YJHEC 1146 ARTICLE IN PRESS No. of pages: 21
DISK / 29/7/03 / Hilda(CE)/ Suresh(TE) DTD 4.3.1 / SPS

ACADEMIC PRESS

Available online at www.sciencedirect.com
SCIENCE DIRECT®

Journal of Housing Economics xxx (2003) xxx–xxx

JOURNAL OF HOUSING ECONOMICS

www.elsevier.com/locate/jhe

Housing market segmentation and hedonic prediction accuracy

Allen C. Goodman[a] and Thomas G. Thibodeau[b,*]

[a] Economics Department, Wayne State University, Detroit, MI 48202, USA
[b] E.L. Cox School of Business, SMU, Dallas, TX 75275, USA

Received 19 December 2002

Abstract

In an earlier paper, Goodman and Thibodeau [Journal of Housing Economics 7 (1998) 121] examined housing market segmentation within metropolitan Dallas using hierarchical models (Hierarchical Linear Models: Applications and Data Analysis Methods, Sage, Newbury Park, 1992) and single-family property transactions over the 1995:1 – 1997:1 periods. Their preliminary results suggested that hierarchical models provide a useful framework for delineating housing submarket boundaries and that the metropolitan Dallas housing market is segmented by the quality of public education (as measured by student performance on standardized tests). This paper examines whether delineating submarkets in the manner proposed by Goodman and Thibodeau improves hedonic estimates of property value. We include two additional housing submarket constructions in our evaluation: one using census tracts and one using zip code districts. Using data for 28,000 single-family transactions for the 1995:1 – 1997:1 period, we estimate hedonic house price equations for most of Dallas County as well as individually for each submarket. The parameters of the hedonic house price equations are estimated using a 90% random sample of transactions. The remaining observations are used to evaluate the prediction accuracy of the alternative housing submarket constructions. The empirical results indicate spatial disaggregation yields significant gains in hedonic prediction accuracy.

* Corresponding author.
E-mail addresses: allen.goodman@wayne.edu (A.C. Goodman), tthibode@mail.cox.smu.edu (T.G. Thibodeau).

1051-1377/$ - see front matter © 2003 Published by Elsevier Inc.
doi:10.1016/S1051-1377(03)00031-7

1. Introduction

Within urban and real estate analyses, it has become clear that submarkets differentiated by housing or neighborhood type serve important purposes in urban analyses, and even more important purposes in home and property assessment. Analysis of both point estimates of property value, as well as the variances of these estimates are critical with respect to assessment for purposes varying from property tax collection, to the valuation of residential mortgage backed securities. With the emergence of hedonic price, repeat sales, and various hybrid statistical methods, the identification, and proper characterization of housing submarkets has maintained critical importance.

This paper examines several frameworks for formulating housing submarkets, and presents methods for estimating them. We consider the following features:

Contiguity of submarkets. It is convenient, although not essential, to group adjacent properties, and/or neighborhoods together.

Hierarchical nature of submarkets. Submarkets may have hierarchical features. Neighborhoods are located within school districts, within municipalities, and within suburbs. Some of these features are nested hierarchically; some are not.

Point prediction and variance estimation. House price prediction accuracy is important. Also important is the appropriate specification of predicted variance.

Comparing formulations. Testing the formulations requires methods for considering both nested and non-nested alternatives.

Analysts have taken different approaches to identifying submarket boundaries within metropolitan areas. Zip code districts have frequently been used to identify submarkets, in large part because they were the only identifiers available on Multiple Listings Service databases. Goodman (1977) compared census block group to census tract data in evaluating neighborhood attributes, and Goodman (1981) implicitly clustered submarkets by census tracts within the different New Haven municipalities.

Goodman and Dubin (1990) propose methods for analyzing non-nested submarkets. Dale-Johnson (1983) and Bourassa et al. (1999) use factor analysis and statistical clustering techniques to assign properties to housing submarkets. Goetzmann and Spiegel (1997) examine how neighborhood amenities influence house prices using zip code districts to delineate housing submarkets.

Goodman and Thibodeau (1998) propose to identify housing submarket boundaries by developing and estimating the parameters of a hierarchical model for house prices. The basic idea is that all homes within a spatially concentrated area share amenities associated with the property's location. Consequently, the housing characteristics that determine a property's market value are nested in a hierarchy—properties within neighborhoods, neighborhoods within school zones, school zones within municipalities, and so on. The authors use the hierarchical model to delineate areas where variation in public school quality explains variation in the hedonic coefficient for property size for the 18 elementary school zones within a suburban Dallas school district. They conclude that hierarchical models provide a useful framework for delineating housing submarket boundaries. Brasington (2000, 2001) makes particular use of their findings in examining school quality and community size. His 2001 paper indicates that using both school districts and municipalities to measure communities, the rate of tax and public services capitalization into house prices is smaller for larger communities.

This paper extends the earlier analysis by comparing hedonic prediction accuracy for four alternative ways of delineating Dallas County housing submarkets: (1) no spatial disaggregation; (2) using zip code districts to delineate submarkets; (3) using census tracts to delineate submarkets; and (4) using the Goodman–Thibodeau (GT) technique for identifying housing submarkets. Our results provide a preliminary investigation of the benefits and problems of implementing the hierarchical modeling approach to defining housing submarkets.

