joseph Gyourko is Professor of Real Estate & Finance and Director of the Zell/Lurie Real Estate Center at the Wharton School, University of Pennsylvania. His research interests include urban finance and economics, as well as real estate finance. In addition, Professor Gyourko is a Fellow of the Urban Land Institute and a Nonresident Senior Fellow of The Brookings Institution.

Joseph Gyourko
256 S. 37th Street
Zell/Lurie Real Estate Center at Wharton
University of Pennsylvania
Philadelphia, PA 19104-6330
Phone: 215/898-9687
Fax: 215/573-2220
Email: gyourko@wharton.upenn.edu

Richard Voith is Senior Economist at the Federal Reserve Bank of Philadelphia. Dr. Voith is widely published in the areas of urban and transportation economics. He is on the board of the Southeastern Pennsylvania Transportation Authority, the regional transportation authority.

Richard Voith
Federal Reserve Bank of Philadelphia
10 Independence Mall
Philadelphia, PA 19127
Phone: 215/574-3812
Fax: 215/574-4364
Email: richard.voith@phil.frb.org

Authors' Note: This paper was submitted in completion of a project on "The Impact of Housing-Related Tax Expenditures on Residential Land Use Within the Metropolitan Area" (prepared for the Lincoln Institute of Land Policy).

Contents

The Price Elasticity of the Demand for Residential Land: Estimation and Implications of Tax Code-Related Subsidies on Urban Form

I. Introduction

A host of metropolitan issues ranging from central city decline to suburban traffic congestion and the perception of disappearing green space are hotly debated in academia and the press.[1] Each of these issues is intimately related to the society's adoption of less dense land usage patterns. The impact and desirability of any policy addressing these issues, or of an economic change influencing them, is dependent upon the nature of the demand for land. Unfortunately, very little convincing evidence exists regarding the parameters of the demand for land.

Estimating these parameters turns out to be a difficult task for reasons identified by Bartik (1987) and Epple (1987). Those authors showed that the nonlinearity of the underlying hedonic price function relating house value to a trait bundle effectively allows consumers to choose both quantities and marginal prices of all traits—including lot size. Under these circumstances, ordinary least squares (OLS) estimation is likely to result in an upwardly biased price and income elasticities.[2] Identification of the underlying demand function places onerous data requirements that are seldom satisfied. First, repeated observations on the market of interest are needed. Second, the distribution of preferences must not change across the repeated observations of the market. Third, the data must include instruments that shift the household's budget constraint but which are uncorrelated with unobserved tastes that could be influencing the consumed trait set. If these conditions are satisfied, consistent estimates of the parameters of the demand function can be obtained using the two-stage, instrumental variable estimation procedure described by Bartik (1987) and Eppel (1987).

Fortunately, we are able to estimate the price and income elasticity of demand for residential land with a unique data set on house transactions from Montgomery County, PA. This data base includes all single family housing transactions spanning a 20-year period, 1970-1997, in the most populous suburban county in the Philadelphia metropolitan area. Montgomery County, PA, stretches from the Philadelphia city border to the metropolitan fringe. All observations have been geocoded so that street addresses are known in addition to a wealth of structural trait data. We treat each year as a single observation on the market and make the assumption that the distribution of preferences does not differ much across years. These data have been used in other contexts by Voith (1993, 1996), so that their quality has been tested and not found wanting.

The conclusion of our estimations is that the price elasticity of demand for residential land is well above zero, most probably around -1. It is noteworthy that this is 45 percent lower than the estimate resulting from OLS estimation, confirming Bartik's (1987) and Epple's (1987) conclusion that OLS estimates can be biased upward.[3] The income elasticity of demand for land is higher, being around 1.5.

The implications of the price elasticity differential alone are not inconsequential for urban form. Given that the implicit benefits to housing via the tax code are significant, representing 10 to 20 percent of user costs on an annual basis according to most estimates (e.g., see Poterba (1991)), the implied change in the quantity demanded of land is potentially significant for individual owners and for the metropolitan area in general. The extent of increases in land usage depends, of course, on how much capitalization of the implied subsidies occurs. In one extreme where land is supplied perfectly inelastically and the urban boundary is fixed, there can be no land consumption impacts. In the polar opposite case of a monocentric city on featureless plane with land supplied perfectly elastically, after-tax land prices fall by the full amount of the tax code-related subsidy, and, therefore, lead to a very large increase in metropolitan size (and commensurate decrease in residential density). Assuming a price elasticity of residential land of -1.0 and a fall in after-tax land prices of 15 percent, our calculations show that the radius of the metropolitan area increases by about 7 percent. Using a price elasticity of residential land equal to -1.45, the radius of the metropolitan area increases by 10.5 percent.

