Housing Prices and Neighborhood Homeownership and Vacancy

Edward Coulson

Herman Li

Department of Economics
Penn State University
University Park PA, 16802

Premliminary Draft; October 1, 2009

Prepared for CURE, the Conference in Urban and Regional Economics, Milan, IT, October, 2009

Abstract: We use panel data, endogenous switching and instrumental variables to assess the externality value of neighborhood homeownership. Using the cluster samples of the 1985, 1989 and 1993 American Housing Survey, we find some variation in the estimate of this value, depending on the assumptions made about the unobservable neighborhood and unit attributes. The most reliable estimates suggest that the externality value of ownership is approximately $1000/year, exceeding the deadweight loss of the mortgage interest tax deduction for all but the highest income households.

Acknowledgments: We thank Sung-Jae Ju, Joris PInkse, Mark Roberts and Brent Ambrose for helpful suggestions, without implicating them in the following.

Introduction

Homeownership is heavily subsidized by the federal government through various tax expenditures and other programs that directly or indirectly provide substantial encouragement for households to become homeowners. Prominent among these is the exemption from tax of the implicit rental income generated by owner-occupation, and the deductibility of mortgage interest payments. Capital gains from the sale of owner-occupied housing are also subject to exclusions and there are other subsidies in the tax code that accrue to owner-occupiers. Using data from 1990, Gyourko and Sinai (2003) estimate that the treatment of implicit rent and mortgage deductibility alone reached almost 200 billion dollars in tax expenditures. The level of tax subsidization has surely reached much higher levels since then. Moreover there are numerous state and local programs designed to foster ownership, particularly in neighborhoods with below average incomes and environments. All recent presidential administrations have been seen as fostering homeownership; examples cited in Gabriel and Rosenthal (2004) bear witness to the political popularity of promoting homeownership

Considerable doubt has been cast on the desirability of these policies, especially in light of the economic downturn of 2007-8. While this paper is not the appropriate venue for a discussion of that downturn, it is appropriate to note that the usual recitation of its chronology lays some blame on households entering ownership without the financial means to do so, and that the public subsidization of homeownership did nothing to discourage such risk-taking. Quite the opposite. However even before recent events focused US attention on the tax treatment of ownership, there were doubts about the interest deduction. A prominent example of this is the Presidents Cmmission Advisory Panel on Tax Reform, which in its final report of November 2005 suggested that it be replaced with a 15% tax credit[1].

The justification for the tax treatment of housing, or any subsidization of ownership should not rest on its status as a merit good-- ! ! ! ! ! ! B !E ! ! ! ! ! ! ! ! -- but with the more compelling justification that ownership creates external benefits; that ownership not only creates private benefits, but benefits for

---

[1] It is worth noting that in light of the exclustion of implicit rent, the mortgage interest deduction merely puts debt and cash purchases of houses on the same footing. Even in the absence of the MID, there would still be subtantial financial advantages to ownership.

the neighborhood and broader community. Over the past two decades or so a research program has grown around the identification of external effects that are created by ownership. Three sets of effects have been so identified and on that account debated in the literature:

a. Maintenance and appearance: Owner-occupiers spend more resources than renters on the upkeep and maintenance of their property. Renters have little incentive to do so directly, since the return on such investment accrues not to them but to the landlord, and the landlord cannot credibly commit to properly compensating tenants for proper maintenance, or what amounts to the same thing, punishing tenants for excessive wear and tear (Henderson and Ioannides (1982)). Landlords do have an incentive to maintain the property, but this comes at a higher cost when the landlord is absentee. Galster (1983) and Harding et al (2000) both come to the conclusion that owner-occupied properties are better maintained than rental properties. DiPasquale and Glaeser (1999) report some similar findings.

b. Family life: Green and White (1997) and Haurin, Parcels and Haurin (2002) both contain evidence to suggest that children growing up in owner-occupied dwellings have higher high school graduation rates and cognitive test scores. Aaronson (1998) notes that this seems to be due to the longer spells that owners have in their place of residence. It should be noted that a recent paper by Barker and Miller (2009) casts doubt on the link between childhood outcomes and ownership, stating that the regressions run by these authors is subject to omitted variable bias, and that car ownership is at least as important as homeownership in this regard.

c. Citizenship: In a widely-cited paper, DiPasquale and Glaeser (1999) provide substantial evidence that homeowners are more involved with local organizations and community, have greater knowledge of their local elected officials, and vote with greater frequency. Not all of this is necessarily productive. Fischel (2001) notes that homeowners will be more active adherents of NIMBYism (i.e. not in my back yard) and that this may devolve to mere rent-seeking. Contrary evidence is obtained by Englehart et al (2009) who use a randomized treatment effect to obtain exogenous shifting of households into ownership. They find little evidence of increased civic or neighborhood involvement by owners.

The lacuna in the above ! ! ! ! ! ! ! ! ! is only indirectly measured; there is little or no sense that the behaviors identified by the above authors is at

all valuable to those that live nearby. Conceptually, it would be a straightforward thing to calculate the externality value of homeownership, even if the behaviors are not directly observable. If

! ! ! ! ! ! ! ! ! ! ! ! ! ! ! !

to live near owner-occupiers. A hedonic regression, one that correlates housing prices (where this can be either the flow rent, or the asset price) to the numerous structural and locational characteristics embodied in them, should include some measure of the ownership propensity in the neighborhood, and that characteristic should, if the aforementioned externalities exist, have a positive coefficient in the regression. Indeed, Nelson (1979), Kohlhase (1991), and likely dozens of other authors have found that the ownership rate within a census tract has a positive influence on housing prices in that tract. However the obvious problem is that there is unobserved heterogeneity across neighborhoods that can cause the correlation to be spurious. Coulson, Hwang and Imai (2002, 2003) tried come to grips with the problem of consistently estimating the hedonic price of neighborhood homeownership in the presence of this heterogeneity. These authors found that even controlling for unobservable person and neighborhood effects (and tenure choice), neighborhood ownership had a positive impact on housing prices.

