SERC DISCUSSION PAPER 110

Do Homeowners Benefit Urban Neighborhoods? Evidence from Housing Prices

Mika Kortelainen (Government Institute for Economic Research VATT, University of Manchester)
Tuukka Saarimaa (Government Institute for Economic Research VATT)

May 2012

This work is part of the research programme of the independent UK Spatial Economics Research Centre funded by a grant from the Economic and Social Research Council (ESRC), Department for Business, Innovation & Skills (BIS) and the Welsh Assembly Government. The support of the funders is acknowledged. The views expressed are those of the authors and do not represent the views of the funders.

Do Homeowners Benefit Urban Neighborhoods? Evidence from Housing Prices

Mika Kortelainen*
Tuukka Saarimaa**

May 2012

* Government Institute for Economic Research VATT, University of Manchester
** Government Institute for Economic Research VATT

Acknowledgements
We thank James Banks, Jan Brueckner, Essi Eerola, Rachel Griffith, Christian Hilber, Teemu Lyytikäinen, Elias Oikarinen and seminar participants at Helsinki, Jyväskylä, Manchester, Turku and VATT for useful discussions. Oskari Harjunen provided excellent research assistance. Tuukka Saarimaa is grateful to the Academy of Finland and the Finnish Cultural Foundation for financial support and SERC at the London School of Economics and Political Science for hospitality.

Abstract

Homeownership is heavily subsidized in many countries mainly through the tax code. The adverse effects of lenient tax treatment of owner-occupied housing on economic efficiency and growth are large and well documented in the economics literature. The main argument in favor of subsidizing owner-occupied housing is that it creates positive externalities that offset these adverse effects. This paper tests whether homeowners create positive externalities to their immediate neighborhood that capitalize into housing prices in multi-storey buildings. Using semiparametric hedonic regressions with and without instrumental variables we find no evidence of positive externalities from neighborhood homeownership rate. This result is robust to relaxing the identification assumptions of our instrument using a recently developed set identification method. Our results suggest that the adverse efficiency effects of lenient tax treatment of owner-occupied housing are not offset by positive externalities.

Keywords: Homeownership, neighborhood effects, partial linear model, set identification
JEL Classifications: D62, R21, R28

1 Introduction

Homeownership is heavily subsidized in many western countries. In most cases, the subsidy is channeled through the tax code by excluding imputed rental income and capital gains from homeowners' taxable income while allowing them to deduct mortgage interest payments (Hendershott and White, 2000; Englund, 2003). The adverse effects of lenient tax treatment of owner-occupied housing on economic efficiency and growth are well documented in the economics literature.[1] Furthermore, it has been reported that the tax benefits are regressive and benefit mostly middle- and high-income households (Hills, 1991; Poterba, 1992; Poterba and Sinai, 2008; Saarimaa, 2011). Given these facts, the main argument in favor of subsidizing homeownership has to be that it creates positive externalities or social benefits.

The case for positive externalities from homeownership is based on the hypothesis that homeowners put more weight on the condition and amenity levels of their neighborhood than renters. This is because in most cases a house is the single most important asset in a homeowner's wealth portfolio, and thus, a homeowner's wealth level depends on the quality of their immediate neighborhood. This should create incentives for homeowners to engage in activities that improve neighborhood quality. If homeowners' actions improve neighborhood quality, neighborhoods with high homeownership rates are more desirable for prospective buyers and higher neighborhood quality translates into higher housing prices.[2]

---

[1] These include inefficient allocation of the capital stock (Berkovec and Fullerton, 1992; Skinner, 1996; Gervais, 2002), suboptimal household wealth portfolios (Brueckner, 1997; Flavin and Yamashita, 2002; Chetty and Szeidl 2012) and frictions in the labor market (Oswald, 1999; Coulson and Fisher, 2009; Head and Lloyd-Ellis, 2012).

[2] It is somewhat unclear whether all homeowners' actions create external benefits in a broader sense. Homeowners may, for example, oppose the building of social housing in their neighborhoods, and thus, only shift the possible harm to other neighborhoods and households. Nevertheless, these activities should raise house prices in the neighborhood.

DiPasquale and Glaeser (1999) and Hoff and Sen (2005) present formal models where homeowners are able to reap the benefits from their investment in neighborhood amenities (broadly speaking) even if they move away from the neighborhood. This happens because improved neighborhood quality gets capitalized into house values. Renters, on the other hand, are unable to capitalize on their investment to the same extent, because their housing costs increase and landlords capture the increased return on housing capital. Moreover, *ex ante* contracting for these contingencies is very difficult. In fact, if rents rise sufficiently improvements in neighborhood quality may even result in a welfare loss for some renters.

Empirical research on the effects of homeownership is growing, but the results so far are mixed. A typical empirical strategy has been to look at the effects of homeownership on individual outcomes and behavior, such as voting, civic participation or child achievement. The results from this literature remain inconclusive because of major endogeneity issues in estimation. For example, Dipasquale and Glaeser (1999) find that homeowners are more politically involved. Green and White (1997), Boehm and Schlottmann (1999), Aaronson (2000) and Haurin et al. (2002) find that homeownership is associated with improved child outcomes. Hilber (2010) finds evidence that homeowners are more likely to engage in creating and maintaining neighborhood specific social capital especially in areas with inelastic housing supply where social capital is likely to capitalize into housing prices.

However, some recent studies seem to indicate that some if not all of the positive effects of homeownership found in the earlier literature are driven by inadequate control of unobservable factors that are correlated with homeownership. For example, Barker and Miller (2009) argue that the beneficial effects of homeownership on several measures of child welfare are overestimated in earlier literature. Similarly, using an exogenous social experiment as their identifying assumption, Engelhardt et al. (2010) find that homeownership has no effect on political involvement of low-income households.

In this paper, instead of looking at the effects of homeownership on outcomes or behavior of individuals, we simply ask: are houses more valuable in neighborhoods with higher homeownership rates. If homeowners' investments increase neighborhood quality, they should be reflected in housing prices.[3] Our strategy is to estimate a hedonic house price model where neighborhood homeownership rate is included as an explanatory variable. Coulson et al. (2003) and Coulson and Li (2011) use this strategy and find a positive association between neighborhood homeownership rate and prices of single-family homes. Both papers use data from the American Housing Survey. A clear problem in using survey data with a hedonic model is that house prices are not from actual transactions but estimated by the owners in the survey. Banzlaf and Farooque (2012) show that although self-reported house values are correlated with transaction prices, they do not reflect local public goods as well as transaction prices. Furthermore, the analyses in Coulson et al. (2003) and Coulson and Li (2011) are limited to neighborhoods composed of single-family homes. It is not clear whether their results extend to other housing structures or neighborhoods consisting of other than single-family homes.

Our analysis differs from and complements previous studies in a number of ways. First, we use geo-referenced house transaction data which facilitates the use of Geographic Information Systems (GIS) and more precise measurement of neighborhood attributes. Second, precise geo-referencing together with semiparametric econometric techniques allow us to control for unobserved neighborhood attributes that do not vary within a relatively small neighborhood.[4] Third, we concentrate on housing units in multi-storey apartment buildings, which

---

[3] Recently Rossi-Hansberg et al. (2010) and Autor et al. (2011) have found that residential investments have significant price spillover effects on neighboring houses.