2. Housing submarkets

Hedonic methods have provided an important means of analyzing commodities that had previously seemed extraordinarily complex. The characterization of a house as a bundle of lot size, rooms, bathrooms, floor space, as well as heating types, hardwood floors, and other qualitative characteristics permitted an explicit characterization that had been heretofore impossible.

Most authors follow Rosen's (1974) characterization of hedonic price functions being formed as envelops of bid (by buyers) and offer (by sellers) functions. Due to either supply- or demand-related factors, the normal arbitrage that would be expected to equalize prices both within and across metropolitan areas may work either slowly, or not at all. Straszheim (1975) notes "variation in housing characteristics *and prices* by location is a fundamental characteristic of the urban housing market" (p. 28). A metropolitan housing market may be segmented into smaller submarkets due to either supply- or demand-related factors. Submarkets may be defined by structure type (e.g., single-family detached, row house, town home, and condominium), by structural characteristics (property age—housing consumers may have strong preferences for newly constructed properties or for historic properties), or by neighborhood characteristics (e.g., public education and public safety). Alternatively, housing markets may be segmented by household income and race. Higher income households may be willing to pay more for housing (per unit of housing services) to maintain neighborhood homogeneity. Finally, racial discrimination may produce separate housing submarkets for white and minority households.[1]

Consider both the estimated values and the predicted variance of the hedonic price function for a set of potentially segmented markets. Let P denote the house price, z_i the ith housing characteristic, and β_i the unknown hedonic coefficient. Compare the pooled and potentially segmented submarket j samples:

$$\ln P = \sum_i \beta_i z_{i+}\varepsilon \quad \text{(Pooled)}, \tag{1a}$$

$$\ln P_j = \sum_i \beta_{ij} z_{ij} + \varepsilon_j \quad \text{(Submarket)}. \tag{1b}$$

[1] See Goodman and Thibodeau (1998) for more discussion of the segmentation literature.

Improper pooling constrains all $\beta_{ij} = \beta_i$ irrespective of whether attribute z_i even exists in submarket j. Pooled estimation of (1a) leads to an estimate of ε (and related variance σ^2), that is a weighted average of ε_j's (and related variances σ_j^2's).

How important these problems are depends on the purpose of the exercise. For overall estimation (across a metropolitan area), the pooling problems may not matter. For property tax assessment, or for the valuation of individual (or groups of) properties within a metropolitan area, they may be critical. Assuming the estimation of k parameters for each submarket, with n submarkets, and m_i observations per submarket, the standard nested test for pooled v. submarkets is $F_{k(n-1),\sum mi-k}$. This test, however, requires nested submarkets, and it is conditional on the number and the composition of the submarkets.[2]

2.1. Identifying submarket boundaries

Submarket specification has typically been performed on an ad hoc basis. Researchers stratify a sample based on prior expectations related to municipal boundaries, school districts, racial divisions, or housing types. Hedonic regressions are estimated separately for the individual submarkets and F tests determine whether the resulting reduction in sum of squared residuals is significant. If the reduction is significant, then the posited submarkets are assumed to be appropriate, conditional on the particular specification of submarkets.

Although researchers (including the authors) often impose submarket boundaries, rather than actually modeling them, if submarkets impact housing prices, the factors that define the submarkets would be expected to affect the prices. Moreover, the nesting of these factors is important. One can draw on a parallel literature in education and evaluation for an analogy. Suppose one is looking at the determinants of pupil achievement, holding pupil ability constant. There may be separate and hierarchically nested *classroom* impacts, *school* impacts, and perhaps *district* impacts.[3]

For a single-family detached house, we consider the value of the house, nested within a neighborhood, within a school district, and within a metropolitan area. Some of these effects may be nested hierarchically, such as blocks within neighborhoods. Others, such as ethnic areas, religious parishes, or housing types, may cross school or municipal boundaries, and will not necessarily be nested, hierarchically or at all.

Our previous application of hierarchical models to housing market analysis has limitations. We assumed that the quality of public education is capitalized (exclusively) in the hedonic coefficient for the square feet of living area. The underlying assumption is that school quality is capitalized in property size. We used square feet of living space to measure property size. There are alternative models that capture this relationship. For example, we could assume that school quality is capitalized in lot

---

[2] Other maintained hypotheses include the premise that the functional form is the same across submarkets, and that the variable specification is also the same across submarkets.

[3] There is a considerable literature on improving the efficiency of such estimates through hierarchical linear modeling (Bryk and Raudenbush, 1992).

size, or in both lot size and the square feet of living space. There might also be a separate impact of school quality as a housing characteristic.[4]

3. The empirical hedonic specification

One objective is to determine the role that various housing characteristics play in producing accurate predictions of market values. To satisfy this objective, we examine two broad classes of hedonic specifications: (1) a parsimonious specification and (2) an expanded specification. The parsimonious specification relates the log of transaction price to dwelling size, a polynomial in dwelling age, and month of sale. The expanded specification includes numerous additional structural characteristics and is given by

$$
\begin{aligned}
\ln(V_{i,t}) = \beta_0 &+ \beta_1 * \ln(\text{AREA}) + \beta_2 * \ln(\text{SERVQ}) + \beta_3 * \text{AGE} \\
&+ \beta_4 * \text{AGESQ} + \beta_5 * \text{AGECUBE} + \beta_6 * \text{BATHS} \\
&+ \beta_7 * \text{GHSYS} + \beta_8 * \text{OHSYS} + \beta_9 * \text{NACSYS} \\
&+ \beta_{10} * \text{WACSYS} + \beta_{11} ast \text{WETBAR} + \beta_{12} * \text{FIREPL0} \\
&+ \beta_{13} * \text{POOL} + \beta_{14} * \text{DTGAR} + \beta_{15} * \text{CARPORT} \\
&+ \beta_{16} * \text{NOGAR} + \sum_{t=1}^{T} \rho_t * \text{SOLD}_t + \zeta_{i,t},
\end{aligned} \qquad (2)
$$