This paper moves beyond Coulson, Hwang and Imai in a number of dimensions. First, like

D .!I ! !J !)3113 ! ! ! ! ! ! ! !BI !) ! !more

below) to identify neighborhoods, but we also note a prominent number of vacancies in the sample. This has prompted us to address the (increasingly relevant) issue of the hedonic price of neighborhood vacancy. While not as prevalent as rental properties, unoccupied properties are potentially a drag on neighborhood property values (Ioannides and Zabel, 2008). But neighborhood vacancy is also subject to the same endogeneity concerns as neighborhood ownership, so it will be necessary to account for this in our estimation procedures.

Second, we estimate separate hedonic parameter vectors for rental and owner-occupied property. Hedonic modeling nearly always chooses one or the other to comprise the sample (or in the case of Coulson, Hwang and Imai (2002) or Bajari and Kahn (1998) constrains the partial effects to be the same across the two tenure types). In this paper we model rent and value determination as the result of a switching regression process, with endogenous selection into each tenure mode[2]. But in light of the previous paragraph, we also model occupancy on the same principal. Importantly, it

[2] Charlier et al (2001) develop a semiparametric estimator for this model.

is only for occupied properties that we observe a price, thus, we follow Hotchkiss and Pitts (2005) among others[3], in creating a double selection model. First the house is selected into occupancy; conditional on occupancy we observe a price, and that price rent or value-- is the result of endogenous switching into either the rental or ownership market.

Our third innovation is to exploit the panel nature of the neighborhood cluster data. Our first identification assumption is this: that the prices we observe in the data are a function of the current conditions of the house and neighborhood, but the ownership and occupancy status of those are due to conditions that existed prior to the time of the survey. Migration and tenure decisions are very costly and are the result of long and slow adjustment processes. Thus we have natural exclusion restrictions, that the price of a unit at time t does not depend on conditions at t-1; but the ownership and vacancy rate do. This allows for the creation of natural instruments that we use in the estimation of the model.

In the next section, we describe the data set we use in our estimation. We do this first because the nature of the data informs the construction of the empirical model. We then describe that empirical model and our estimation strategy. This is followed by some preliminary results and some conclusions based on these results.

Data

Our data comes from the American Housing Survey (AHS). While the AHS takes a couple of different forms, our data is from the National Sample, which is a biennial survey of over 50,000 housing units from across the USA. It is important to note that this sample repeatedly surveys the same units, so it is a panel of the units, but not necessarily the same occupants. The AHS records data on the price and physical structure of the units (size, assortment of rooms and other characteristics) and the occupants (including income and some limited financial data, as well as numerous demographic characteristics) as well as the quality of the location and housing unit, as evaluated by the occupants.

---

[3] Hotchkiss and Pitts estimate a model of labor force participation and wages. The wage is observed only those who participate, and then there is selection into the full-time or part-time market, whose data generating processes are distinct.

There are both rental and owner-occupied units in the sample. The price given in each record corresponds to the tenure type: rental units report the monthly rent, while the owners provide an estimate of the current market value of the unit. This estimate is of course subject to error (as is perhaps the rent reportage), however Kiel and Zabel (1997) note that while the error has ! ! .! ! ! ! ! ! ! ! ! , this bias is uncorrelated with housing attributes, so that only the intercept term is affected.

Importantly for our purposes, in the 1985, 1989 and 1993 waves of the survey, for a limited number of respondents) ! ! !, the sample also included ighborhood !! ! ! !normally the 10 housing units that are nearest to the respondent in question. These contiguous units are only sampled in the given years and in particular not surveyed in 1987 or 1991 waves. We assemble a panel data set consisting of the kernels and the surrounding cluster for each of the three surveys. The units are classified as being vacant, rented, or owner-occupied.

In the United States, owner-occupied housing is strongly associated with single family structures. The reasons for this are open to debate (Coulson and Fisher (2009)) but Glaeser and Sacerdote (2000) note that the types of social interactions that occur in the former type of housing can be different than those in the latter. For this combination of reasons, we limit our analysis to clusters which are entirely composed of single family households. To include multiple structure types would complicate our analysis immensely, and to use anything other than single family housing as our sample would unduly limit the size of our database.

Table 1 presents the count of clusters. As can be seen, there are as few as six and as many as sixteen houses in a cluster. As hinted above, 11 is the most prominent number, but as noted by Myers (2004) when construction contiguous to the cluster occurs, new units may be added to that cluster, and there were a few instances of this in our sample. Also, attrition may occur, leading to less than eleven units[4].

Given our focus on single family structures, and the aforementioned correlation of structure and tenure types, it is of interest to note that there is in fact substantial variation in both occupancy

---

[4] We noted a few instances of clusters having 2, or similarly small numbers of units. These were eliminated from our sample.

and ownership rates across clusters. Tables 2a and 2b provides evidence on that point. Note that there is a large number of neighborhoods with 100% ownership, and an aggregate ownership rate (.81) that is higher than the overall US ownership rate, but both of these facts befit our single family sample. Nonetheless Table 2 demonstrates a wide variety of ownership rates that appear in our sample.