[4] For recent examples of these methods in hedonic models, see McMillen (2010) and McMillen and Redfearn (2010).

are the prevailing housing structure in urban areas. The problem in using data from neighborhoods with single-family houses is that, at least in our application of Finnish data, they are almost exclusively owner-occupied. This makes it difficult to separate the neighborhood effects of homeownership from amenities offered by a neighborhood consisting of single-family houses, such as housing structure, open space, gardens and so forth. Concentrating on apartment buildings in a built-up urban area also means that housing supply is relatively inelastic facilitating capitalization, as it is costly to increase density in the study area.[5]

Finally, it is likely that any neighborhood effects of homeownership are tied to the particular housing structure in the neighborhood.[6] For example, homeowners living in single-family houses have a much higher degree of authority on house exterior, yard or garden appearance than their counterparts living in multiunit apartment buildings. Single-family homeowners may also have more connection to neighborhood amenities and the actions of their neighbors so they have more to gain from increased neighborhood quality (Glaeser and Sacerdote, 2000). Furthermore, homeowners' incentives may also be tied to housing structure. As Linneman (1985), Glaeser and Shapiro (2003) and Glaeser (2011) point out, the major maintenance problems in multi-unit buildings are building, not unit specific. This leads to an obvious common pool problem as the owners of individual units in a multi-unit building have an incentive to free-ride on the expense of other owners in the building when it comes to investments into common facilities and building attributes. This may apply to neighborhood amenities as well. It is of obvious interest to see whether homeowners create externalities in an urban environment consisting mostly of multi-storey apartment buildings or whether these externalities are confined to single-

[5] Brasington (2002) and Hilber and Mayer (2008) find that capitalization of local amenities is stronger in areas where housing supply is constrained. Hilber (2010) finds that investment in neighborhood social capital is higher among homeowners in built-up neighborhoods.

[6] In fact, Barker and Miller (2009) find that the effect of homeownership on child outcomes varies depending on whether the family lives in a single-family home or in a multi-unit structure.

family housing neighborhoods. These aspects are of interest because, for instance, in Finland over 40 percent of units in multi-storey apartment buildings are owner-occupied.[7]

A caveat in this type of research is that the neighborhood homeownership rate is likely to be endogenous in a simple regression model either due to omitted variables or simultaneity. We address to this problem by using a semiparametric approach with and without instrumental variables. In fact, concentrating the analysis on multiunit buildings offers a natural choice for an instrument. The arguments by Linneman (1985), Glaeser and Shapiro (2003) and Glaeser (2011) suggest that housing structure, and in particular, the number of housing units in a building should drive the homeownership rate of a neighborhood and should be a valid instrument. We elaborate on this below.

We find no evidence of positive externalities from neighborhood homeownership rate that capitalize into housing prices in different models we estimate. We also test the robustness of these results by relaxing the identification assumptions of our instrument using the method proposed by Nevo and Rosen (2012). In this method, the instrument is allowed to be correlated with the error term of the hedonic regression, while providing a meaningful set identification result. Using this method we cannot reject the null hypothesis that the true effect of neighborhood homeownership on housing prices is zero. Although these more robust results are somewhat imprecise, they are consistent with the fact that the adverse efficiency effects of lenient tax treatment of owner-occupied housing are not offset by positive externalities. This means that the tax favored status that homeownership enjoys in many countries should be questioned.

[7] Homeownership in urban areas and multi-storey buildings is common in other countries as well. For example, according statistics reported in Focus on London (2011), homeownership rate in the City borough is nearly 60 percent and 36 percent in Westminster. According to U.S. Census Bureau, homeownership rate in New York City and Chicago city is 33 and 48, respectively. According to Eurostat's Urban Audit, homeownership rate in Madrid is 78, in Paris 33 and in Rome 64. Unfortunately, we have not found data on homeownership rates according to building type.

The rest of the paper is organized as follows. Section 2 introduces the institutional background. In section 3 the econometric model is presented. Section 4 introduces the data and section 5 the results. Section 6 concludes.

2 Institutional Background

We utilize data from the city of Helsinki, which is the capital of Finland. Helsinki has a population of almost 600,000 and the city can be thought of as one housing and labor market region, although commuting from surrounding cities is easy and frequent. Roughly 48 percent of households living in Helsinki are homeowners. This is a much lower percentage than the national average of 65 percent. The difference is mostly due to a large number of students living in Helsinki and also because single-family units are more often owner-occupied and single-family units are rare in a dense urban area.

In Finland owner-occupied units in multi-unit buildings are part of cooperatives that are incorporated as limited liability companies. This form of ownership is considered as home owning just as much as owning a single-family house and the same tax benefits accrue to both types of homeowners.[8] Membership of a cooperative is obtained by buying the shares on the open market, and the shares can be traded freely. The shares are treated as private property and can be used as collateral on mortgage loans just as single family houses. The company owns all the common facilities and usually the lot as well. The executive board consists of shareholders and the board is responsible for maintenance and investment decisions concerning common facilities.

[8] The transaction tax is lower (1.6 percent) when buying housing company shares compared to buying a property (4 percent).

In some cases, the housing company owns the building but the lot is owned by the city of Helsinki, which leases the lot to the company. This is an important distinction because neighborhood quality capitalizes into land value and land owners benefit from any improvements in neighborhood quality. Thus, if lot rents correctly reflect land value households living on rented lots do not have incentives to make investments into the neighborhood any more than renter households. However, the contracts are long term (up to 60 years), land rents are well below market rents and renewed infrequently. This means that neighborhood investments are not reflected in lot rents, but are capitalized into housing prices.[9] We can control for whether the unit is situated at own lot in our econometric models.

Owning a share does not require the owner to live in the unit in question and the owner can freely rent the unit out. In this case, the household living in the unit is registered as a renter household. In fact, around half of the privately owned rental units in Finland are rented out this way by private individuals. This institutional setting creates within building variation in homeownership rates across the city.

3 Data

We use data from two sources. First, we have access to individual transaction price data provided to us by the Technical Research Centre of Finland (VTT). The data set is based on transactions where major real estate brokerage firms acted as intermediary. We concentrate on transactions made in Helsinki in 2006 and 2007. This choice is made because we have data on neighborhood characteristics only for a single year of 2006. It is plausible that neighborhood characteristics change slowly so using transaction data from two years should not be a problem. After dropping

[9] In most cases, lot rents are tied to the cost of living index and do not follow local house prices.

observations with missing characteristics and insufficient address information for geo-coding purposes, the total sample size is 7,472 units.

The second data source is the Grid Database produced by Statistics Finland, which is used to measure neighborhood characteristics (Statistics Finland, 2010). The grid size is 250 meters x 250 meters and each grid includes, in addition to grid coordinates, information on socio-economic and age structure of the population, building characteristics, employment and service levels. The housing units are geo-coded to the grids.

Descriptive statistics of the data are reported in Table 1. The housing characteristics included in the data are the unit's floor area, age, condition (evaluated for internal purposes by the broker as good, satisfactory or poor), maintenance charge (includes heating, maintenance, property taxes, interest on company debt etc.), indicator that the building is situated on own or rented lot, elevator, floor level and the total number of floors in the building. The data also include the address so we are able to measure exact road distances to the central business district (CBD), to the nearest commuter rail or subway stop and to sea shore using the GIS techniques. Table 1 also includes descriptive statistics of the neighborhood variables used in the econometric analysis.

Table 1. Descriptive statistics.