where

$V_{i,t}$	is the transaction price of the ith house sold in month t
AREA	square feet of living area
LNAREA	ln (AREA)
SERVQ	square feet of servant's quarters
LNSERVQ	log(SERVQ) (ln (SERVQ) = 0 if there are no servant's quarters)
AGE	age of the dwelling in decades
AGESQ	AGE squared
AGECUBE	AGE cubed
BATHS	the number of bathrooms (two one-half bathrooms are counted as one full bath)
CHSYS	central heating system (the omitted heating system category)
GHSYS	dummy variable for (non-central) gas heating system
OHSYS	dummy variable for other heating system—other heating systems include floor furnaces, wall heating systems, radiator heating systems, and no heating systems
NACSYS	dummy variable for no air conditioning system
WACSYS	dummy variable for window air conditioning system
CACSYS	dummy variable for central air conditioning system (omitted category)

[4] See Goodman and Thibodeau (1998) for additional limitations of this procedure.

YJHEC 1146 ARTICLE IN PRESS No. of pages: 21
DISK / 29/7/03 / Hilda(CE)/ Suresh(TE) DTD 4.3.1 / SPS

WETBAR	dummy variable for the presence of a wetbar
FIREPL0	dummy variable for no fireplace
FIREPL	dummy variable for the presence of at least one fireplace (omitted category)
POOL	dummy variable equal to 1 if swimming pool present and 0 otherwise
ATGAR	dummy variable equal to 1 if the property has an attached garage and 0 otherwise (the omitted category)
DTGAR	dummy variable equal to 1 if the property has a detached garage and 0 otherwise
CARPORT	dummy variable equal to 1 if the property has either an attached or a detached carport and 0 otherwise
NOGAR	a dummy variable equal to one if the property has no covered parking facility
SOLD_t	dummy variables for month of sale

Within each broad category of hedonic specifications (parsimonious vs. expanded) we examine four ways to delineate housing submarkets. The first simply ignores within metropolitan area spatial variation in house prices; the second defines submarkets using zip code districts; the third combines census tracts; while the final housing submarket construction uses the GT procedure.

4. Hierarchical models

4.1. Specification

Housing submarkets exist when the per unit price of housing exhibits spatial variation. We examine a two-level model of house price determination.[5] In the *Level 1 Model*, submarket house prices are determined by property structural characteristics:

$$Y_{ij} = X_{ij}\beta_j + r_{ij}, r_{ij} \sim N(0, \Omega_j) \tag{3}$$

for $i = 1, \ldots, n_j$ transactions within submarket j, and for $j = 1, \ldots, J$ submarkets. Y_{ij} denotes the house price for property i within submarket j, and X_{ij} denotes the structural characteristics for property i located within submarket j. Ω_j is a (potentially non-constant) diagonal matrix. The representation for the general linear model is obtained by stacking the submarket observations. Let $Y = (Y_1^{\mathrm{T}}, Y_2^{\mathrm{T}}, \ldots, Y_J^{\mathrm{T}})^{\mathrm{T}}$, $\beta = (\beta_1^{\mathrm{T}}, \beta_2^{\mathrm{T}}, \ldots, \beta_J^{\mathrm{T}})^{\mathrm{T}}$, $r = (r_1^{\mathrm{T}}, r_2^{\mathrm{T}}, \ldots, r_J^{\mathrm{T}})^{\mathrm{T}}$,

$$X = \begin{pmatrix} X_1 & 0 & 0\ldots0 \\ 0 & X_2 & 0\ldots0 \\ 0 & 0 & 0\ldots X_J \end{pmatrix} \quad \text{and} \quad \Omega = \begin{pmatrix} \Omega_1 & 0 & 0\ldots0 \\ 0 & \Omega_2 & 0\ldots0 \\ 0 & 0 & 0\ldots\Omega_J \end{pmatrix},$$

[5] Goodman and Thibodeau (1998) discuss the estimation procedures in detail.

so that

$$Y = X\beta + r, \quad r \sim N(0, \Omega). \tag{4}$$

In the framework of hierarchical models, the hedonic coefficients of the structural characteristics in the *Level 1 Model* vary across submarkets. The *Level 2 Model* is given by

$$\beta_j = W_j\delta + u_j, \tag{5}$$

where W_j is a matrix of predictors, δ is a vector of (assumed) fixed effects, and $u_j \sim N(0, \tau)$. The general *Level 2 Model* linear representation is obtained by stacking the appropriate matrices to obtain $\beta = W\delta + u$, $u \sim N(0, T)$. The *Combined Model* is obtained by substituting (11) into (10)

$$Y = XW\delta + Xu + r \quad \text{or} \quad Y = A_1\Theta_1 + A_2\Theta_2 + r, \tag{6}$$

where $A_1 = XW$, $A_2 = X$, $\Theta_1 = \delta$, and $\Theta_2 = u$.