Model

The preceding considerations suggest an empirical model for housing prices- values (P) and rents (R) based on the following log linear approximation:

$$\ln P_{i\,jt} = X_i\varphi + \bar{Z}_{jt}\gamma + \bar{H}_{jt}\delta + \bar{O}_{jt}\theta + \alpha_{it} + \tau_{jt} + \varepsilon_{ijt} \quad (1)$$

$$\ln R_{i\,jt} = X_i\varphi' + \bar{Z}_{jt}\gamma' + \bar{H}_{jt}\delta' + O\bar{V}_{jt}\theta' + \alpha'_{it} + \tau'_{j} + \varepsilon'_{ijt} \quad (2)$$

The log-linear form is a convenience in that it will allow us to compare the sizes of the coefficients in the rental and owner equations. Note that 2 ! ! ! .! 2 !K ! neighborhood clusters, and t= 1985, 1989, 1993 indexes time periods. Also

P_{ijt}= price of housing unit i in neighborhood j at time period t

R_{ijt}=rent of housing unit i in neighborhood j at time period t

X_i= vector physical characteristics of housing unit i. These are assumed for the time being to be unchanging over time but it is a simple matter to allow alterations over time.

$\bar{Z}_{jt}$=vector of demographic characteristics of neighborhood j. This consists of averages of demographic characteristics of the cluster residents. That is, with n_j as the number of units in cluster j,

$$\bar{Z}_{jt} = \sum_{i=1}^{n_j} Z_{ijt}$$

V_{ijt}=1 if house i in neighborhood j at time t is vacant (and =0 if occupied by an owner or renter).

H_{ijt} =1 if house i in neighborhood j at time t is owner-occupied (and =0if occupied by a renter) and is only observed if the house is occupied.

$\overline{H}_{jt}$= homeownership rate in neighborhood j at time t. This is calculated in the same way as $\bar{Z}_{jt.}$

$\bar{V}_{t}$= vacancy rate in neighborhood j at time t. This is calculated in the same way as $\bar{Z}_{jt.}$

! ! ! ! ! . . .! ! ! in the hedonic equation describing the owner-occupied market. The prime modifiers indicate analogous parameters in the rental market.

$_{it}$=unobserved characteristics of the housing unit that are possibly changing over time.

$_{jt}$= unobserved characteristics of the neighborhood that are changing over time.

$_{ijt}$=random error term

Again, primes indicate parameters from the renter equation. As noted, we observe only the monthly rent for properties that are occupied by renters, ! ! ! ! ! ! ! ! ! units that are owner-occupied, and neither for vacant units.

For reasons noted below, we will limit the estimation to the two latter years in our sample, 1989 and 1993. Jointly estimating these two years in a panel is equivalent, in this linear setting, to estimating the difference regression, so for illustrative purposes consider the difference regression in price

$$\Delta \ln P_{i\,jt} = \Delta X_{it}\varphi + \Delta \bar{Z}_{jt}\gamma + \Delta \overline{H}_{jt}\delta + \Delta \bar{O}_{jt}\theta + \Delta\alpha_{it} + \Delta\tau_{jt} + \Delta\varepsilon_{ijt} \quad (3)$$

Under the usual fixed effects assumption, $\Delta\alpha_{it} + \Delta\tau_{jt} = 0$ and OLS estimation of the above equation provides consistent estimates of the parameters. But now allow for the possibility that the change in these variables is not zero. We make two assumptions to aid in identification:

1. Price or rent is a function of current values of characteristics, however the decisions that led to occupation, tenure, and the demographic composition of residents, took place before the current period. These decisions are, however, based on data available at time t-1, in a role of forming expectations about time t characteristics.

3!! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! !

forecasted based on their current values. Thus $E(\Delta\alpha_{it}|\alpha_{it-1}) \neq 0$! ! !

Now consider the *switching problem.* We note that P and R are only observed when the unit is occupied, and the respective prices are only observed when the corresponding tenure choice is made. Thus both the occupancy and tenure decision for a given property impacts whether a price is observed for that property and nature of price that we do observe. Thus we potentially face a two part selection issue. The first selection is into occupied status, and the second selects into the rental or owner market, with the resulting hedonic equation being relevant.

! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! !

$$O_{ijt}^{*} = F_{ijt}\varphi + w_{ijt} \qquad (4)$$

where F is a vector of variables that affect the time on the market that a house remains vacant. Without belaboring the notation, w_{ijt} contains unobservable elements that are specific to the housing unit and specific to the neighborhood. This data, recall is from an era when the default rate was very low, and housing vacancies were almost never due to foreclosure, but rather due to the simple frictions that arise in a market with significant search costs. As is usual in this kind of context, O* is unobserved but we do observe the indicator function for the unit being occupied (in which case O=1) :

$$O_{ijt} = 1\ if\ O_{ijt}^{*} > 0 => F_{ijt}\varphi > -w_{ijt} \qquad (5)$$

$$= 0\ if\ O_{ijt}^{*} < 0 => F_{ijt}\varphi < -w_{ijt}$$

Similarly, we consider the propensity of an occupied building to be owner-occupied. For any given unit, the process through which it comes to be owner-occupied or rented is a function of the size of

! !) ! ! ! ! ! ! ! !!J ! ! !2: 91 ! ! !2: : 1 .! !

did not practice asset diversification (Caplin et al,) but rather had upwards of 90% of their portfolio tied up in a single asset, their home. Thus owner-occupiers were presumably purchasing less risky