Variable	Mean	Std. Dev.	Minimum	Maximum
Dwelling characteristics:				
Price (€)	172,585	113,801	60,000	1,500,000
Floor area (m^2)	53.7	25.4	11	362
Age (in years)	54.7	22.5	2	136
Condition (broker estimate):				
Good	0.55	0.50	0	1
Satisfactory	0.39	0.49	0	1
Poor	0.06	0.23	0	1
Situated at own lot	0.76	0.43	0	1
Elevator	0.32	0.47	0	1
Floor level	3.04	1.65	1	9
Total number of floors in the building	4.92	1.69	2	9
Maintenance charge (€/m^2/month)	2.90	0.75	0	8
Road distances (km):				
CBD	5.80	4.49	0.32	19.2
Nearest train or subway stop	1.25	0.80	0.002	5.83
Sea	1.23	1.31	0.01	7.28
Neighborhood characteristics (grid):				
Homeownership rate	0.49	0.14	0.00	0.98
Household median income (€)	31,756	8,847	13,778	144,092
Share of college educated adults	0.28	0.11	0.02	0.72
Unemployment rate	0.09	0.05	0.00	0.42
Share of pension households	0.21	0.09	0.02	0.81
Share of households with children	0.12	0.07	0.01	0.62
Number of service jobs per capita	0.44	1.00	0.00	31.02
Number of buildings	21	12	2	67
Population	834	567	22	2374
Mean floor area of units	55.5	13.3	30.3	185.8
Number of units per building	35.4	18.0	1.1	117.8

Note: The data consist of 7,472 dwelling transactions from Helsinki in 2006 and 2007. All observations are from multi-storey buildings.

4 Econometric Model

4.1 Model Specification

This section discusses our empirical approach. In general, the functional form of the hedonic price function is unknown and nonlinear (see e.g. Ekeland et al., 2004). Several recent papers have allowed nonlinearities by estimating flexible non- or semiparametric hedonic house price models (see e.g. Bajari and Kahn, 2005;

Redfearn, 2009; McMillen, 2010). However, in this paper we have to control for large number of dwelling and neighborhood characteristics (see Table 1), which makes the use of a fully nonparametric model infeasible. Moreover, due to endogeneity issues with our key explanatory variable (homeownership rate), we will use a fairly simple hedonic regression model of the partial linear form. As a starting point, we consider a model

$$(1) \qquad p_{ij} = \mathbf{x}'_{ij}\beta + \delta * amenity_j + f\left(lo_i, la_i\right) + u_{ij},$$

where p_{ij} denotes the log of transaction price of dwelling i in neighborhood (or grid) j, $\mathbf{x}_{ij}$ is a vector of dwelling and neighborhood characteristics, $f(lo_i, la_i)$ is an unknown function of longitude and latitude coordinates of the dwelling (lo_i, la_i) and u_{ij} is the error term. The function $f(.)$ captures unobservable neighborhood characteristics that do not vary within the local estimation window as explained further below.[10]

We assume that homeowners can make investments into neighborhood amenities and social capital, which we denote for simplicity with a scalar *amenity*.[11] More precisely, these are not all neighborhood amenities, but just the ones related to the prevalence of homeowners in the neighborhood. The upshot of this strategy is that we can remain agnostic of what exactly these amenities are. According to hedonic theory, the parameter δ can be interpreted as households' mean marginal willingness to pay (MWTP) for a small change in neighborhood amenities (see

[10] The model is a variant of partial linear models introduced by Robinson (1988). See Rossi-Hansberg et al. (2010) for a recent housing market application of the model.

[11] It is not clear whether neighborhood specific social capital is capitalized into housing prices. DiPasquale and Glaeser (1999) assume that it is not because newcomers do not instantaneously gain social capital meaning that it cannot capitalize, but Hilber (2010) finds empirical evidence of the contrary. Social capital could enhance the ability of neighbors to invest in amenities that require efforts from more than one household. However, it could be that some neighborhood investments of homeowners' do not capitalize into housing prices even when they create externalities.

Rosen, 1974). Naturally, we expect a positive impact for an amenity, i.e. $\delta > 0$. The variables in $\mathbf{x}$ and the coordinates (lo, la) are observed by homebuyers and the econometrician, but *amenity* and u are observed only by the homebuyers, which means that δ is not identified without further assumptions. However, according to theory in DiPasquale and Glaeser (1999), Hoff and Sen (2005) and Hilber (2010) in each neighborhood, the amenity level is a function of the neighborhood's homeownership rate, i.e. $amenity = g(ownrate)$. If homeowners produce positive neighborhood externalities we expect that $\partial amenity / \partial ownrate > 0$.

If we make the simplifying assumption that $amenity = \alpha * ownrate$ and plug this into (1) we get[12]

$$(2) \qquad p_{ij} = \mathbf{x}'_{ij}\beta + \delta\alpha * ownrate_j + f\left(lo_i, la_i\right) + u_{ij}.$$

Estimating equation (2) (and assuming no endogeneity problems) identifies the product $\delta\alpha$, not the mean MWTP for neighborhood amenities. However, $\delta\alpha$ is the relevant parameter for housing and tax policy purposes because it takes into account both the valuation that homebuyers have for neighborhood amenities and, in a sense, how productive homeowners are in "producing" these amenities. A sufficient test for positive externalities is to test whether $\delta\alpha > 0$. It is important to stress that in this setup homebuyers value the neighborhood's amenity level, not homeownership rate *per se*. Moreover, we do not have to assume that homebuyers observe the neighborhood's homeownership rate. They simply need to observe the amenity level.

This argumentation is easily extended to neighborhood disamenities for which the MWTP would be negative ($\delta < 0$). In this case, if homeowners create positive

[12] The parameter α can be interpreted as "the intensity of neighborhood amenity production by homeowners" (see also Kling et al., 2007). This linearity assumption is made for simplicity and the following arguments do not hinge on the assumption.

externalities to neighborhoods (or reduce negative externalities) they should lower the disamenity level ($\alpha < 0$). Again a sufficient test for positive externalities is to test whether $\delta\alpha > 0$. Furthermore, a negative sign for $\delta\alpha$ would always indicate that homeowners inflict negative neighborhood externalities. This could arise if increasing homeownership rate in a neighborhood leads to lower amenity levels ($\delta > 0$ and $\alpha < 0$) or to higher disamenity levels ($\delta < 0$ and $\alpha > 0$). Of course, for a full evaluation of the pro-homeownership tax policies, we should also know how effective current policies are in encouraging homeownership. This is beyond the scope of this paper.

4.2 Instrument Choice

In equation (2) neighborhood homeownership rate is possibly correlated with the error term even with a rich set of controls and the partial linear specification with dwelling coordinates. This could be either due to sorting according to some unobservable neighborhood characteristic that affects prices and attracts homeowners or due to simultaneity of prices and homeownership.[13] Regardless of the reasons for endogeneity an instrumental variable strategy is needed.