4.2. How good are the submarkets?—three tests

Any set of submarket segmentations must address a validity issue. Three potential tests are available. A first test, following Schnare and Struyk (1976) involves the reduction of the squared error. Presumably, reduction of prediction error is important in formulating submarkets. How big any reduction should be, to be valuable, is an important question which is unresolved by standard statistical methods.

The second test is the F test for submarkets. As noted above, assuming the estimation of k parameters for each submarket, with n submarkets, and m_i observations per submarket, the standard nested test for pooled v. submarkets is $F_{k(n-1),\sum mi-k}$. This test, however, requires nested submarkets, and it is conditional on the number and the composition of the submarkets.

The third test, following Goodman and Dubin (1990), formulates a non-nested test among sample formulations using the J test, originally proposed by Davidson and MacKinnon (1981). Consider, in Fig. 1, the simplest example of a sample that could conceivably be split either North and South (the solid line), or East and West (the dashed line).

The two submarket formulations may be considered as the North–South formulation

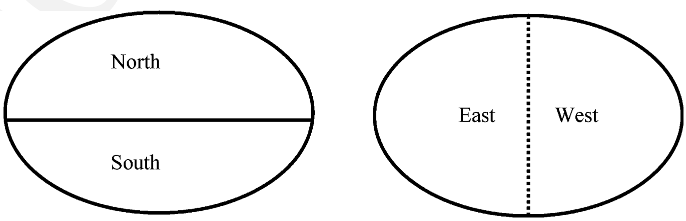

Fig. 1. Potential submarket stratifications.

$$H_0\colon\ \mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \varepsilon_0,$$

and the East–West formulation

$$H_1\colon\ \mathbf{y} = \mathbf{Z}\boldsymbol{\gamma} + \varepsilon_1.$$

H_1 cannot be written as a restriction on H_0, so conventionally nested F tests of covariance are not appropriate.

One possibility for testing the restrictions involves an artificial nesting of the two models. Following Davidson and MacKinnon (1993) and Greene (2003), define $\mathbf{Z}_1$ as the set of $\mathbf{Z}$ that are not in $\mathbf{X}$, and $\mathbf{X}_1$ likewise with respect to $\mathbf{Z}$. A standard F test can be carried out to test the hypothesis that in the augmented regression

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_1\boldsymbol{\gamma}_1 + \boldsymbol{\mu}_1,$$

vector $\boldsymbol{\gamma}_1 = 0$, with the test then reversed (with $\mathbf{Z}$ as the null hypothesis). Greene notes that this compound model may have an "extremely large" number of regressors (in this problem the number of elements of $\mathbf{Z}_1$ will always equal the number of elements of $\mathbf{X}$ unless specific submarkets are identical). This is potentially troublesome if one is comparing more than two alternative well-specified hedonic formulations, with large numbers of regressors.

The Davidson and MacKinnon J test allows the researcher to test H_0 against the alternative H_1 with the *single* parameter α:

$$\mathbf{y} = (1-\alpha)\mathbf{X}\boldsymbol{\beta} + \alpha(\widehat{\mathbf{Z}\boldsymbol{\gamma}}) + \boldsymbol{\mu}, \tag{10}$$

and reversing the test with

$$\mathbf{y} = (1-\alpha')\mathbf{Z}\boldsymbol{\gamma} + \alpha'(\widehat{\mathbf{X}\boldsymbol{\beta}}) + \boldsymbol{\mu}'. \tag{11}$$

In testing H_0 vs. H_1 and vice versa, all four possibilities may occur (reject both, neither, or either one of the two), similar to the non-nested F test. Multiple alternatives may also be tested where a vector of test statistics α (for each alternative) is distributed as an F distribution.[6]

In sum, we evaluate the prediction accuracy for eight alternative hedonic specifications—two alternative hedonic specifications for four alternative housing submarket constructions. The parsimonious specification explains variation in (the log of) house price as a function of dwelling size, dwelling age, month of sale. The expanded hedonic specification includes additional structural characteristics (e.g., number of bathrooms, type of space heating system, type of air conditioning system, presence of wetbar, fireplace, swimming pool, and type of garage). Each alternative specification is examined for four housing submarket constructions: (1) no housing submarkets within Dallas County; and housing submarkets defined using (2) zip code districts; (3) census tracts; and (4) the GT procedure.

[6] Other single parameter tests (as noted by Dubin and Goodman, 1989; Davidson and MacKinnon, 1981) include the JA test and the Cox tests.

5. The data

The database contains 28,561 transactions of single-family homes sold in Dallas, Texas between 1995:1 and 1997:1. The primary information source is the Dallas Central Appraisal District (DCAD), which estimates values for tax purposes for all real property in Dallas County.

Table 1 provides descriptive statistics for the transactions data. The average transaction price for the 28,561 properties sold over the 1995:01 – 1997:01 period is \$118,229 (\$58.20 per square foot). The average property has 1867 square feet of living space and was 28.8 years old (DWELAGE) at the time of sale (AGE = DWELAGE/10). Definitions for the variables listed in Table 1 are provided above.

TAAS95 is the average pass rate for third, fourth, and fifth grade students in the neighborhood elementary school.[7] The pass rate for each grade is obtained by averaging the pass rate for the reading and mathematics portions of the exam. Across Independent School Districts (ISDs), the average pass rate ranges from 58% for transactions within the Dallas Independent School District (DISD) to 93.7% for properties in the Highland Park Independent School District. Within the DISD, average pass rates range from below 20% to over 90%.[8]

Variables M9501 through M9612 are dummy variables for the month of sale (M9501 for January 1995, etc.) The omitted variable in the hedonic is for properties sold in January 1997.