!! ! ! ! ! . as a latent variable H*:

$$H_{ijt}^{*} = K_{ijt}\omega + u_{ijt} \qquad (6)$$

and if H=1 if the unit is owner-occupied:

$$H_{ijt} = 1\ if\ H_{ijt}^{*} > 0 => K_{ijt}\omega > -u_{ijt} \qquad (7)$$

$$= 0\ if\ H_{ijt}^{*} < 0 => K_{ijt}\omega < -u_{ijt}$$

Again, there are unit and neighborhood specific elements to u. Then the mean functions for two prices are:

$$\mathrm{E}(\ln P_{i\,jt} \left| P_{ijt} observed\right) = \mathrm{E}(\ln P_{i\,jt} \left| F_{ijt}\rho < -w_{ijt}\ and\ K_{ijt}\omega > -u_{ijt} \right)$$

$$= X_i\varphi + \bar{Z}_{jt}\gamma + \bar{H}_{jt}\delta + \bar{O}_{jt}\theta + E(\tau_{jt} + \alpha_{it} | F_{ijt}\rho < -w_{ijt}\ and\ K_{ijt}\omega > -u_{ijt}) \quad (8)$$

and for rents:

$$\mathrm{E}(\ln R_{i\,jt} \left| R_{ijt} observed\right) = \mathrm{E}(\ln R_{i\,jt} \left| F_{ijt}\rho < -w_{ijt}\ and\ K_{ijt}\omega < -u_{ijt} \right)$$

$$= X_i\varphi' + \bar{Z}_{jt}\gamma' + \bar{H}_{jt}\delta' + \bar{O}_{jt}\theta' + E(\tau'_{jt} + \alpha'_{it} | F_{ijt}\rho < -w_{ijt}\ and\ K_{ijt}\omega < -u_{ijt}) \quad (9)$$

The two-step estimator would be constructed as follows:

1. A bivariate probit model of O and V is estimated. The log likelihood for this is:

$$\log L = \textstyle\sum_{O=0} \log\left[1 - \Phi(-F'_{ijt}\rho)\right] + \sum_{O=1,H=0} \log\left[\Phi_2(-F'_{ijt}\rho, K'_{ijt}\omega, \sigma_{uw}) + \right.$$
$$\textstyle\sum_{O=1,H=1} \log\left[\Phi_2(-F'_{ijt}\rho, -K'_{ijt}\omega, \sigma_{uw})\right. \qquad (10)$$

where the first term is the contribution of the unoccupied units to the likelihood, the second term is that of renter homes, and the third is the contribution of owner units. We view housing market

participants as being able to forecast the unobserved housing and neighborhood quality using the vectors F and K and make decisions accordingly. Note that under fixed effects, the econometrician can view them too, but we wish to be more flexible than that. Hotchkiss and Pitts (2005; see also Poirier, 1980) note that the hedonic regressions become:

$$\ln P_{i\,jt} = X_i\varphi + \bar{Z}_{jt}\gamma + \bar{H}_{jt}\delta + \bar{O}_{jt}\theta + \sigma_{ug}\lambda_{P1} + \sigma_{wg}\lambda_{P2} \qquad (11)$$

$$\ln R_{i\,jt} = X_i\varphi' + \bar{Z}_{jt}\gamma' + \bar{H}_{jt}\delta' + \bar{O}_{jt}\theta' + \sigma_{ug'}\lambda_{R1} + \sigma_{wg'}\lambda_{R2} \qquad (12)$$

where , !! ! ! ! ! ! ! and are interpretable as the ability of the household to form expectations of the unobservables based on t-1 information, as long as F and K come from that time period, which will be the case. The conditional expectation terms are given as:

$$\lambda_{P1} = \frac{\varphi\left(-F'_{ijt}\rho\right)\left[1-\Phi\left[\frac{\left(K'_{ijt}\omega-\sigma_{uw}F'_{ijt}\rho\right)}{(1-\sigma_{uw}^2)^{\frac{1}{2}}}\right]\right.}{\Phi_2(-F'_{ijt}\rho,-K'_{ijt}\omega,\sigma_{uw})} \qquad (13)$$

$$\lambda_{P2} = \frac{\varphi\left(K'_{ijt}\omega\right)\Phi\left[\frac{\left(F'_{ijt}\rho-\sigma_{uw}K'_{ijt}\omega\right)}{(1-\sigma_{uw}^2)^{\frac{1}{2}}}\right]}{\Phi_2(-F'_{ijt}\rho,K'_{ijt}\omega,-\sigma_{uw})} \qquad (14)$$

$$\lambda_{R1} = \frac{\varphi\left(-F'_{ijt}\rho\right)\Phi\left[\frac{\left(K'_{ijt}\omega-\sigma_{uw}F'_{ijt}\rho\right)}{(1-\sigma_{uw}^2)^{\frac{1}{2}}}\right]}{\Phi_2(-F'_{ijt}\rho,K'_{ijt}\omega,-\sigma_{uw})} \qquad (15)$$

$$\lambda_{R2} = \frac{\varphi\left(-K'_{ijt}\omega\right)\Phi\left[\frac{\left(F'_{ijt}\rho-\sigma_{uw}K'_{ijt}\omega\right)}{(1-\sigma_{uw}^2)^{\frac{1}{2}}}\right]}{\Phi_2(-F'_{ijt}\rho,-K'_{ijt}\omega,\sigma_{uw})} \qquad (16)$$

Now consider *the endogeneity problem.* The analytics are similar but now applied to the other housing units in the neighborhood. The decisions that come together in neighborhood j to create the cluster occupancy and ownership rate depend in part on the neighborhood unobservables. The problem is ameliorated in one respect, that these in the !) ! ! ! !! ! ! ! ! $_i$. However if ! ! ! ! ! ! ! ! .! ! ! ownership and occupancy rates in neighborhood j will be correlated with the error term even under fixed effects.