Our instrument choice is based on the fact that there are economies of scale in producing housing services that are related to building size in terms of number of units per building (see e.g. Glaeser and Shapiro, 2002; Glaeser, 2011). This may arise through fixed costs in setting up building maintenance and management. It could also be cheaper to arrange the cleaning and maintenance of common facilities or tenant selection and monitoring in one building compared to many buildings

[13] Since we control for a large number of neighborhood attributes and also include nonparametric function of coordinates in our hedonic specification, we expect simultaneity (or reverse causality) to be a more serious problem. We discuss the endogeneity problem and its implications further in Section 5.

dispersed across space. Thus, if there are economies of scale, we should expect institutional landlords to own big buildings rather than many small ones for a given amount of investment in real estate.[14] Furthermore, residents in a multi-unit building face a common pool problem because major maintenance problems are building, not unit specific (see Linneman, 1985; Glaeser and Shapiro, 2003; and Glaeser, 2011). The owners of individual units in a multi-unit building have an incentive to free-ride on the expense of other owners in the building when it comes to investments into common facilities and building maintenance. This may apply to neighborhood amenities as well. Naturally, the common pool problems increase as the number of occupants increases. However, these common pool problems are not present when a single landlord owns the whole building.

These arguments suggest that housing structure and the number of housing units per building in a neighborhood, in particular, should drive the homeownership rate of a neighborhood. More precisely, building size and homeownership rate should be negatively correlated.[15] This means that the number of units per building in a neighborhood is a natural choice for an instrument. The identifying assumption is that the number of housing units per building is not correlated with the error term in the hedonic regression model.

In order to make this assumption plausible, we control directly for a number of household socioeconomic characteristics along with neighborhood population level and total number of buildings. We also control for the floor level of the transacted unit and the total number of floors in the building of the transacted unit. Moreover, we control for the maintenance charge of the unit which should capture the effects that building size may have on maintenance expenses and their effect on transaction

---

[14] A single landlord (a small investor) may also like to own units in the same building due to lower monitoring costs because they know the building, the board that make maintenance decisions etc.

[15] Some recent studies also find empirical evidence on this. For example, Hilber (2011) and Lerbs and Oberst (2011) report that housing structure and especially building size is an important driver of whether a housing unit is owner-occupied.

price. In addition, as a robustness check we use the method by Nevo and Rosen (2012) that relaxes some of the assumptions of our instrumental variable and still produces a meaningful set identification result or bounds for the parameter of interest.

4.3 Estimation

The partial linear model can be estimated as follows (Robinson, 1988).[16] The first step is to use nonparametric regression to regress p and each variable in the parametric part ($\mathbf{x}$) individually on the variables in the nonparametric part, in this case the coordinates (lo, la). For this step we employ local linear regression with nearest-neighbor bandwidths (Cleveland and Devlin, 1988; Li and Racine, 2007). Then form the residuals from each of these regressions, say e_p and e_{x1},..., e_{xk}.[17] The second step is to use OLS to run a regression of e_p on e_{x1},..., e_{xk}. The coefficients for the residuals are consistent estimates of the parameters in the parametric part of the model and the standard errors are valid as well. When instrumental variables (say $\mathbf{z}$) are used, we simply run the nonparametric regression using the instrument as the dependent variable and coordinates on the right hand side and save the residuals (say $e_{\mathbf{z}}$). In the second stage, instead of OLS, we run 2SLS using the residuals ($e_{\mathbf{z}}$) as an instrument for the residual of the endogenous variable (Li and Stengos, 1996; Li and Racine, 2007).

The model produces a location specific intercept (an estimate of the function $f(.)$), which captures unobservable neighborhood quality that stays constant within

[16] See Li and Racine (2007) for a more detailed discussion of partial linear model and its estimation using different estimation techniques.

[17] This can be seen as spatial differencing akin to a spatial fixed effect where you subtract the weighted mean value of a given observation's nearest neighbors from each variable of the observation. Since only coordinates enter the nonparametric part, proximity is simply geographic distance. The mean is basically calculated using inverse distance as weight so that nearer observations get more weight.

the locally fitted regression or local window. This mitigates both endogeneity and spatial autocorrelation problems that may arise due to omitted variables that vary at a broader spatial scale than the local window used in the estimation. Thus, the intercept term in the partial linear model works in a similar way as a spatial fixed effect but with one important difference: in the fixed effects specification it is assumed that there is a discontinuous jump at the fixed effect area borders (say zip codes) whereas the intercept in the partial linear model varies smoothly over space. Another way to illustrate the difference is to consider a housing unit situated just at a zip code border. In this case, the fixed effects estimator uses only observations within the same zip code in calculating the fixed effects transformation. Obviously, using nearby units from both sides of the zip code border can capture unobservable neighborhood quality in some cases in a more plausible way. We report results also from a zip code fixed effects models for comparison.

Using a smooth intercept to control for unobservable neighborhood characteristics may be problematic when some unobservable characteristics, in fact, vary discretely over space. An example would be binding school catchment zone boundaries where school quality might change discretely just at the boundary. We do not expect this to be a major issue in Helsinki for several reasons. First, the school choice system allows the students to attend schools outside of their school catchment area. Second, according to PISA tests variance in quality of Finnish schools is relatively low (or at least much lower than in most other OECD countries). Third, unlike in many countries Finnish pupils are not tested regularly using standardized tests and when these tests are carried out the results are not made public. The reason is that Finnish education policy makers are concerned that publication of test scores and school quality would lead to larger differences among schools and segregation based on school quality.

A key issue in the estimation process is choosing the bandwidth or the number of nearest neighbors used in the nonparametric regressions in the first step. A smaller local window (i.e. the number of neighbors) means that the nonparametric part can capture neighborhood quality and unobserved factors at a more local level. However, in our case the size of the local window or the number of nearest neighbors is constrained from below because the neighborhood variables are measured based on fixed sized grids. If the local window is too small, in some cases all the observations within a local window will be in the same grid and there would be no variation in the neighborhood variables making identification impossible. The average number of transactions in a grid in our data is about 9. However, the maximum number is 143, which is the lower limit for the size of our local window. The average number of observations in a zip code is about 106 whereas the maximum is 446. We select the number of nearest neighbors so that there will be enough independent variation for each neighborhood characteristic. We also experiment with bandwidths of different size using 150, 200 and 300 nearest neighbors or transactions. In all specifications, we use the second order Gaussian Kernel in the nonparametric step.

5 Results

5.1 Main Results

Results from our hedonic models are presented in Table 2. Note that we report only the results concerning neighborhood homeownership rate in Table 2, while the full results from the hedonic models are reported in the Appendix. Homeownership rate and other neighborhood characteristics are standardized so that the coefficient measures the percentage change in prices as a neighborhood's homeownership rate increases by one standard deviation. As a benchmark we estimated simple OLS models with and without zip code fixed effects. The OLS model without zip code fixed effects and controlling for other neighborhood characteristics produces a negative and statistically significant coefficient for neighborhood homeownership rate. At face value, this rather surprising result means that increasing neighborhood homeownership rate by one standard deviation (roughly 14 percentage points) decreases house values by approximately 6 percent. Including zip code fixed effects reduces this effect substantially to 2 percent.

Table 2. Point estimates from hedonic models.