Transactions are assigned longitudes and latitudes using MAPINFO, a geocoding software program. Properties are also assigned to their respective elementary school zone. The area includes 283 elementary school zones located in 11 Dallas County independent school districts. The elementary school zone boundaries are also geocoded using MAPINFO.

Each transaction is associated with its zip code district and census tract. There are 86 zip code districts and 415 census tracts in the area. Zip code districts are typically much larger than census tracts (or elementary school zones) and frequently cross municipal and elementary school boundaries. Table 2 provides the frequency distributions for the number of elementary school zones, independent school districts, and municipalities included in zip code districts and in census tracts. The top half provides the geography for zip code districts. Eight zip code districts are contained entirely within a single elementary school zone while one zip code district contains 20 elementary school zones. Exactly half of the zip codes cross at least one independent school district boundary and one zip code district contains portions of 9 municipalities. The bottom half of the table provides similar information for census tracts. Over half of the census tracts cross at least one elementary school zone boundary while 68 tracts (16%) span independent school district boundaries and 88 tracts (21%) span municipal boundaries. The groupings are clearly non-nested.

[7] The Texas State Department of Education makes these scores publicly available on the Internet.

[8] School quality is a multidimensional vector of attributes in itself. Dubin and Goodman (1982) use principal components analysis to reduce 25 dimensions to 5 or 6 (depending on the submarket).

Table 1
Descriptive statistics for Dallas County transaction data

Variable	*N*	Mean	Std. dev.	Minimum	Maximum
PRICE	28561	118228.45	106042.18	6500.00	1500000.00
AREA	28561	1866.77	812.4267086	528.0000000	11882.00
DWELAGE	28561	28.7777039	17.3044371	0	97.0000000
PRICEPSF	28561	58.1978671	24.7805265	10.3448276	199.6656761
LNPRICE	28561	11.4419771	0.6553438	8.7795575	14.2209757
LNAREA	28561	7.4523721	0.3901937	6.2690963	9.3827799
LNSERVQ	28561	0.0484155	0.5404616	0	8.2940496
AGE	28561	2.8777704	1.7304437	0	9.7000000
AGE2	28561	11.2758930	12.0371923	0	94.0900000
AGE3	28561	52.3099141	81.5702639	0	912.6730000
BATHS	28561	2.0908932	0.7653321	0	9.5000000
GHSYS	28561	0.0651588	0.2468101	0	1.0000000
OHSYS	28561	0.0195021	0.1382840	0	1.0000000
NACSYS	28561	0.0083330	0.0909059	0	1.0000000
WACSYS	28561	0.0949897	0.2932058	0	1.0000000
WETBAR	28561	0.1263261	0.3322223	0	1.0000000
FIREPL0	28561	0.2712090	0.4445915	0	1.0000000
POOL	28561	0.1355695	0.3423368	0	1.0000000
DTGAR	28561	0.1270264	0.3330083	0	1.0000000
CARPORT	28561	0.0589615	0.2355568	0	1.0000000
NOGAR	28561	0.0859214	0.2802528	0	1.0000000
TAAS95	28561	71.9596863	16.8855474	19.4000000	98.2000000
M9501	28561	0.0252792	0.1569747	0	1.0000000
M9502	28561	0.0173663	0.1306344	0	1.0000000
M9503	28561	0.0408249	0.1978878	0	1.0000000
M9504	28561	0.0384090	0.1921850	0	1.0000000
M9505	28561	0.0490179	0.2159092	0	1.0000000
M9506	28561	0.0515388	0.2210979	0	1.0000000
M9507	28561	0.0459018	0.2092757	0	1.0000000
M9508	28561	0.0498932	0.2177281	0	1.0000000
M9509	28561	0.0386891	0.1928564	0	1.0000000
M9510	28561	0.0383740	0.1921008	0	1.0000000
M9511	28561	0.0350478	0.1839039	0	1.0000000
M9512	28561	0.0339624	0.1811356	0	1.0000000
M9601	28561	0.0270649	0.1622754	0	1.0000000
M9602	28561	0.0342425	0.1818546	0	1.0000000
M9603	28561	0.0422954	0.2012659	0	1.0000000
M9604	28561	0.0497532	0.2174384	0	1.0000000
M9605	28561	0.0518889	0.2218067	0	1.0000000
M9606	28561	0.0475474	0.2128102	0	1.0000000
M9607	28561	0.0519590	0.2219481	0	1.0000000
M9608	28561	0.0492980	0.2164933	0	1.0000000
M9609	28561	0.0407549	0.1977253	0	1.0000000
M9610	28561	0.0373936	0.1897278	0	1.0000000
M9611	28561	0.0350128	0.1838153	0	1.0000000
M9612	28561	0.0329820	0.1785927	0	1.0000000

YJHEC 1146 ARTICLE IN PRESS No. of pages: 21
DISK / 29/7/03 / Hilda(CE)/ Suresh(TE) DTD 4.3.1 / SPS

Table 2
Dallas County geography

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Total
Number of zip code districts including																					
Elementary school zone	8	7	6	8	8	8	4	12	6	5	3	4	1	2	1	0	2	0	0	1	86
Independent school district	43	38	5																		86
Municipality	27	19	24	10	0	3	0	2	1												86
Number of census tracts including																					
Elementary school zone	129	155	90	23	14	1	2	0	0	1											415
Independent school district	347	63	5																		415
Municipality	327	68	17	3																	415