The remainder of the paper is organized as follows. The next section outlines the econometric issues first raised by Bartik (1987) and Epple (1987) involved in estimating the price elasticity of demand for a single trait such as residential lot size. Section III then describes the Montgomery County, PA, data base in more detail. This is followed with Section IV's presentation of the specifications estimated and a discussion of key results. The implications for urban form are contained in Section V. A brief summary concludes the paper.

II. Econometric Issues

Although using hedonic techniques to estimate market prices of individual traits in bundled goods is standard fare in empirical studies of housing markets, estimates of the underlying demand functions for these traits are rare. Determining the price elasticity of demand for a single trait such as residential lot size is fraught with more than the typical identification problems involved in any situation in which demand (or supply) must be estimated. Bartik (1987) and Epple (1987) pointed out the unique identification problems in their critique of Rosen's (1974) suggested methodology for estimating the supply and demand schedules for bundled traits.

Rosen (1974) analyzed the issue as a standard identification problem and suggested the following two-step procedure for estimating the supply and demand functions for traits of bundled goods such as housing. First, compute individual equilibrium trait prices based on estimates of a hedonic price function such as that for housing shown in equation (1):

$$(1) \qquad V_i = f(Z_{ik}; \beta_k) + \in_i,$$

where: V_i is the observed value of house i;

Z_{ik} is a vector of housing traits;
β_k is vector of parameters;
$\in_i$ is the random error term.

The market price of a trait in the bundle such as residential land, *l*, is given by $p_{il} = \partial V_i/\partial Z_{il}$. Note that if the hedonic price function is nonlinear, the price of residential land will vary across houses.

The second step is to estimate an inverse demand or marginal bid function using the trait prices as the dependent variable:

(2) $p_{il} = \partial V_i/\partial Z_{il} = g(Z_{il}, X_i, E_i, D_i; \gamma_1, \gamma_2, \gamma_3, \gamma_4) + \mu_i,$

where: $Z_{il,}$ is the amount of residential land;
X_i is consumer expenditure on other goods;
E_i is non-housing expenditure;
D_i is a vector demand shifters;
γ_j are coefficient vectors; and
μ_i is an error term.

A companion marginal offer function would be estimated along with (2) and would contain individual supplier traits (S_i). Rosen (1974) suggested that two-stage least squares (2SLS) be employed with the supplier traits being appropriate instruments for the endogenous Z and X variables in the marginal bid function.

Bartik (1987) and Epple (1987) correctly pointed out that the real problem with estimating hedonic demand parameters lies not in traditional supply-demand interaction, as no individual consumer's behavior can affect suppliers because a single consumer cannot influence the hedonic price function itself. Rather, the crux of the problem lies in the nonlinearity of the hedonic price function which implies that consumers simultaneously chose both quantity and marginal price of the housing trait. Epple (1987) illustrated this with a graph similar to that in Figure 1 which has the hedonic price of the trait on the vertical axis and the quantity of the trait on the horizontal axis. Even though the distribution of supply is exogenously given in this example, the nonlinearity of the hedonic price function means that the price changes with any quantity chosen. Hence, a choice of price necessarily implies a choice of quantity (and *vice versa*)—even when only demand parameters need to be estimated.

This leads OLS estimation of (2) to be biased. This can be seen more clearly by following Bartik (1987) and decomposing the error term in our equation (2) into an unobserved tastes component (T_i) and a purely random component η_i as in equation (3),

(3) $p_{il} = \partial V_i/\partial Z_{il} = g(Z_{il}, X_i, E_i, D_i; \gamma_1, \gamma_2, \gamma_3, \gamma_4) + T_i + \eta_i.$

Running OLS on (3) will be biased because the consumer's choice of both the Z's and

X's are correlated with the unobserved tastes component in the residual.[4] Essentially, a household with a strong preference for a given trait in the Z vector will choose more of it. And, as Bartik also noted, individual supplier traits—Rosen's suggested instruments—are also likely to be correlated with unobserved tastes.[5]

The solution to this problem is an instrumental variables estimation, but of a different type than that employed for the standard supply-demand identification problem. Appropriate instruments are those that exogenously shift the household budget constraint yet are uncorrelated with unobserved tastes. The reason is that a shift in the budget constraint shift will be correlated with the observed Z (and X) vectors, yet uncorrelated with tastes. Bartik and Epple recommend two classes of factors that can shift the budget constraint exogenously, and, hence, are candidate instruments. One is income or wealth. The other is the set of variables that shift the hedonic price function, assuming that average tastes do not change across those shifts. This assumption is very important because of what it requires of the data. In particular, it suggests that multiple observations on the market that satisfy two conditions are needed. First, the distribution of tastes must be unchanged across observations. And, second there must be forces that shift individuals' budget constraints across observations.