Our solution to this is to instrument for the ownership and occupancy rates. Natural instruments are available, these being the predicted owner and occupancy rates from the selection equations above.

$$\bar{\bar{V}}_{jt} = \frac{\sum_{i\in j} \Phi(-F'_{ijt}\rho)}{n_j} \qquad (18)$$

$$\bar{\bar{O}}_{jt} = \frac{\sum_{i\in j} \Phi(-K'_{ijt}\omega)}{n_j} \qquad (19)$$

Specification and Estimation

Table 3 presents variables means and standard deviations, stratified by year and tenure status; for convenience we present this for occupied units only. The first pair of variables are rent and value means for the respective tenures. These are logarithms: the corresponding values are around \$90,000 and the rents are approximately \$400/month. The next set of variables are the individual demographic variables which are aggregated to the neighborhood level to create measures of neighborhood attributes. (These are also reported in the table at the individual level.) The household income (ZINC) and number of children (CHILD) variables are self-explanatory. RACE=1 if the head of household is white, and zero otherwise. The measures of school quality (SCH), public transportation (PUBTRANS), and shopping (SHPADEQ) are =1 if the locational ! ! ! ! !!B ! ! .! ng apathetic or ignorant ones are aggregated to zero[5]. By this measure shopping is largely seen as adequate, and public services not so.

[5] E ! ! ! ! !

Structural characteristics include BATHS, the number of bathrooms. In matching observations across time, we inspected the coded number of full baths and the number of half-baths in the unit. Somewhat surprisingly, we found that these two variables were often not the same from observation to observation of the same house, but that the *sum* of the two were (almost always) identical. We conclude that there is some confusion about the distinction between full and half baths in the minds of the survey respondents and we correspondingly just add up the two numbers for every observation and use the total as the measure of BATHS. There is a very slight increase over time in this variable that can be accounted for by renovation. AGEHOUSE is the age of the unit. This the year of the survey minus the year that the house was built. The latter is coded into the AHS as a categorical variable. We take the middle year of the category as our measure of construction year. GARAGE, AIRSYS and the three heating indicators are all =1 if the give attribute exists. Regional and center city dummies are included in our variable list, but are only used in those specifications where fixed effects are not. Finally, two measures of size, LOT and UNITSF, are included in the model. A time dummy is also included in some specifications.

The next series of tables present the results of a series of models. Table 4 merely gives the binary regression with just the ownership rate on the right hand side. While this is an obviously misspecified model, it does give evidence of a surprisingly robust correlation between the ownership rate and housing values over the two years.

The first four columns of table 5 present the full hedonic model, using both years of data and the housing characteristics, without any treatment of unobservables. The first two present the parameters estimates and the associated standard errors for the price of owner-occupied property. The assumption here is that there are no unobservable neighborhood or property characteristics. Note that the signs of the parameter estimates are usually of the expected sign and are significant at the usual levels. Among these are average income in the neighborhood, the measure of shopping adequacy, age, number of bathrooms, lot size, house size, garage spaces, and air conditioning. The sizes of these parameter estimates are not unexpected: an additional garage space adds approximately 9% to the value, air conditioning 11%, and a bathroom 12%, and an additional 1000 square feet about 8%. The time dummy for 1993 is positive, and the age variable is negative (though not statistically significant. Interestingly, additional neighborhood ownership is shown to decrease the average price. Noting that HOMOWN is measured in decimal terms, the parameter indicates that a 10 percentage point rise in neighborhood ownership causes prices to fall by about one-half of one

percent, although a similar increase in the occupancy rate causes prices to rise about 2%. The rental equation (in the next two columns) shows similar findings. Neighborhood ownership is conditionally associated with lower housing prices (and is even more precisely estimated than in the owner equation) and neighborhood occupancy is associated with higher prices.

The next two columns provide estimates for the (individual) fixed effects model. The underlying assumption behind this model is that there are individual and neighborhood fixed effects, but that they are constant over time. There are a couple of things to note from this part of the table applying as well to subsequent analyses. First and most obviously, the unchanging variables (the regional and central city indicators) drop out of the model. Second, conditional on the individual means being concentrated out of ! ! .! ! ! ! ! ! ! ! correlated, and the latter drops out of the model. The age variable therefore enters with a positive coefficient the appreciation rate of housing in this time period is greater than the physical depreciation rate of housing. Third, the impact of the physical attribures of housing become harder to identify, because those characteristics are for the most part unchanging and are on that account folded into the fixed effects. They largely are insignificant in this and subsequent tables. This does not at all imply that the *unobservable* physical attributes are constant over time to the contrary, depreciation and renovation are often unobserved (although there is some information on this in the AHS) and it is precisely these unobserved facets of the physical structure that are most often subject to change.

Observable neighborhood effects can change over time, and the identification of their impact is most revealed when the fixed effects are factored out. Most importantly, controlling for these (fixed) unobservables causes the ownership rate coefficient estimate to switch signs and become positive. Holding neighborhood constant, an increase in the ownership rate by 10% raises house prices by about 5%, and rents by about 3% although the t-statistic on the latter is only about 1.00. The occupancy rate is also positive, though in neither equation is it statistically significant at the usual levels. The fact that the ownership rate is now positive in conjunction with the negative sign from the previous model suggests that ownership in 1989 was most prevalent in neighborhoods with adverse unobservables and the sign of the ownership coefficient estimate had a negative bias. But when those (fixed) unobservables are controlled for, the bias disappears, and neighborhood ownership has a positive impact.

We turn now to the specification of the selection equations, in preparation for estimation of the hedonic equations under the assumption of time-varying unobservables. This involves specifying vectors F and K that determine the occupancy and tenure status of the housing units. Given the panel nature of the data, and the dynamics that are implicitly at work, we use lagged values of the ownership indicator and occupancy indicator. In line with the discussion on the matchability and risk of the housing unit, we also construct the Mahalanobis distance between the housing unit and the average housing unit in the sample. I !)2: 99 ! ! ! ! ! of a housing unit contributed to its time on the market, and by extension, the probability of it being ! ! ! ! ! ! !!I !)2: 99 ! ! ! ! ! ! ! hedonic prices as weights. In this problem, the determination of atypicality must precede the hedonic calculation (and this is congruent with our identification assumption above) so we use the somewhat simpler Mahalanobis distance to measure this. Letting A be the entire vector of variables in the hedonic regression, the Mahalanobis distance is

$$M_i = (A_i - \bar{A})' cov(A)^{-1} (A_i - \bar{A}) \qquad (20)$$

and this variable is included in the F vector that comprises the determinants of vacancy probability. We further speculate that M_i is a determinant of risk, in the sense that a more atypical property also has more variable future price path, and therefore include this measure in K, the vector of variables associated with the probability of ownership. We also include lagged values of the neighborhood occupancy rate in the probability of occupancy equation and the lagged neighborhood ownership rate in the probability of ownership equation. As noted, we estimate these jointly using the likelihood function (.), but stratify the sample by time and estimate the 1989 and 1993 samples separately. (It is the use of lagged values in these equations that constrains us to only use these two waves in the estimation of the hedonic equation.)