Housing submarkets are constructed using zip code districts, census tracts, and elementary school zones. A zip code district containing at least 200 transactions is classified as a separate housing submarket. If a district had fewer than 200 transactions, it is combined with another zip code district until there are at least 200 transactions in the submarket. This procedure produced 55 zip code district defined submarkets. Similarly, census tract submarkets are constructed by combining adjacent census tracts until the submarket has about 200 transactions. This procedure yielded 82 census tract defined submarkets. Finally, housing submarkets are constructed by estimating parameters of the GT hierarchical submarket model. This is accomplished in two steps. First, MAPINFO was used to identify spatially adjacent elementary school zones. The parameters of the hierarchical model are estimated *for each pair* of adjacent elementary school zones. If the estimated coefficient for the dwelling size-test score interaction variable in the hierarchical model is statistically different from zero, the school zones are assigned to separate submarkets. If the estimated coefficient of the dwelling size-test score variable is not statistically different from zero, then the two zones are assigned to the same submarket. In the second step, pairs of elementary school zones assigned to the same submarket are combined and the parameters of the hierarchical model re-estimated to test whether the combination of elementary school zones satisfies the housing submarket criteria. This procedure produced 90 housing submarkets for Dallas County.

6. Estimation results

6.1. Characteristics of estimation and prediction samples

To evaluate the prediction accuracy of the eight alternative models, the sample of 28,561 transactions is separated into two subsamples: an estimation subsample and a prediction subsample. The estimation sample is a 90% random sample of all transactions. These transactions are used to estimate the parameters of the alternative hedonic models. The remaining transactions (e.g., the prediction sample) are excluded from the estimation sample and are used to evaluate prediction accuracy for the alternative hedonic and submarket specifications. The same estimation and prediction subsamples are used for each alternative specification. Consequently, any variation in prediction accuracy cannot be attributed to differences in the underlying sample (although particular results may be artifacts of the particular samples drawn).

Table 3 provides descriptive statistics for: (1) all transactions; (2) the estimation subsample; and (3) the prediction subsample. Summary statistics are provided for transaction price, square feet of living area, and dwelling age. The estimation subsample contains 25,699 transactions and the prediction subsample contains 2862 transactions. The distributions of transaction prices, dwelling size, and dwelling age for the estimation and prediction samples are very similar to the sample of all transactions. The mean transaction price for the estimation sample is $118,128 (compared to $118,229 for all transactions) while the mean transaction price for the prediction sample is $119,133. The distribution of transaction prices in the prediction

Table 3
Descriptive statistics for single-family transactions

	All transactions ($N = 28,561$)	Estimation sample ($N = 25,699$)	Prediction sample ($N = 2862$)
Transaction price (\$)			
Mean	118,229	118,128	119,133
Std. dev.	106,042	106,296	103,746
Q3	134,500	134,125	136,000
Median	87,500	87,000	89,000
Q_1	62,500	62,300	63,500
Q_3–Q_1	72,000	71,825	72,500
Sq. ft. living area			
Mean	1867	1866	1872
Std. dev.	812	813	810
Q3	2205	2206	2191
Median	1690	1687	1713
Q_1	1312	1312	1314
Q_3–Q_1	893	894	877
Dwelling age (years)			
Mean	28.8	28.8	28.4
Std. dev.	17.3	17.3	17.1
Q_3	42.0	42.0	42.0
Median	27.0	27.0	26.0
Q_1	14.0	14.0	14.0
Q_3–Q_1	28.0	28.0	28.0

sample has a slightly smaller variance. Properties in the prediction sample were also slightly larger and younger than properties in the estimation sample, but the differences are very small.

6.2. An illustration

Before reporting results for all of Dallas County, we examine the prediction accuracy of alternative submarket constructions for one zip code district. Zip code district 75217 has 805 transactions and spans 12 census tracts and 14 elementary school zones. The 12 census tracts were combined to form two complete census tract submarkets (and portions of two additional tract defined submarkets) while the hierarchical model estimation results reduced the 14 elementary school zones to two complete housing submarkets (and portions of three additional submarkets). To insure that the same transactions will be used to evaluate the alternative housing market constructions, only the transactions common to zip code district 75217, the two complete census tract submarkets, and the two complete GT submarkets are included in this illustration.

Table 4 provides descriptive statistics for the prediction sample residuals for the expanded hedonic specification. Descriptive statistics are reported for: (1) the dollar amount of the error; (2) the absolute value of the dollar error; and (3) the propor-

Table 4
Summary statistics for prediction sample residuals

	Zip code district	Census tracts	Goodman–Thibodeau
Residual, e_i			
Mean	\$1580	\$1489	\$1707
Std. dev.	10,001	9167	9028
Q_3	7472	5632	5566
Median	2233	2380	2808
Q_1	−3114	−1644	−1851
Q_3–Q_1	10,585	7276	7417
$\|e_i\|$			
Mean	\$7911	\$6742	\$6728
Std. dev.	6225	6320	6191
Q_3	11,546	10,747	10,669
Median	6510	4363	4132
Q_1	2811	2002	1902
Q_3–Q_1	8735	8746	8766
PPE			
Mean	−0.0347	−0.0219	−0.0148
Std. dev.	0.3684	0.3353	0.3243
Q_3	0.1887	0.1315	0.1302
Median	0.0570	0.0488	0.0631
Q_1	−0.0805	−0.0527	−0.0479
Q_3–Q_1	0.2692	0.1843	0.1780

Expanded hedonic specification for zip code district 75217 ($N = 52$) ln (sales price) = f(structural characteristics, month of sale).

tional error, or PPE $= e/P$, where e is the computed residual and P is the observed transaction price. The table lists the mean, standard deviation, median, first (Q_1), and third (Q_3) quartiles and interquartile range (Q_1–Q_3) for the 52 residuals in the prediction sample. The residual is the difference between the actual transaction price and the unbiased hedonic prediction of house price.