We are fortunate to have a unique data base to deal with both requirements. Our data are from tax assessment files containing observations on transactions of single family detached homes in Montgomery County, PA, over the period 1970-1996. We treat the sales in each year as an observation on the market for homes, with the maintained hypothesis being that the distribution of preferences does not change over time (or at least does not change in a way that we cannot control for in the estimation).[6] Hedonic regressions of the type illustrated in equation (1) are estimated by year to determine annual market prices of a square foot of lot size (p_l).

Were we not concerned with the price of residential land and residential lot size being simultaneously determined, a simple inverse demand function (or marginal bid) could be estimated via OLS with the price of land, p_{il}, regressed on the quantity of land, Z_{il}, and other appropriate terms such as demand shifters (denoted D) and non-housing expenditures (E). Because of the simultaneity, however, we must estimate the marginal bid function via 2SLS in which Z_{il} and E_i are instrumented for by a set of variables that shift the household's budget constraint without being correlated with tastes.[7] Thus, the specification we estimate via 2SLS is of the form described by equation (4),

(4) $p_{il,t} = h(\hat{Z}_{il,t}, D_{it}, \hat{E}_{it}; \alpha_{1,t}, \alpha_{2,t}, \alpha_{3,t}) + \theta_{iy}$

where: $\hat{Z}_{l,}$ and $\hat{E}_t$ are instruments as described more fully below;

α_j are coefficient vectors; and

θ_{it} is the error term.

Both price and income elasticities can be computed from the α_1 and α_3 coefficients.

Before getting to the specifics of the estimation and the results, the next section more fully describes the data employed in estimating equations (1) and (4).

III. Data

The core data base was created from tax assessment files for Montgomery County, PA, the most populous suburban county adjacent to the city of Philadelphia. The data begin in 1970 and end in 1997. Montgomery County extends from the Philadelphia border to the metropolitan fringe. All observations are geocoded so that precise location within the county is known, allowing matching of observations to census tracts and local jurisdictions. There are 53 such jurisdictions in our sample. The tax assessment data cover all properties in the county and include information on a variety of housing characteristics in addition to sales price. We focus our attention on the over 100,000 observations on sales transactions on single family homes.

Housing and Neighborhood Traits

Table 1 reports summary statistics on structural traits and neighborhood characteristics used in the hedonic equation. Note that these are averages over the full sample across all years. More detailed breakdowns by year are available upon request. The mean age of the homes is nearly 32 years, with a standard deviation of 28 years. The single family homes average 1,973 square feet of living space which is close to the national average for new homes constructed in the United States.[8] Residential lot size averages 19,116 square feet or about 0.44 acres. Other structural traits are categorical in nature and are included at the bottom of the table. These include central air conditioning, which is in just over a third of the single family homes in the sample. Nearly four-fifths have garages, while only one-in-twenty have pools.

In addition to the structure traits, various neighborhood controls are included in some of the hedonic models discussed below. These variables include a couple of density measures—one population based, the other a physical housing measure. Both measures are based on census tract-level calculations. There is an average of 2928 people per square mile in Montgomery County, with a standard deviation nearly as large (row 4, Table 1), with the fraction of a census tract covered by single family housing averaging nearly 16 percent in the sample (row 5, Table 1). Location within the metropolitan area is captured by a variable measuring the highway travel time from the census tract of the house to the central city. This variable is based on a calculation made in 1987 and does not vary over time. Finally, there is a dummy variable for the presence of a train station in the house's census tract. Montgomery County contains the Main Line, which was named for the train line into the city of Philadelphia. The table shows that 42 percent of the homes are in tracts with a train station (bottom row, Table 1).

Table 2 reports house price by year, including the number of observations. This and all other monetary values always are in 1990 dollars. The table depicts the large changes in real prices that have buffeted the Philadelphia market in general and Montgomery County in particular. For example, there was an 80-percent real increase from the 1982 recession to the peak in 1988-89. Since then home prices have trended down over 15 percent.