	Panel A: Linear model		
	OLS	**OLS**	**2SLS**
Homeownership rate	-0.057**	-0.024**	-0.071
	(0.008)	(0.003)	(0.048)
Zip code fixed effects	no	yes	yes
F-test for instrument	..	..	19.80
R^2	0.91	0.94	0.94
Observations	7,472	7,472	7,472
	Panel B: Partial linear model		
	nn = 150	**nn = 200**	**nn = 300**
Homeownership rate,	-0.015**	-0.019**	-0.022**
not instrumented	(0.003)	(0.003)	(0.003)
Homeownership rate,	-0.024	0.004	0.031
instrumented	(0.038)	(0.038)	(0.041)
F-test for instrument	26.73	25.20	20.25
Observations	7,472	7,472	7,472

Notes: The dependent variable is the natural log of transaction price. All models include the control variables reported in Table 1 and quarter of sale dummies. Heteroscedasticity robust standard errors are reported in the parentheses. The marking nn refers to the number of nearest neighbors or bandwidth size used in the local regressions in the partial linear model. ** and * indicate 1 and 5 percent significance levels, respectively.

The partial linear models with small bandwidths (150, 200 and 300 nearest neighbors) produce similar results. In absolute terms, the coefficient is slightly smaller or of similar size as in the OLS zip code fixed effects model. As the bandwidth size is increased the coefficient on homeownership rate increases in absolute terms, although not very much. This might indicate that smaller bandwidths work better in mitigating endogeneity problems or control better for unobserved neighborhood variables.

There are two possible explanations for the negative coefficient. First, the result could be the true causal effect so that in of multi-storey buildings homeowners' net effect on neighborhood amenity level and house prices is indeed negative (e.g. $\delta > 0$ and $\alpha < 0$). One possible mechanism is outlined in Linneman (1986), Glaeser and Shapiro (2003) and Glaeser (2011). Multiple owners in a building create a common pool problem in investment decisions concerning building specific projects and in neighborhoods with high homeownership rates buildings could be in worse shape. Second, the result could be explained by reverse causality. Households may be reluctant to own expensive homes because it distorts the allocation of their wealth portfolio as explained in Henderson and Ioannides (1983), Brueckner (1997) and Flavin and Yamashita (2002). For this reason, households with a clear preference for homeownership may sort into lower price neighborhoods, whereas households with a clear preference for an expensive neighborhood (sea shore etc.) may choose to rent and not distort their wealth portfolio too much towards housing capital.[18] If this is the case the result is due to endogeneity and the effect is not the true causal effect of neighborhood homeownership rate on house prices.

We find the latter explanation more plausible and therefore an IV strategy is needed. Using 2SLS with zip code fixed effects again produces a negative effect, but the effect is not statistically significant. The first stage F-test statistic is 24.6 indicating that the instrument has good explanatory power. The partial linear model with the smallest bandwidth (150) also produces a negative, but statistically insignificant effect. With larger bandwidth (200 and 300) the coefficient is positive, but again the coefficient is not statistically significant in either model. The

[18] Another potential and related reason for sorting is that user-cost-to-rent ratios can vary across neighborhoods and homeownership might be lower in neighborhoods with higher user-cost-to-rent ratios due to differences in expected capital gains (see e.g. DiPasquale and Glaeser, 1999; Amior and Halket, 2011).

instrument works well also with the partial linear model as can be seen from the high *F*-test values. Furthermore, in all cases, the partial correlation of the instrument and homeownership rate in the first stage is negative as expected.

Our preferred results from the partial linear models with instrumental variables indicate that neighborhood homeownership rate has no effect on housing prices, and thus, there is no evidence of positive externalities.

This result is contrary to the theoretical predictions by DiPasquale and Glaeser (1999) and Hoff and Sen (2005). However, the results could be explained by at least two things. First, homeowners' incentives to invest in neighborhood amenities could depend on building type. Glaeser and Sacerdote (2000) argue that single-family homeowners are physically more connected to their immediate neighborhood compared to homeowners living in multi-storey buildings. This means that the latter group benefits less from neighborhood quality, and thus, has weaker incentives to invest in their neighborhood. Common pool problems may weaken the incentives further.

Second, households' willingness to pay for neighborhood amenities may also depend on building type. If this is the case, our results are in line with Autor et al. (2011) who studied housing market externalities in the context of rent control abolishment in Cambridge Massachusetts. As rent control ended, landlords in Massachusetts had more incentives to invest in the quality of their dwellings and this could have a spillover price effect on surrounding dwellings. They find strong spillover effects, but the effects were much stronger on individual houses compared to condominiums. This could be explained by the fact that households living in condominiums are less connected to their neighborhood than households living in single-family houses. This means that neighborhood quality is less important to condominium households and they are not willing to pay as high a price premium on neighborhood quality as households living in single-family houses.

The theoretical models by DiPasquale and Glaeser (1999) and Hoff and Sen (2005) do not differentiate homeowners' incentives according to building type. Therefore we cannot claim that our results refute these models. Our results do suggest, however, that homeowners' investment incentives are not strong in multi-unit buildings and the predictions from DiPasquale and Glaeser (1999) and Hoff and Sen (2005) may be confined to homeowners living single-family houses.

5.2 Robustness Check Using Set Identification

The assumption that our instrument is exogenous may not be true. For example, the prevalence of large buildings may be associated with unpleasant neighborhoods because they block sunlight and view. This could induce correlation between the instrument and the error term. In this section, we present results using a set identification method developed by Nevo and Rosen (2012) that relaxes this assumption and allows for an imperfect instrument.[19] To illustrate the method consider the following model specification

$$\tilde{p} = \gamma * \widetilde{ownrate} + u, \tag{3}$$

where $\tilde{p}$ and $\widetilde{ownrate}$ denote the log of transaction price and homeownership rate from which we have netted out the effects of the unit's coordinates (lo, la) and control variables ($\mathbf{x}$). Let z denote our instrument. Instead of assuming that $corr(z,u) = \rho_{z,u} = 0,$ we need to make assumptions about the signs of correlations between the error term, the endogenous regressor and our instrument. More specifically, we need to assume that the correlations between the endogenous

[19] See Nevo and Rosen (2012) for the formal presentation of linear model and Reinhold and Woutersen (2011) for the extension to the partial linear model.

regressor and the error term and between the instrument and the error term have the same sign implying that $\rho_{z,u} \cdot \rho_{\widetilde{ownrate},u} \geq 0$.

We believe that the sign assumptions for the correlations are plausible because of the following arguments. First, we assume that the correlation between the endogenous regressor and the error term is negative ($\rho_{\widetilde{ownrate},u} < 0$) due to simultaneity. This is a plausible assumption in our case as explained above. Households may be reluctant to own units in expensive neighborhoods because of portfolio diversification reasons, and thus, homeowners may sort into neighborhoods with lower house prices. Since we cannot necessarily control for this aspect well enough in our regressions, we expect the non-zero correlation between the error term and the regressor to be negative rather than positive.

Second, we argue that the correlation between the instrument and the error term is also negative ($\rho_{z,u} < 0$). For example, if sunlight or prevalence of green areas affect prices (i.e. are in the error term of our hedonic regression model) a negative correlation arises because neighborhoods with big buildings will have lower levels of these amenities. Street crime incidence may also be higher in neighborhoods with tall buildings as reported by Glaeser and Sacerdote (2000). This would also induce a negative correlation between the instrument and the error term. In addition, we know from the first stage of our instrumental variable models presented in Table 2 that the partial correlation between our instrument and homeownership rate is negative. These three assumptions together provide us with both a lower and an upper bound for the true parameter value, even when the instrument is not exogenous.