The mean transaction price for zip code district 75217 was \$38,502. The mean prediction error was under \$1800 for the zip code, census tract, and GT defined submarkets. The standard deviation for the residual distribution was largest for the zip code submarket (\$10,001). The GT submarket construction yielded more efficient estimates of house value and reduced the standard deviation of the prediction sample residuals by 10% to \$9028. The standard deviation of the prediction sample residuals for the census tracts submarkets was \$9167.

The descriptive statistics for the distribution of the proportional error (PPE) also illustrate the dominance of the GT submarket construct *in this example*. The mean proportional error is 3.5% for the zip code submarket, 2.2% for the census tract submarkets, and 1.5% for the GT submarkets. In addition, the GT submarkets yield the lowest standard deviation of the PPE distributions: 0.37 for zip code 75217, 0.34 for the two census tract submarkets and 0.32 for GT.

In sum, spatial disaggregation for zip code district 75217 produced more accurate hedonic predictions of market value—both the census tract and GT submarkets yielded more accurate estimates of market value. Furthermore, in the GT submarkets, predictions defined by meaningful economic criteria were more efficient than the arbitrarily defined census tract defined submarkets, even though each construction contained the same number of submarkets. The conclusions we draw from this example are: (1) spatial disaggregation will generally increase the prediction accuracy of hedonic house price estimates; and (2) the procedure used to construct housing submarkets can contribute to the efficiency of the resulting market value predictions.

6.3. Dallas County results

Table 5 contains descriptive statistics for the distributions of: (1) the residuals; (2) the absolute values of the residuals; and (3) PPE $= e/P$, where e is the computed residual and P is the observed transaction price. The left half of Table 5 provides these statistics for the parsimonious hedonic specification (house price $= f$(size, age, month of sale)) while the right side of Table 5 provides these statistics for the expanded hedonic specification. The summary statistics indicate substantial increases in hedonic prediction accuracy associated with spatial disaggregation. For the parsimonious specification, the mean error for the Dallas County specification is over \$1000 while the mean errors are \$71 for the zip code submarket model and \$96 for the GT submarket model. Spatial disaggregation also significantly reduces the hedonic prediction residual variances. The standard error of the prediction sample residuals for the Dallas County parsimonious specification is \$50,946. The zip code submarket model reduces this by 20.7% to \$40,425; the census tract submarket model by 24.9% to \$38,275; and the GT submarket construction by 28.8% to \$36,283. The submarket models reduce the mean and variance of the proportional error by half. The mean PPE is −0.105 for the Dallas County model, −0.048 for the zip code model, −0.046 for the census tract model, and −0.040 for the GT model. The PPE distribution variance for the parsimonious specification is 0.1548 for the Dallas County hedonic, 0.0865 for zip code submarkets, 0.0715 for census tract submarkets, and 0.0670 for GT submarkets.

The residual distributions for the expanded specification show similar increases in prediction accuracy, but the differences between the three submarket constructions are much smaller. The mean prediction error for the Dallas County specification is about half the mean error for the parsimonious specification (\$516 vs. \$1053). The standard deviation for the Dallas County specification is \$46,326, a \$4620 reduction compared with the Dallas County parsimonious specification. The standard deviations for the submarket residual distributions are all about \$35,000, a 24.3% reduction compared to the Dallas County specification. The mean PPE for the Dallas County model is −0.082. The submarket models lower the mean PPE to −0.037 for the zip code submarket model; −0.040 for the census tract model; and −0.035 for the GT submarket model. The variance of the PPE for the expanded specification is about 8.7% lower than the variance of the PPE for the parsimonious specifications.[9]

[9] The biggest improvements occurred for the pooled and the zip code estimates, suggesting that the GT method had already captured some of the "expanded hedonic" specification.

Table 5
Summary statistics for prediction sample residuals ($N = 2862$)