Supply and Demand Shifters

The data base also contains a wealth of information on supply and demand shifters that are used as instruments in the 2SLS demand estimation. Summary statistics on these variables are reported in Table 3. The supply shifters include a new construction variable, a measure of how extensive new construction is, a measure of census tract size, the amount of vacant land available for residential development, and the total number of sales in the tract. Except for the tract size which is measured only in 1980 in these data, each of these variables varies over time. The average number of new homes per tract per year is about nine, although there is substantial variance in this variable. Table 3 indicates that new homes typically are a small fraction (2.3 percent) of the total number of single family homes in a tract. The typical census tract in Montgomery County is just under three square miles in size. There is a lot of land that is vacant which could be used for residential development, 28.4 percent on average with a very large standard deviation in this variable reflecting the fact that the inner-ring tracts are largely built up and the new areas on the fringe not so. There were approximately 65 home sales per tract per year in the sample. The vector of supply shifters we use also includes two demographic variables (which also could be considered demand shifters)—the fraction of households in the tract that moved between 1975 and 1980 and the percentage of households with heads in the 35-54 age range. These data were computed from the STF3 files of the 1980 census. Mobility is relatively high as indicated by the fact that over 35 percent of households in these tracts had a different residence in 1975. Finally, nearly a quarter of household heads in these tracts are between 35 and 54 years of age.

We use a time series on the thirty-year mortgage rate as both a supply and demand shifter. The average loan rate is 9.5 percent, although this varies widely over time as our series spans the high inflation late 1970s and early 1980s as well as the low inflation late 1990s. Other demand shifters include city and suburban employment growth. City employment growth has averaged a -1 percent per year over the entire sample period, while suburban growth has exceed 2 percent per year.[9] There is substantial variance in these data over time. When these variables are used in the demand estimation, lagged values are employed to deal with endogeneity issues. In addition, they are interacted with municipality dummies.

Non-Housing Expenditures

While our data are very strong in terms of housing, location controls, and potential supply and demand shifters, the Montgomery County tax assessment files do not contain detailed information on household income. The only income data in the files is median income at the tract level for 1980. Consequently, we use data from the *American Housing Surveys (AHS)* in conjunction with this median figure to impute income at the household level. Using the observations on Philadelphia suburbs that can be identified from the *AHS*[10], we begin by defining household income (y) in deviation in mean form as in equation (5),

(5) $y_i = \ln Y_i - \ln Y$,

where y_i is the income for household i and Y is the sample mean across all years. Income in deviation from mean form then is regressed on a set of housing traits common to both the Montgomery County and *AHS* data sets ($x_i = X_i - X$, also in deviation from mean form) and a set of time dummies as shown in equation (6),

(6) $y_i = x_i\delta_1 + \text{TimeDummies}\delta_2 + \lambda_i$,

where λ is an error term. The coefficient vectors δ_1 and δ_2 are then used to impute household incomes in the Montgomery County data base. This is done in a way that incorporates the median tract income information that is available. For the purposes of exposition, denote that tract mean value (which does not vary over time) as Y_c. An increment to that value is imputed, introducing time series and cross section variance from the *AHS*. This increment, denoted $y_{i,m}$, is imputed via the following equation,

(7) $y_{i,m} = x_{i,m}\delta_1 + \text{TimeDummies}\ \delta_2$,

where $x_{i,m}$ represents a housing trait vector in deviation from mean form analogous to that in equation (6), with δ_1 and δ_2 being the coefficient vectors estimated in equation (6). Imputed household income for the Montgomery County observations then is $\hat{Y}_i = Y_c + y_{i,m}$. Finally, non-housing expenditures (denoted E, which is what is required by theory for the demand estimation) is computed using a capitalization rate c to convert house values into service flows as in equation (8),

(8) $E_i = Y_i - cV$,

where V is house value (the sales price in the Montgomery County data) and the cap rate is assumed to be 7 percent.[11] The end result is a mean non-housing expenditure value of E = \$38,885 with a standard deviation about that mean of \$17,321.

IV. Specifications and Results

Hedonic Price Functions

The first task in determining the price elasticity of the demand for residential lot size is to estimate the hedonic price of an added square foot of lot size via a specification as in equation (1). The structural trait variables used include those listed above in Table 1. They are age of the property and its square (AGE, AGE2), a 0-1 dummy for central air conditioning (CENTAIR), a 0-1 dummy for the presence of garage (GARAGE), a 0-1 dummy for the presence of pool (POOL), the square footage of living space and its square (LIVAREA, LIVAREA2), and the square footage of lot (LOTSIZE) in cubic form.[12]

Because we want to estimate the price of a square foot of generic lot, our hedonic specification includes controls for location within the metropolitan area and density controls. The former include controls for the presence of a train station in the same census tract as the house observation (STATION), travel time to the central city and its square (HTIME), and a series of municipality dummy variables. The two density measures are the population density (POPDEN) and land density (LANDDENS) variables discussed in the previous section.[13] Summary statistics on all but the municipalities are reported in Table 1, with the distribution of observations by municipality within the county available upon request.