Table 6 presents the results of this estimation. As can be seen, the lagged rate (whether occupancy or ownership) plays a significant role in determining the status of units, as is natural. On the other hand, the Mahalanobis distance is quite insignificant in the two occupancy equations; this is not congruent with the results of Haurin (1988), but rather suggests a simpler model of housing units search randomly by prospective buyers. The Mahalanobis measure is positive in both ownership equations, with a relatively precise standard error in 1989 and somewhat less so in 1993. The positive sign is somewhat puzzling, since this is somewhat contradictory to its interpretation as

a measure of riskiness. Two interpretations are possible: one is that homeowners are risk-loving, which, given the concentration of their portfolios is not impossible (Caplin et al, 1997). The second interpretation is that some combinations of attributes are more valuable for the owner-occupier than in other tenure forms, because ownership yields the ability to modify the attribute package to suit the consumption needs of the occupant. It is not inconceivable that atypical packages are more amenable to such renovation.

In any case, the selection terms described above (13) through (16) were inserted into the hedonic as suggested by equations (11) and (12). The first columns of Table 7 present the fixed effects model with the selection terms added. The implicit assumption within this context is that the unobservable house-specific effects are time-varying, but that the unobservable neighborhood effects are constant. As can be seen, the key parameters, the coefficients on ownership rate are basically unchanged. The ownership rate in the owner equation is 0.47, unchanged. It is still statistically significant, with a t-ratio of 3.53. The ownership rate coefficient in the rental equation declines to .23 but still insignificant, with a t-ratio of about 0.76. The occupancy rate is very imprecisely estimated in both equations. Note that both the ownership selection term (labeled lambda in the Table) and the occupancy selection term (labeled lambda2) seem to have much more importance in the rental equation than in the owners equation. For lambda, the coefficient in the owners equation is about half the size of that in the renter equation, and the t-ratio in the former is about 1.2 compared to 3.3. The coefficient of lambda2 is basically zero in the owners equation, while it is negative and significant for renters.

Turning finally to the IV estimates, wherein the implicit assumption is that neighborhood effects change over time, we construct instruments for the ownership rate and occupancy rate as suggested by (11) and (12). The corresponding estimates are presented in the final column. The results suggest a much larger role for the ownership rate, in the sense that its coefficient in the owners equation rises to 1.15, although its standard error rises to about 1.2. So the estimate is rather imprecise. The ownership rate coefficient falls to .13 in the renters equation, and is very imprecisely estimated. The same conclusion applies to the estimate of the coefficient of the occupancy rate in both equations.

So the most robust conclusion is that the ownership rate does have a positive effect on the value of owner-occupied housing. Under the most conservative assumption about the nature of the

unobservables, the (one-sided) prob-value of this inference is only about .16 so there is significant uncertainty about the exact parameter value. Under slightly less conservative assumptions, the estimate is more precise: a ten percent increase in the neighborhood ownership rate raises property values in the neighborhood by about 4.7% (which is itself a more conservative number than that obtained under the most conservative assumptions).

Taking this latter number for purposes of illustration, note that the typical neighborhood in our sample has 11 houses. We ignore the effect of neighborhood ownership on rental units. Assuming nine of these eleven are owner-occupied, the transition of the tenth unit from rental to owner would raise the price of each of the other nine by .47*(.091)= 4.3%. Assuming a typical property value of $90,000, this amounts to about $3850 per housing unit. Thus the externality benefit of ownership in a neighborhood is 9*3849=$34,644. If a 3% annual capitalization rate is applied (assuming an infinite lived asset), this yields an annuity of $1039 per year in externality value.

Poterba (1992) calculates the deadweight loss that accompanies the use of the mortgage interest tax deduction for 1990 taxpayers (i.e. under the 1986 Tax Reform Act). This date is quite congruent with our use of data from 1989 and 1993. This loss varies considerably across income groups and Poterba gives an estimate for those with income of 30, 50 and 250 thousand dollars. The annualized deadweight loss is $53, $326, and $1631 respectively. Thus, the calculations above suggests (subject to considerable variation , obviously) that the benefits of the mortgage interest deduction outweigh the costs, for all but the highest income households.

Conclusions

We have attempted to estimate the externality value of homeownership in the context of the small neighborhood clusters constructed by the American Housing Survey in 1989 and 1993. Our estimation strategy considers a wide variety of assumptions on the nature of the unobservable housing and neighborhood effects. The estimates range rather substantially, but a benchmark model with modest assumptions about those effects suggest that transiting a home from rental to ownership in a typical neighborhood would create about $1000 per year in externality value, suggesting that the mortgage interest deduction may very well pass the cost-benefit test. We observe little or no price effects from neighborhood vacancies in any of our specifications.