The upper bound can be obtained using the standard IV regression or IV with the partial linear model as above. The intuition can be easily seen from the probability limit of the standard IV estimator:

$$\text{(4)} \qquad \operatorname{plim}\left(\gamma_{IV}\right) = \gamma + \frac{\rho_{z,u}}{\rho_{z,\widetilde{ownrate}}} \cdot \frac{\sigma_u}{\sigma_{\widetilde{ownrate}}} \geq \gamma,$$

where ρ again denotes the correlation and σ denotes the standard deviation. With our assumptions, the second term of the probability limit is positive (since $\rho_{z,u}, \rho_{z,\widetilde{ownrate}} < 0$) and the IV estimator gives the upper bound for the true parameter γ. Similarly, the lower bound can be obtained using OLS because

$$\text{(5)} \qquad \operatorname{plim}\left(\gamma_{OLS}\right) = \gamma + \rho_{\widetilde{ownrate},u} \cdot \frac{\sigma_u}{\sigma_{\widetilde{ownrate}}} \leq \gamma,$$

where σ again denotes the standard deviation. Thus, in our case, the above assumptions about the signs of the correlations are enough to identify bounds for the true parameter value.

However, Nevo and Rosen (2012) show that tighter bounds can be obtained if we are willing to make a further assumption that the correlation between the endogenous regressor and the error term is larger than the correlation between the instrument and the error term, i.e. $\left|\rho_{\widetilde{ownrate},u}\right| \geq \left|\rho_{z,u}\right|$. This condition implies that while we allow $\rho_{z,u} \neq 0$, the instrument needs to be “less endogenous” than the endogenous regressor. We think that this is a reasonable assumption in our application. Since we are controlling for a large number of factors affecting house prices, we expect the correlation between the error and the instrument to be quite negligible and at least smaller than the correlation between the endogenous regressor and the error term. When this condition is satisfied, one can estimate the lower bound using a generated instrumental variable suggested in Nevo and Rosen (2012). In our case, the generated instrumental variable is defined as $z_{Nevo-Rosen} = \sigma_{ownrate} \cdot z - \sigma_z \cdot ownrate$.[20]

Naturally, these correlation assumptions cannot be directly tested because one never observes the error term. However, the assumptions used here are less strict than the usual assumptions needed for a valid instrument, which assume a zero correlation between the instrument and the error term.

We also need to estimate the confidence interval for the identified set. Following Reinhold and Woutersen (2011), we use the regular bootstrap to estimate confidence intervals for the true parameter value. We sample units with replacement and get an estimate for the lower and upper bound for each subsample. We generate 400 subsamples for each bandwidth size. We then stack all the estimates of the upper and lower bound in an ordered vector (with length 800) and calculate the 2.5% and 97.5% percentiles, which give us the 95% confidence interval.

The results using the Nevo and Rosen (2012) set identification strategy are presented in Table 3. We estimate both a linear model with zip code fixed effects and partial linear models with different number of nearest neighbors using the same model specifications as in Table 2. The figures in Table 3 correspond to the 2.5% and 97.5% percentiles of the bootstrapped parameter estimates. The bounds are rather large in all specifications and grow significantly in the partial linear model as the bandwidth is increased. However, in each case, zero is in the identified set of the true parameter value, and thus, we cannot reject the null hypothesis of a zero effect. Interestingly, all the estimates from the partial linear models reported in Table 2 are within the confidence interval. However, the estimates from the partial linear models without instrumenting are quite close to the lower boundary of the confidence interval.

[20] Again see Nevo and Rosen (2012) for details.

Table 3. Bootstrapped confidence intervals for Nevo-Rosen set estimates.

	Linear model	**nn = 150**	**nn = 200**	**nn = 300**
Upper bound	0.003	0.048	0.085	0.125
Lower bound	-0.168	-0.072	-0.046	-0.031
Observations	7,472	7,472	7,472	7,472

Notes: The table reports 95 percent confidence intervals for the true parameter value. The confidence intervals are based on 400 bootstrap repetitions per model.

6 Conclusions

Homeownership is heavily subsidized in many western countries. The adverse effects of lenient tax treatment of owner-occupied housing on economic efficiency and income distribution are well documented in the economics literature. Given these facts, the main argument in favor of encouraging homeownership has to be that it creates positive externalities. If homeowners create positive externalities to their neighbors by improving neighborhood quality, neighborhoods with high homeownership rates are more desirable for prospective buyers, which should translate into higher housing prices.

In this paper, we tested this hypothesis using semiparametric hedonic regression models with and without instrumental variables. Our strategy was to estimate a hedonic house price model using data from multi-storey buildings where neighborhood homeownership rate is included as an explanatory variable. We found no evidence of positive externalities from neighborhood homeownership rate that capitalize into housing prices. We also tested the robustness of these results by relaxing the identification assumptions of our instrument using the method proposed by Nevo and Rosen (2012). In this method, the instrument is allowed to be correlated with the error term of the hedonic regression, while providing a meaningful set identification result. Using this method, we could not reject the null hypothesis that the true effect of neighborhood homeownership rate on house prices is zero. Unfortunately, the bounds for the true effect were estimated with some imprecision.

Our results are in line with the results reported by Barker and Miller (2009) and Engelhardt et al. (2010) who find that the beneficial effects of homeownership on several outcomes may have been exaggerated in earlier literature. The results also suggest that building type may influence residents' incentives to invest in their neighborhood. Explicit consideration of building type should be a fruitful avenue in future research concerning various neighborhood effects.

Our results are also policy relevant as they suggest that the adverse efficiency effects of lenient tax treatment of owner-occupied housing are not offset by positive neighborhood externalities. This means that the tax favored status that homeownership enjoys in many countries should be scrutinized, at least when a large portion of the tax subsidy is directed to owner-occupiers in multi-storey buildings as in Finland. Possible reforms include taxing the imputed rental income from owner-occupied housing the same way as other capital income, reducing the mortgage interest deduction or making landlords' rental income tax free. The last two suggestions would not result in a neutral tax system, but could be steps toward a more level playing field for homeowners and renters in terms of housing costs.

Of course the results from this paper concern homeownership in multi-storey buildings and may not generalize to other building types. On the other hand, because of agency problems related to renting in single-family houses, they are likely to be owner-occupied even in the absence of tax subsidies. If this is the case, tax subsidies to homeowners in single-family houses are also redundant in the sense that they do not increase homeownership.

References

Amior, M. and J. Halket (2011): Do Households Use Homeownership to Insure Themselves? Evidence Across U.S. Cities. Mimeo.

Autor, D.H., C.J. Palmer and P.A. Pathak (2011): Housing Market Spillovers: Evidence from the End of Rent Control in Cambridge Massachusetts. MIT mimeo.

Bajari, P. and M. E. Kahn (2005): Estimating Housing Demand with an Application to Explaining Racial Segregation in Cities. *Journal of Business and Economic Statistics* 23(1), pp. 20–33.

Banzhaf, H.S. and O. Farooque (2012): Interjurisdictional Housing Prices and Spatial Amenities: Which Measures of Housing Prices Reflect Local Public Goods? NBER Working Paper No. 17809.

Barker, D. and E. Miller (2009): Homeownership and Child Welfare. *Real Estate Economics* 37(2), pp. 279–303.

Berkovec, J. and D. Fullerton (1992): A General Equilibrium Model of Housing, Taxes, and Portfolio Choice. *Journal of Political Economy* 100(2), pp. 390–429.

Boehm, T.B. & A.M. Schlottmann (1999): Does Home Ownership by Parents Have an Economic Impact on Their Children? *Journal of Housing Economics* 8, pp. 217–232.