A semi-log functional form was estimated, as is generally the case with housing data.[14] The model reported on here is described by equation (9). The model was estimated on each annual cross section, with the time subscript being suppressed for convenience sake.

(9) Ln HV_i = f{AGE, AGE2, CENTAIR, GARAGE, POOL, LIVAREA, LIVAREA2, LOTSIZE, LOTSIZE2, LOTSIZE3, STATION, HTIME, HTIME2, POPDENS, LANDDENS, MUNICIPALITY DUMMIES}

The equilibrium price of a square foot of lot can be calculated from (9) as $p_l = \partial HV_i/\partial LOTSIZE$, with the partial evaluated with the coefficients estimated in equation (9) and each observation's HV value.

Table 4 reports regression summary statistics for the estimation of equation (9), along with the mean price of residential lot size (p_l) implied by the hedonic estimation.[15] In general, the annual hedonic regressions do a good job, typically explaining between 60 percent and 70 percent of the variance in house prices. Depending upon the year, the results for p_l suggest that land amounts to between 10 percent and 20 percent of total house value (on average). The mean price of an added square foot of lot size for the sample also exhibits substantial intertemporal variance. For example, Table 2 showed that real house prices rose by nearly 80 percent over the 1982-1989/9 period. Table 4 suggests that the price of residential land doubled over the same period. Also, the land price series takes more discrete jumps. For example, the house price series trends down from 1989, while the price of land falls discretely in a short period of time.[16]

Estimating Demand Via 2SLS

An inverse bid function of the type illustrated above in equation (4) is then estimated using the estimate of p_l from the hedonic as the dependent variable. The results presented below are from a log-log specification (i.e., p_l, Z_l, E are all in log form) that also includes a set of year dummies as controls. Thus, the functional form and specification are as in equation (10).

(10) $lnp_{l,\,i} = \alpha_l \ln \hat{Z}_{l,\,i} + \alpha_{DS} D + \alpha_y \ln \hat{E}_i + \theta_i$,

where all terms are as defined above, with α_l and α_y being used to compute the price and income elasticities of demand respectively.[17]

Before getting to the results, it is worth reviewing the instruments for the Z_l and E terms used in the 2SLS estimation. Recall that appropriate instruments are those that exogenously shift the household's budget constraint, but are uncorrelated with preferences. We use a variety of supply and demand shifters for this purpose, summary statistics for which can be reviewed above in Table 3. The supply shifters include a variety of new construction, tract size, and vacant land variables that capture actual and potential changes in housing activity in individual census tracts (NEWHOME, %NEWHOME, TCTAREA, VACLAND, HOMESALE) as well as two demographic variables reflecting broader mobility (PCTMOVED) and age distribution (PCT35_54). These latter two variables also reasonably can be interpreted as demand shifters. This is also the case with the mortgage interest rate variable (MORTRATE) which can exogenously shift the budget constraint, but could work through both demand and supply.

Variables that are only demand shifters include the first and second lags of the city and suburban employment growth rates (CITEMPGR and SUBEMPGR). Lagged values are used because of potential endogeneity problems arising from the use of contemporaneous employment growth rates. Experimentation showed that including a second lag was useful. In addition, the four employment growth rate variables are interacted with the municipality dummies. The reason interactions with location controls are included is that it is quite possible that demand shifters exogenously impacting the budget constraint will do so differently in different parts of the metropolitan area.[18] For similar reasons, the mortgage interest rate variable is also interacted with the municipality dummies.[19] Finally, a set of year dummies is included as possible demand shifters.

While the interaction terms expand the number of instruments, there is no problem for estimation given our very large sample size. The instrument set explains about 40 percent of the variance in lot size (Z_l in equation (4)) and about 50 percent of the variance in non-housing expenditures (Y in equation (4)). Complete results are available upon request.

The price elasticity resulting from the 2SLS estimation of the inverse bid function for residential lot size is -1.03 with a very small standard error of 0.01.[20] This is well within the range of estimates of the price elasticity of demand for housing services in general.[21] This estimate also is nearly 40 percent lower than the -1.45 figure resulting from a simple OLS estimate of (11), further confirming Bartik (1987) that OLS estimates can be substantially upwardly biased.