References

E !B !)3111 ! B! ! ! !C ! !I !!*Journal of Urban Economics*, 47, 356-369

E !C ! !F ! !)311: ! I ! !D ! !*Real Estate Economics*, 37, 279-303

Andrew Caplin, Charles Freeman, Joseph Tracy, Sewin Chan (1997) *Housing Partnerships : A New Approach to a Market at a Crossroads,* MIT Press

Erwin Charlier, Bertrand Melenberg, and Arthur van Soest, (2001) "An analysis of housing expenditure using semiparametric models and panel data," *Journal of Econometrics*, . 101, 71-107

F !D ! !M !G !)311: ! ! ! !

Edward Coulson, Seok-joon Hwang and Susumu Imai (2002) ! ! ! -Occupation in !*Journal of Housing Research,* 13, 153-174

Edward Coulson, Seok-joon Hwang and Susumu Imai (2003) !C ! ! -Occupation ! !) ! !J ! ! !K !I .!*Journal of Housing Research*, 14, 21-48

E !E ! !F !H !)2: : : ! J ! ! ocial Capital: Are Homeowners Better D !!*Journal of Urban Economics,* 45, 354-384

H ! !F .! !E !F .! !H !H .!H !C ! !)311: ! !B ! ! Social Benefits of Homeownership? Experimental Evidence for Low-Income I !*Journal of Urban Economics,* forthcoming

William Fischel (2001) *The Homevoter Hypothesis: How Home Values Influence Local Government Taxation, School Finance, and Land-Use Policies*. Cambridge: Harvard University Press, 2001

Stuart Gabriel and Stuart Rosenthal (2004) I ! ! !2: 91 ! !2: : 1 ! ! ! ! !*Journal of Urban Economics*, 57, 101-127

G. Galster)2: 94 ! F !F ! !D -Tenure Differences in House Maintenance and D !*Land Economics ,* 59, 107-113

Edward Glaeser and Bruce !)3111 ! ! !D ! !I !*Journal of Housing Economics,* 9, 1-23

!H ! ! ! !)2: : 8 ! ! !C ! !I !F ! !D . ! *Journal of Urban Economics*, 41.

Joseph Gyourko and Todd Sinai (2003) **"**The Spatial Distribution of Housing-Related Ordinary Income Tax Benefits, *Real Estate Economics,* 31, 527-575

K !I !! !K! .! !D!G!)3111 ! E ! ! !C !D ! ! ! I ! ! !! *Real Estate Economics*, 28,: 663 681

Donald Haurin)2: 99 ! !E ! ! ! ! ! !I !*American Real Estate and Urban Economics Association Journal,* 16, 396-410

D. Haurin, T. Parcells, and R. Haurin (2002) "The Impact Of Home Ownership On Child Outcomes" *Real Estate Economics*, 30, 635-666.

K! !I ! ! !! !J !)2: 93 ! B! ! !I ! !D !*American Economic Review ,* 72, 98-113

Katherine A. Kiel and Jeffrey E. !)2: : 8 ! F ! ! ! ! !B !I ! Survey for Creating House P !J !*Journal of Real Estate Finance and Economics* 189-202

Janet L !)2: : 2 ! !J ! ! ! ! ! !I ! !*Journal of Urban Economics*, 30, 1-26

D !L ! !)311 ! E ! ! ! ! ! ! ! ! ! ! !*Journal of Urban Economics,* 56*, 279-302*

Jon !)2: 8: ! B ! .!M ! .! ! ! ! ! !B !*Journal of Environmental Economics and Management,* 6, 320-331

Dale J. Poirier)2: 91 ! ! ! ! ! ! !*Journal of Econometrics*, 12, 209-217

James Poterba)2: : 3 ! ! !I ! ! .! !B !*American Economic Review,* 82, 237-242

Table 2a: Distribution of Occupancy Rates

Occupancy Rate in Cluster	Number of Neighborhoods in 1989	Number of Neighborhoods in 1993
0	1	1
0.09091	0	0
0.18182	2	0
0.36364	0	1
0.45455	1	0
0.54545	2	4
0.6	0	0
0.63636	3	6
0.7	1	1
0.72727	6	14
0.75	2	2
0.76923	0	1
0.77778	3	6
0.8	5	7
0.8125	0	0
0.81818	19	23
0.83333	0	2
0.85714	0	1
0.875	1	1
0.88889	2	3
0.9	11	6
0.90909	84	75
0.91667	5	3
0.92308	0	1
0.9375	1	1
1	219	208

Table 2b: Distribution of Ownership Rates

Ownership rate in cluster	Number of Neighborhoods in 1989	Number of Neighborhoods in 1993	Ownership rate in cluster	Number of Neighborhoods in 1989	Number of Neighborhoods in 1993
0	9	9	0.533333	1	0
0.090909	1	1	0.538462	0	0
0.1	1	2	0.545455	8	8
0.111111	1	0	0.555556	3	2
0.125	1	0	0.571429	2	2
0.142857	0	1	0.583333	0	0
0.153846	1	0	0.6	13	4
0.166667	1	0	0.625	2	2
0.181818	0	1	0.636364	8	7
0.2	2	2	0.666667	6	7
0.222222	0	1	0.7	11	11
0.25	0	1	0.714286	1	2
0.272727	1	2	0.727273	14	14
0.285714	0	1	0.75	3	8
0.3	0	1	0.777778	9	6
0.307692	0	0	0.8	22	16
0.333333	2	2	0.818182	27	23
0.363636	0	2	0.833333	3	4
0.375	1	0	0.857143	3	3
0.4	3	2	0.875	5	5
0.416667	0	1	0.888889	6	9
0.428571	3	2	0.9	22	30
0.444444	2	4	0.909091	48	37
0.454546	4	3	0.916667	0	0
0.466667	0	1	1	109	120
0.5	8	11			