Brasington, D. (2002): Edge versus Center: Finding Common Ground in the Capitalization Debate. *Journal of Urban Economics* 52, pp. 524–541.

Brueckner, J. K. (1997): Consumption and Investment Motives and the Portfolio Choices of Homeowners. *Journal of Real Estate Finance and Economics* 15(2), pp. 159–180.

Chetty, R. and A. Szeidl (2012): The Effect of Housing on Portfolio Choice. Mimeo.

Cleveland, W.S and S.J. Devlin (1988): Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting. *Journal of the American Statistical Association* 83, pp. 596–610.

Coulson, N. E., S.E. Hwang and S. Imai (2003): The Value of Owner-occupation in Neighborhoods. *Journal of Housing Research* 13(2), pp. 153–174.

Coulson, N. E. and L.M. Fisher (2009): Housing Tenure and Labor Market Impacts: The Search Goes on. *Journal of Urban Economics* 65(3), pp. 252–264.

Coulson, N. E. and H. Li (2011): Measuring the External Benefits of Homeownership. Mimeo.

DiPasquale, D. and E. L. Glaeser (1999): Incentives and Social Capital: Are Homeowners Better Citizens? *Journal of Urban Economics* 45(2), pp. 354–384.

Ekeland, I., J.L. Heckman and L. Nesheim (2004): Identification and Estimation of Hedonic Models. *Journal of Political Economy* 112(1), 60–109.

Engelhardt, G.V., M.D. Eriksen, W.G. Gale and G.B. Mills (2010): What are the Social Benefits of Homeownership? Experimental Evidence for Low-Income Households. *Journal of Urban Economics* 67, pp. 249–258.

Englund, P. (2003): Taxing Residential Housing Capital. *Urban Studies* 40(5–6), pp. 937–952.

Flavin, M., and T. Yamashita (2002): Owner-occupied Housing and the Composition of the Household Portfolio. *American Economic Review* 92(1), pp. 345–362.

Focus on London (2011) Housing: A Growing City. Greater London Authority, London. Available at: http://data.london.gov.uk/datastore/applications/focus-london-housing.

Gervais, M. (2002): Housing Taxation and Capital Accumulation. *Journal of Monetary Economics* 49(7), pp. 1461–1489.

Glaeser, E. L. (2011): Rethinking the Federal Bias towards Homeownership. *Cityscape* 13(2), pp. 5–37.

Glaeser, E. L. and B. Sacerdote (2000): The Social Consequences of Housing. *Journal of Housing Economics* 9(1–2), pp. 1–23.

Glaeser, E.L. and J.M. Shapiro (2003): The Benefits of the Home Mortgage Interest Deduction. " In J.M. Poterba (ed.): *Tax Policy and the Economy* 17, pp. 37–82. Chicago, IL: The University of Chicago Press.

Haurin, D., T.L. Parcel and J.R. Haurin (2002): Does Homeownership Affect Child Outcomes? *Real Estate Economics* 30(4), pp. 635–666.

Head, A. and H. Lloyd-Ellis (2012): Housing Liquidity, Mobility and the Labour Market. *Review of Economic Studies* forthcoming.

Hendershott, P. H. and M. White (2000): The Rise and Fall of Housing's Favored Investment Status. *Journal of Housing Research* 11(2), pp. 257–275.

Hilber, C. (2010): New Housing Supply and the Dilution of Social Capital. *Journal of Urban Economics* 67(3), pp. 419–437.

Hilber, C. (2011): The Determinants of Homeownership: Panel Data Evidence from Europe. London School of Economics, mimeo

Hills, J. (1991): Distributional Effects of Housing Subsidies in the United Kingdom. *Journal of Public Economics* 44(3), pp. 321–352.

Hoff, K. and Sen A. (2005): Homeownership, Community Interactions, and Segregation. *American Economic Review* 95(4), pp. 1167–1189.

Kling, J.R., J.B. Liebman and L.F. Katz (2007): Experimental Analysis of Neighbourhood Effects. *Econometrica*, Vol. 75, No. 1, pp. 83–119.

Lerbs O. and C. Oberst (2011): Explaining Spatial Variation in Homeownership Rates: Results for German Regions. CESifo Working Paper No. 3377.

Li. Q. and T. Stengos (1996): Semiparametric Estimation of Partially Linear Panel Data Models. *Journal of Econometrics* 71(1–2), pp. 389–397.

Li, Q., and J.S. Racine (2007): *Nonparametric Econometrics: Theory and Practice.* Princeton University Press.

Linneman, P. (1985): An Economic Analysis of the Homeownership Decision. *Journal of Urban Economics* 17(2), pp. 230–246.

McMillen, D.P. (2010): Issues in Spatial Data Analysis. *Journal of Regional Science* 50(1), pp. 119–141.

McMillen, D.P. and C.L. Redfearn (2010): Estimation and Hypothesis Testing for Nonparametric Hedonic House Price Functions. *Journal of Regional Science* 50(3), pp. 712–733.

Nevo, A. and A.M. Rosen (2012): Identification with Imperfect Instruments. *Review of Economics and Statistics*, forthcoming.

Oswald, A.J. (1999) The Housing Market and Europe's Unemployment: A Nontechnical Paper. University of Warwick, mimeo.

Poterba, J.M. (1992): Taxation and Housing: Old Questions, New Answers. *American Economic Review* 82(2), pp. 237–242.

Poterba. J.M. and T. Sinai (2008): Income Tax Provisions Affecting Owner-Occupied Housing: Revenue Costs and Incentive Effects. NBER Working Paper No. 14253.

Redfearn, C.L. (2009): How Informative Are Average Effects? Hedonic Regression and Amenity Capitalization in Complex Urban Housing Markets. *Regional Science and Urban Economics* 39, pp. 297–306.

Reinhold, S. and T. Woutersen (2011): Endogeneity and Imperfect Instruments: Estimating Bounds for the Effect of Early Childbearing on High School Completion. Mimeo.

Robinson, P. (1988): Root-N-Consistent Semiparametric Regression. *Econometrica* 56(4), pp. 931–954.

Rosen, S. (1974): Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy* 81(1), pp. 34–55.

Rossi-Hansberg E., P.-D. Sarte and R. Owens III (2010): Housing Externalities. *Journal of Political Economy*, 118(3), pp. 485–535.

Saarimaa, T: (2011): Imputed Rental Income, Taxation and Income Distribution in Finland. *Urban Studies* 48(8), pp. 1695–1714. DOI 10.1177/0042098010377474.

Skinner, J. (1996): The Dynamic Efficiency Cost of Not Taxing Housing. *Journal of Public Economics* 59(3), pp. 397–417.

Statistics Finland (2010): Grid Database 2010. Available at:
http://www.stat.fi/tup/ruututietokanta/griddatabase2010.pdf

Appendix. Additional regression results.

Table A1. Results from linear models.