That said, there is some reason to worry that our -1.03 figure still is an overestimate of the true elasticity. The reason is that our dependent variable, p_l, is itself estimated with error via the underlying hedonic. As is well known, noise in the dependent variable biases regression coefficients towards zero. That is potentially problematic for us because the estimate of α_l is the inverse of the price elasticity given that we are estimating an inverse bid function. On the other hand, the fact that our sample does not include higher density attached housing could bias the estimate in the other direction. While this restriction provided us a homogenous group of housing observations with which to estimate trait prices, it ignores a potentially important margin of adjustment to a different housing type.

V. Implications for Metropolitan Size and Density and Spatial Sorting by Income

Knowing the price elasticity of residential land allows one to predict how individuals change their land consumption in response to price changes and therefore is crucial to evaluating the impacts of public policies that affect the relative price of residential land. In the United States, the tax treatment of housing is one policy that has large impacts on the relative price of residential land, and therefore land consumption and the pattern of metropolitan development. Poterba (1991) estimates the subsidy to user costs to be between 10 and 20 percent depending upon household characteristics.

Consider first the impacts under the assumptions of the traditional Mills-Muth-Alonso monocentric framework: a flat featureless plane, production in the center, a fixed number of identical workers, and constant commuting costs to the center. Land is supplied perfectly elastically so that after-tax land prices fall by the full amount of any tax code-related subsidies to owner-occupied housing. If we assume a 15-percent subsidy, a price elasticity of residential land of around -1.0 implies that total residential land usage would increase by 15 percent, with residential density being 15 percent lower.[22] The radius of the metropolitan area would increase 7.2 percent.[23]

While this use of our estimate of the price elasticity of residential land is instructive, evaluating the impact of a public policy such as the deductibility of mortgage interest and property taxes on urban form is substantially more complex for at least three reasons. First, the extent to which housing tax policies affect the relative price of housing and therefor cause adjustments along the demand curve differs across individuals. Higher income individuals have higher marginal tax rates and hence the tax code provide significant shifts in the relative price of housing for these individuals while low income individuals are likely to find the standard deduction more attractive than itemization.[24] The second complication is that the impact of a subsidy for owner-occupied housing depends on response of supply: how much of the housing related tax expenditure is capitalized into house prices instead of lowering the after-tax cost of housing? To the extent that the housing subsidies are capitalized into housing prices, there should be no impact on the quantity of residential land purchased.

The final complication derives from the fact that the degree of capitalization of housing subsidies is likely to differ across communities. In suburban communities on the urban fringe, the extent to which housing subsidies are capitalized into land values should be minimal because land is supplied relatively elastically. On the other hand, housing subsidies are likely to be capitalized into property values to a greater extent in fully developed communities, or communities with prevent further development through zoning. Thus the effects of a subsidy that affects the relative price of housing may not have uniform effects across the metropolitan area.[25]

VI. Conclusions

This paper presents new evidence on the price and income elasticity of residential land. A data base spanning 28 years of single family, detached home sales in Montgomery County, PA, is used to provide the needed repeated observations on a single market that Bartik (1987) and Epple (1987) show is required to deal with the special endogeneity problems that arise when consumers effectively choose both the price and quantity of a given trait. Our results show that the price elasticity is in the range of -1.0 and the income elasticity is near 1.5.

Notes

[1] See Mills (1997) and Downs (1992).

[2] See Bartik (1987) for a graphical exposition of the intuition behind the upward bias in the price elasticity.

[3] This finding stands in contrast to that of Cheshire and Sheppard (1998) who found that the econometric problem of joint price and quantity choice is of little practical importance when estimating the demand for residential land. The found that OLS and IV estimates were similar. Their instrument set, however, does not satisfy the requirements of the Bartik-Epple methodology. It is not surprising, therefore that their IV estimates are similar to the OLS estimates.

[4] The problem remains even in the special with perfectly elastic supply of traits.

[5] This can be the case even if consumers do not care about those supplier traits. If different suppliers offer different bundles of traits the problem remains. Bartik's example involved landlords who are carpenters. If units owned by carpenters tend to be better maintained, then consumers with a stronger preference for maintenance will choose carpenter landlords even without knowing anything about the landlords' occupations.

[6] See below for more on this. In any event, we believe the assumption is well founded. Across Montgomery County there are small lots and big lots and preferences certainly differ over this trait. However, we know of no reason why the distribution of preferences would change much, if at all, over time. The assumption of no preference shifts across observations on the market would be much more tenuous if we took individual suburbs as our markets. In that case, hedonic prices would be estimated for each locality, with the assumption being that there is no change in the distribution of preferences across the localities. The distribution of preferences well could be different in a large lot Main Line suburb such as Bryn Mawr versus a small lot, inner-ring suburb such as Darby Borough.

[7] The reason an inverse function must be estimated is because it is not feasible for an instrumented variable to be on the left-hand side.