Table 3: Means and Standard Deviations, by year and tenure status (occupied units only)

	rent 1989		rent 1993		owner 1989		owner 1993	
Variable	mean	sd	mean	sd	mean	sd	mean	sd
value= *log of owners estimate of value*	--	--	--	--	11.34	0.80	11.41	0.73
rent= *log monthly rent*	5.95	0.65	6.06	0.64	--	--	--	--
zinc =*household income*	30871.72	22734.27	33985.39	27462.27	43090.19	31954.48	46968.33	34224.04

race= 1 *if white* BT1 0 0 1

heatoil= 1 *if oil heat*	0.07	0.25	0.08	0.26	0.12	0.32	0.11	0.32
heatelec = 1 *if electric heat*	0.19	0.39	0.19	0.40	0.17	0.37	0.18	0.38
Regionne	0.09	0.22	0.08	0.28	0.19	0.40	0.20	0.40
regionmwnc	0.20	0.40	0.19	0.39	0.23	0.42	0.23	0.42
regionwest *=1 for indicated region*	0.32	0.47	0.34	0.47	0.22	0.41	0.22	0.41
centercity= *1 if in center city of msa*	0.41	0.49	0.47	0.50	0.37	0.48	0.37	0.48
lot = *square ft of lot*	14739.74	40330.27	13910.00	30059.29	16823.56	34800.86	18252.40	47824.71
unitsf = *interior square ft.*	1437.56	698.40	1429.27	6.04	1948.54	847.07	1953.84	844.61
mahal = *mahalanobis distance*	4.29	1.26	4.30	1.02	4.19	1.11	4.18	1.21

Table 4

Bivariate Regression Results

	Price, 1989	Rent, 1989	Price, 1993	Rent, 1993
Intercept	10.302 (.08)	5.64 (.09)	10.54 (.07)	5.87 (.12)
Ownership rate	1.200 (.09)	.484 (3.68)	1.01 (.08)	.586 (.08)
R^2	.06	.03	.05	.02
Observations	2817	2864	382	392

Table 5: OLS and Fixed effect regressions

	Panel own		Panel Rent		FE own		FE rent	
Variable	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
homown	-0.05	0.05	-0.17	0.08	0.47	0.13	0.30	0.31
occupied	0.21	0.10	0.17	0.22	0.02	0.15	0.11	0.32
avgincome	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
white	-0.09	0.03	0.12	0.09	0.05	0.12	-0.23	0.34
schadeq	-0.13	0.07	-0.12	0.16	0.09	0.08	-0.13	0.24
shopadeq	0.12	0.05	0.19	0.12	0.12	0.06	0.13	0.20
children	-0.14	0.03	0.04	0.06	-0.06	0.04	0.02	0.09
trans	0.20	0.03	0.09	0.08	-0.09	0.05	-0.15	0.17
agehouse	0.00	0.00	0.00	0.00	0.03	0.01	0.04	0.02
baths	0.12	0.01	0.11	0.04	0.03	0.02	-0.05	0.13
unitsf	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
lot	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
garage	0.09	0.02	0.09	0.04	-0.02	0.04	-0.05	0.11
airsys	0.12	0.02	0.09	0.05	0.03	0.04	-0.13	0.15
heatgas	-0.04	0.04	0.15	0.07	0.10	0.05	0.03	0.11
heatoil	0.21	0.05	0.20	0.10	0.01	0.08	0.32	0.20
heatelec	0.00	0.04	0.23	0.08	0.06	0.06	0.25	0.13
regionne	0.36	0.02	0.40	0.08				
regionmwnc	-0.07	0.02	0.13	0.05				
regionwest	0.54	0.02	0.41	0.05				
centercity	-0.08	0.02	0.04					

Table 6

Selection Regressions

	1989 Occupancy		1989 Ownership		1993 Occupancy		1993 Ownership	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
lagged rate	2.17	0.30	2.22	0.14	1.36	0.32	2.57	0.14
mahalanob	0.03	0.04	0.06	0.03	0.03	0.04	0.04	0.03
Intercept	-0.38	0.36	-0.80	0.18	0.28	0.37	-0.95	0.18

Table 7: Fixed Effects, Selection and Instruments

	Fixed Effects, Selection				Fixed Effects, Selection, and Instruments			
	Owners		Renters		Owners		Renters	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
homown	0.47	0.13	0.23	0.30	1.15	1.20	0.13	14.58
occupied	0.05	0.15	-0.51	0.33	0.42	0.54	-0.69	3.31
avgincome	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
white	0.04	0.12	-0.46	0.33	-0.04	0.30	-0.41	0.88
schadeq	0.09	0.08	-0.03	0.23	0.11	0.09	-0.02	0.36
shopadeq	0.12	0.06	0.10	0.19	0.10	0.10	0.12	0.27
children	-0.06	0.04	0.04	0.08	-0.06	0.05	0.05	0.10
trans	-0.09	0.05	-0.15	0.16	-0.11	0.05	-0.12	1.30
agehouse	0.04	0.01	0.01	0.02	0.04	0.01	0.01	0.06
baths	0.03	0.02	-0.08	0.12	0.03	0.02	-0.08	0.25
unitsf	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.02
lot	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
garage	-0.02	0.04	-0.04	0.11	-0.02	0.04	-0.04	0.16
airsys	0.03	0.04	-0.11	0.14	0.03	0.04	-0.12	0.36
heatgas	0.09	0.05	0.03	0.10	0.09	0.06	0.03	0.41
heatoil	0.00	0.08	0.26	0.19	0.01	0.08	0.26	0.32
heatelec	0.05	0.06	0.28	0.13	0.05	0.06	0.27	0.55
lambdap	0.42	0.34	0.81	0.24	0.47	0.40	0.84	0.78
lambda2p	0.00	0.05	-0.46	0.23	-0.07	0.16	-0.48	1.23
_cons	9.42	0.38	6.21	9.41	8.56	1.19	6.58	36.93