Variable	OLS, no FE Coeff.	OLS, no FE Std. Err.	OLS, FE Coeff.	OLS, FE Std. Err.	2SLS, FE Coeff.	2SLS, FE Std. Err.
Dwelling:						
log(floor area)	0.854**	0.013	0.848**	0.005	0.845**	0.005
log(age)	-0.028*	0.013	-0.056**	0.005	-0.050**	0.008
Good	0.159**	0.009	0.150**	0.006	0.150**	0.006
Satisfactory	0.067**	0.009	0.066**	0.007	0.066**	0.007
Own lot	0.081**	0.012	0.034**	0.005	0.033**	0.006
Elevator	0.000	0.008	-0.002	0.004	-0.001	0.004
Floor level	0.016**	0.002	0.016**	0.001	0.016**	0.001
Number of floors	-0.007*	0.003	-0.012**	0.001	-0.012**	0.002
Maintenance charge	-0.010*	0.005	-0.011**	0.003	-0.011**	0.003
Distance CBD	-0.026**	0.004	-0.019**	0.005	-0.013*	0.007
Nearest train or subway stop	0.013*	0.009	0.018**	0.006	0.020**	0.006
Distance to sea	-0.013*	0.008	-0.029**	0.008	-0.019*	0.013
Neighborhood:						
Homeownership rate	-0.057**	0.008	-0.024**	0.003	-0.071	0.048
log(median income)	0.052**	0.015	0.027**	0.007	0.062*	0.035
Share of college educated	0.080**	0.017	0.042**	0.006	0.063**	0.022
Unemployment rate	-0.010**	0.007	0.001	0.003	-0.004	0.007
Share of pension h'holds	0.007	0.006	0.010**	0.003	0.025	0.015
Share of h'holds with children	-0.052**	0.011	-0.025**	0.005	-0.038**	0.014
Service jobs per capita	0.015*	0.006	-0.001	0.002	-0.002	0.003
Number of buildings	0.022**	0.008	0.014**	0.003	0.012**	0.004
Population	0.011	0.010	-0.025**	0.004	-0.023**	0.004
Mean floor area of units	0.030**	0.013	0.008	0.006	-0.006	0.016

Notes: The table reports results from linear models where the dependent variable is the natural log of transaction price. All models include quarter of sale dummies. Standard errors are robust to heteroscedasticity. ** and * indicate 1 and 5 percent significance levels, respectively.

Table A2. Results from partial linear models where homeownership is not instrumented.

	nn = 150		**nn = 200**		**nn = 300**	
Variable	**Coeff.**	**Std.**	**Coeff.**	**Std.**	**Coeff.**	**Std.**
Dwelling:						
log(floor area)	0.843**	0.004	0.844**	0.004	0.846**	0.004
log(age)	-0.056**	0.006	-0.060**	0.006	-0.065**	0.005
Good	0.153**	0.006	0.153**	0.006	0.152**	0.006
Satisfactory	0.067**	0.006	0.067**	0.006	0.067**	0.006
Own lot	0.037**	0.005	0.041**	0.005	0.044**	0.005
Elevator	0.0005	0.003	-0.0005	0.003	-0.002	0.004
Floor level	0.016**	0.001	0.016**	0.001	0.016**	0.001
Number of floors	-0.011**	0.001	-0.012**	0.001	-0.013**	0.001
Maintenance charge	-0.011**	0.003	-0.011**	0.003	-0.011**	0.003
Distance CBD	0.002	0.005	0.0003	0.004	-0.006*	0.003
Nearest train or subway stop	-0.003	0.010	-0.009	0.007	-0.012*	0.005
Distance to sea	-0.047**	0.017	-0.039**	0.012	-0.038**	0.008
Neighborhood:						
Homeownership rate	-0.015**	0.003	-0.019**	0.003	-0.022**	0.003
log(median income)	0.018**	0.006	0.019**	0.006	0.020**	0.006
Share of college educated	0.024**	0.006	0.038**	0.006	0.048**	0.005
Unemployment rate	-0.006	0.003	-0.004	0.003	-0.001	0.003
Share of pension h'holds	0.001	0.003	0.003	0.003	0.006*	0.003
Share of h'holds with children	-0.012**	0.005	-0.014**	0.004	-0.017**	0.004
Service jobs per capita	-0.001	0.002	-0.001	0.002	0.001	0.002
Number of buildings	0.011**	0.003	0.013**	0.003	0.015**	0.003
Population	-0.021**	0.004	-0.025**	0.004	-0.025**	0.003
Mean floor area of units	0.002	0.006	0.001	0.006	0.003	0.006

Notes: The table reports results from partial linear models where the dependent variable is the natural log of transaction price. All models include quarter of sale dummies. Standard errors are robust to heteroscedasticity. The marking nn refers to the number of nearest neighbors or bandwidth size used in the local regressions. ** and * indicate 1 and 5 percent significance levels, respectively.

Table A3. Results from partial linear models where homeownership is instrumented.

	nn = 150		nn = 200		nn = 300	
Variable	**Coeff.**	**Std.**	**Coeff.**	**Std.**	**Coeff.**	**Std.**
Dwelling:						
log(floor area)	0.843**	0.004	0.845**	0.004	0.849**	0.005
log(age)	-0.055**	0.008	-0.064**	0.008	-0.074**	0.009
Good	0.153**	0.006	0.153**	0.006	0.153**	0.007
Satisfactory	0.067**	0.006	0.067**	0.006	0.066**	0.007
Own lot	0.037**	0.006	0.042**	0.006	0.048**	0.006
Elevator	0.001	0.003	-0.001	0.003	-0.002	0.004
Floor level	0.016**	0.001	0.016**	0.001	0.015**	0.001
Number of floors	-0.011**	0.001	-0.012**	0.001	-0.013**	0.001
Maintenance charge	-0.011**	0.003	-0.011**	0.003	-0.010**	0.003
Distance CBD	0.002	0.005	-0.001	0.005	-0.011*	0.005
Nearest train or subway stop	-0.002	0.010	-0.013	0.009	-0.021**	0.008
Distance to sea	-0.046*	0.018	-0.041**	0.013	-0.042**	0.009
Neighborhood:						
Homeownership rate	-0.024	0.038	0.004	0.038	0.031	0.041
log(median income)	0.025	0.028	0.002	0.028	-0.019	0.031
Share of college educated	0.029	0.017	0.028	0.016	0.027	0.017
Unemployment rate	-0.007	0.006	-0.001	0.006	0.006	0.007
Share of pension h'holds	0.004	0.012	-0.004	0.012	-0.012	0.014
Share of h'holds with children	-0.016	0.014	-0.0063	0.014	0.000	0.014
Service jobs per capita	-0.001	0.002	-0.0001	0.002	0.003	0.002
Number of buildings	0.011**	0.003	0.013**	0.003	0.014**	0.003
Population	-0.021**	0.004	-0.025**	0.004	-0.026**	0.004
Mean floor area of units	0.001	0.009	0.006	0.009	0.016	0.011

Notes: The table reports results from partial linear models where the dependent variable is the natural log of transaction price and homeownership is instrumented. All models include quarter of sale dummies. Standard errors are robust to heteroscedasticity. The marking nn refers to the number of nearest neighbors or bandwidth size used in the local regressions. ** and * indicate 1 and 5 percent significance levels, respectively.

BIS Department for Business Innovation & Skills

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Llywodraeth Cynulliad Cymru
Welsh Assembly Government

Spatial Economics Research Centre (SERC)
London School of Economics
Houghton Street
London WC2A 2AE

Tel: 020 7852 3565
Fax: 020 7955 6848
Web: www.spatialeconomics.ac.uk

SERC is an independent research centre funded by the Economic and Social Research Council (ESRC), Department for Business Innovation and Skills (BIS) and the Welsh Assembly Government.