[8] Data on other traits such as the number of bedrooms and bathrooms also are available. Experimentation with different specifications of the hedonic estimating equation showed a slight preference for the use of square foot of living space in lieu of detailed information on rooms. However, none of the key results reported below is sensitive to including any of these other variables.

[9] The suburban growth rate is based on employment growth in all four Pennsylvania suburban counties of Philadelphia—Bucks, Chester, Delaware, and Montgomery.

[10] We use every available annual survey plus all special metropolitan surveys of the Philadelphia metropolitan area (done approximately every 4-6 years) in this effort.

[11] The 7-percent rate is arbitrary in the sense it is not estimated, but it is within the range reported in the literature. We have experimented with small changes about this number. No result reported below is affected in any meaningful way by this. Future versions of this paper will estimate a cap rate as in Linneman & Voith (1991).

[12] We experimented with a variety of specifications that included purely linear specifications and cubics of different variables. The data do appear to prefer that LOTSIZE be entered in cubic form. The cubed term of this variable helps control for important variance, as in its absence, larger lots of over 1 acre tend to have negative prices per square foot in some years. The price of a square foot of land does tend to be lower for larger lot size homes, but we know of no reason why the price systematically would be negative. In any event, the final results regarding the price elasticity of demand are fairly robust with respect to underlying hedonic specific. The elasticity is somewhat larger if LOTSIZE is only included in quadratic form; the cubic form is associated with the lowest price elasticity. Finally, we also experimented with specifications that included detailed information on the number of rooms in general and the number of bedrooms and bathrooms specifically. Those results are very similar to those for the specification that controls for living area.

[13] The inclusion of the population density control (POPDENS) has the greatest impact on calculated residential land prices. This variable typically has a strong positive coefficient in the hedonic estimation (as predicted by theory). Excluding it leads to higher prices per square foot. This also is associated with slightly higher estimated price elasticities in the subsequent 2SLS estimation.

[14] Box-Cox specification searches generally have shown this functional form to fit housing price data in other research.

[15] The number of observations in some years is slightly lower than that provided in Table 2 for sales prices because of missing data for regressors in the hedonic specification.

[16] It should be noted that we estimated three other hedonic models besides the one reported here and computed the price per square foot of residential land for each. Those models include a simple one with only structural housing traits on the right-hand side of the equation (i.e., no location or density measures; a second one adding select location controls (STATION, HTIME, HTIME2); a third one adding POPDENS and LANDDENS to the second model); and a third one that added only municipality dummies to the first model. The resulting prices per square foot of lot size were highly correlated across the models, with the lowest simple correlation coefficient being 0.91. The findings suggested that the fewer location controls, the higher was the resulting land price and the higher the ultimate price elasticity, although the differences may not

be statistically significant. We believe that the better the intrametropolitan location and amenity controls, the better able we are to estimate the price of a generic piece of land. Hence, we report results from the hedonic model in (9).

[17] Results using other functional forms are discussed below.

[18] Voith (1996) presents evidence that housing markets in Montgomery County communities vary in their response to shifts in city and suburban employment changes.

[19] F-tests showed that each set of interactions contributed in a meaningful way to explaining the variance in both $Z_{l,t}$ and Y_t.

[20] The income elasticity (α_y) is a relatively high 1.5, suggesting that lot size is a luxury good. Full regression results, including those on all demand shifters are available upon request.

[21] This particular result is on the lower end of the range based on a variety of specifications estimated. Specifications with all dependent and independent variables in (11) entered linearly yielded higher elasticities in the 1.3 to 1.5 range.

[22] This example assumes all households are identical. The next paragraph takes up the issue of different households—rich (itemizers) and poor (non-itemizers).

[23] Recall that the radius = (area/pi)**.5. In addition, physical house size would also change, but we do not consider that issue in this paper.

[24] Less than 40 percent of homeowners itemize mortgage and property taxes on their federal income tax returns.

[25] See Voith and Gyourko (1998) for a theoretical discussion of the effects of the tax treatment of housing on the differential spatial and demographic impacts on the patterns of metropolitan development.

References

Bartik, Timothy J. 1987. "The Estimation of Demand Parameters in Hedonic Price Models." *Journal of Political Economy* 95, 1 (February): 81-88.

Downs, Anthony. 1992. *Stuck in Traffic: Coping with Peak-Hour Traffic Congestion.* Washington, D.C.: The Brookings Institution.

Epple, Dennis. 1987. "Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products." *Journal of Political Economy* 95, 1 (February): 59-80.

Gyourko, Joseph, and Richard Voith. 1997. "Housing-Related Tax Deductions and America's Urban Form." Zell/Lurie Real Estate Center at Wharton Working Paper (June).

Linneman, Peter, and Richard Voith. 1991. "Housing Price Functions and Ownership Capitalization Rates." *Journal of Urban Economics* 30, 1 (July): 100-112.

Mills, Edwin S., and Luan Sende Lubuele. 1997. "Inner Cities," *Journal of Economic Perspectives* 35, 2 (June).

Voith, Richard. 1993. "Changing Capitalization of CBD-Oriented Transportation Systems: Evidence from Philadelphia, 1970-1988." *Journal of Urban Economics* 33: 361-376.

———. 1996. "The Suburban Housing Market: Effects of City and Suburban Employment Growth." Federal Reserve Bank of Philadelphia Working Paper (July).

Table 1: Summary Statistics on Select House and Neighborhood Traits

Single Family Home Data Base, Montgomery County, PA, 1970-1997, 102,052 Observations

Variable	Mean	Standard Deviation
AGE (years)	31.7	27.8
LIVAREA (sq. ft. of living area)	1973	799
LOTSIZE (sq. ft.)	19,116	14,922
POPDEN (population per sq. mile)	2928	2377
LANDDENS (% tract covered by single family homes)	15.7%	6.4%
HTIME (Travel Time to Central City by road in minutes)	55.7	16.0

Variable	Frequency
CENTAIR (1= yes if central air)	34.5%
GARAGE (1= yes)	78.6%
POOL (1= yes)	5.7%
STATION Train Station in Census Tract (1= yes)	42.0%

Table 2: Real Home Prices, by Year

Single Family Home Data Base, Montgomery County, PA, 1970-1997, 1990 Dollars

Year	Price	Number of Observations
1970	$109,122	451
1971	$111,636	1486
1972	$117,463	1734
1973	$121,552	1843
1974	$121,984	1760
1975	$119,482	1840
1976	$116,965	2464
1977	$116,412	2963
1978	$116,097	3324
1979	$117,202	3351
1980	$112,052	2524
1981	$110,200	2138
1982	$103,336	1976
1983	$107,022	3070
1984	$110,854	3315
1985	$121,515	3909
1986	$141,738	7211
1987	$163,993	5936
1988	$185,787	5794
1989	$184,314	5259
1990	$181,288	4830
1991	$162,278	4864
1992	$164,783	5461
1993	$161,685	5498
1994	$162,086	5779
1995	$157,622	5265
1996	$153,918	5939

1997	$155,134	2102

Table 3: Summary Statistics on Supply and Demand Shifters

Used As Instruments in the 2SLS Demand Estimation

Montgomery County Data Set, 1970-1997, 102,086 Observations

Variable	Mean	Standard Deviation
NEWHOME (# in tract per year)	9.2	18.6
% NEWHOME (% new homes in tract)	2.3	4.2
TCTAREA (sq. miles)	2.9	3.5
VACLAND (% vacant residential land in tract per year)	28.4	52.9
HOMESALE (# in tract per year)	64.1	37.5
PCTMOVED (% with different residence since 1975)	35.7	10.7
PCT35_54 (% between ages of 35 & 54)	24.3	3.6
MORTRATE (30-year rate, annual)	9.5	1.9
CITEMPGR (% annual employment growth rate for city of Philadelphia)	-1.0	1.7
SUBEMPGR (% annual employment growth rate for suburbs of Philadelphia)	2.2	2.1

Table 4: Hedonic Regression Summary Statistics and the Price of Lot Size

by year

Year	Adjusted-R^2	Nobs	Mean LOTSIZE (sq.ft.)	Mean p_l (price per sq.ft.)
1970	0.66	451	18,676	\$1.07
1971	0.65	1486	19,754	\$1.06
1972	0.61	1733	19,890	\$1.21
1973	0.60	1842	19,199	\$1.12
1974	0.66	1758	18,238	\$1.24
1975	0.64	1840	19,195	\$0.89
1976	0.61	2464	18,264	\$0.73
1977	0.66	2963	18,102	\$0.93
1978	0.56	3324	18,444	\$1.08
1979	0.67	3351	17,886	\$0.72
1980	0.64	2522	18,129	\$0.85
1981	0.65	2137	17,982	\$0.98
1982	0.59	1976	18,401	\$0.81
1983	0.64	3068	18,204	\$0.82
1984	0.65	3314	17,942	\$0.97
1985	0.63	3909	18,411	\$0.93
1986	0.60	7210	19,805	\$1.07
1987	0.67	5933	19,399	\$1.46
1988	0.69	5794	19,532	\$1.67