	Parsimonious specification				Expanded specification			
	Dallas County	Zip code districts	Census tracts	Goodman–Thibodeau submarkets	Dallas County	Zip code districts	Census tracts	Goodman–Thibodeau submarkets
Residual, e_i								
Mean	\$1053	\$71	−\$369	\$96	\$516	−\$455	−\$881	−\$502
Std. dev.	\$50,946	\$40,425	\$38,275	\$36,283	\$46,326	\$34,911	\$35,421	\$34,829
Q_3	\$11,504	\$9060	\$8578	\$8746	\$10,467	\$8269	\$8125	\$8344
Median	−\$3309	−\$220	−\$443	\$103	−\$2893	−\$33	−\$340	\$238
Q_1	−\$19,391	−\$10,144	−\$10,294	−\$9857	−\$17,126	−\$9488	−\$9680	−\$9490
Q_3–Q_1	\$30,895	\$19,204	\$18,872	\$18,603	\$27,593	\$17,757	\$17,805	\$17,834
$\lvert e_i \rvert$								
Mean	\$28,678	\$19,185	\$18,639	\$17,887	\$25,433	\$17,103	\$17,287	\$17,264
Std. dev.	\$42,118	\$35,581	\$33,430	\$31,566	\$38,721	\$30,437	\$30,927	\$30,251
Q_3	\$32,001	\$20,542	\$20,064	\$19,104	\$27,873	\$17,845	\$18,367	\$18,317
Median	\$15,479	\$9669	\$9402	\$9281	\$13,940	\$8867	\$9048	\$9023
Q_1	\$7323	\$4310	\$4314	\$4140	\$6091	\$3856	\$3864	\$4035
Q_3–Q_1	\$24,678	\$16,233	\$15,750	\$14,964	\$21,782	\$13,989	\$14,503	\$14,282
PPE								
Mean	−0.1050	−0.0476	−0.0463	−0.0402	−0.0819	−0.0366	−0.0395	−0.0352
Variance	0.1584	0.0805	0.0715	0.0670	0.1033	0.0561	0.0574	0.0561
Q_3	0.1238	0.0976	0.0948	0.0982	0.1132	0.0914	0.0864	0.0932
Median	−0.0394	−0.0025	−0.0052	0.0010	−0.0349	−0.0006	−0.0039	0.0030
Q_1	−0.2495	−0.1203	−0.1278	−0.1172	−0.2170	−0.1131	−0.1190	−0.1111
Q_3–Q_1	0.3734	0.2179	0.2227	0.2154	0.3301	0.2046	0.2054	0.2043

Table 6 provides the frequency distributions of the PPE for each of the eight alternative specifications. The table lists the number of PPEs within ±5%; the number between 5 and 10%; and so on. It also lists the percentage of PPEs within each category and the cumulative percentages for the PPE distributions, and it illustrates the increases in prediction accuracy associated with spatial disaggregation. The prediction accuracy threshold employed by the automated valuation model (AVM) industry is that at least 50% of the predicted house prices must be within 10% of observed transaction prices. With only 34% of the expanded model predicted prices within 10% of observed transaction prices, the Dallas County model does not come close to meeting the AVM industry requirement. Each of the submarket alternatives (zip code, census tracts, and the GT submarkets) meets the industry standard with the expanded specification, but not with the parsimonious specification.

Table 6
PPE distribution summary statistics

	Parsimonious specification				Expanded specification			
	Dallas County	Zip code districts	Census tracts	Goodman–Thibodeau submarkets	Dallas County	Zip code districts	Census tracts	Goodman–Thibodeau submarkets
Frequency								
±5%	442	707	739	751	525	786	781	781
5–10%	434	630	593	606	450	661	665	652
10–15%	366	458	473	475	381	480	432	473
15–20%	308	330	327	338	340	291	315	291
20–30%	489	356	366	350	487	318	331	337
30–40%	308	145	144	144	248	133	149	142
40–50%	181	78	74	65	178	73	80	70
>50%	334	158	146	133	253	120	109	116
Percentage (%)								
±5%	15.44	24.70	25.82	26.24	18.34	27.46	27.29	27.29
5–10%	15.16	22.01	20.72	21.17	15.72	23.10	23.24	22.78
10–15%	12.79	16.00	16.53	16.60	13.31	16.77	15.09	16.53
15–20%	10.76	11.53	11.43	11.81	11.88	10.17	11.01	10.17
20–30%	17.09	12.44	12.79	12.23	17.02	11.11	11.57	11.77
30–40%	10.76	5.07	5.03	5.03	8.67	4.65	5.21	4.96
40–50%	6.32	2.73	2.59	2.27	6.22	2.55	2.80	2.45
>50%	11.67	5.52	5.10	4.65	8.84	4.19	3.81	4.05
Cumulative percentage (%)								
±5%	15.44	24.70	25.82	26.24	18.34	27.46	27.29	27.29
±10%	30.61	46.72	46.54	47.41	34.07	50.56	50.52	50.07
±15%	43.40	62.72	63.07	64.01	47.38	67.33	65.62	66.60
±20%	54.16	74.25	74.49	75.82	59.26	77.50	76.62	76.76
±30%	71.24	86.69	87.28	88.05	76.28	88.61	88.19	88.54
±40%	82.01	91.75	92.31	93.08	84.94	93.26	93.40	93.50
±50%	88.33	94.48	94.90	95.35	91.16	95.81	96.19	95.95
Total (%)	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

7. Are the submarket constructions different?

7.1. F tests

The F test for the statistical significance of spatial disaggregation is given by

$$F_{d,\sum(n_i-v_i)} = \frac{\mathrm{SSE_r}/d}{\mathrm{SSE_u}/\sum(n_i - v_i)},$$

where SSE_r is the sum of squared residuals for the (restricted coefficient) Dallas County hedonic, SSE_u is the sum of squared residuals for the (unrestricted coefficient) submarket hedonics, d is the number of restrictions, n_i is the number of transactions in submarket i and v_i is the number of estimated parameters in submarket i. The F statistics for the three submarket models computed from the estimation sample are reported in Table 7. All submarket F statistics are statistically significant at the 0.0001 level indicating the submarket hedonic equations explain variation in transaction prices better than the Dallas County hedonic.

7.2. Non-nested tests

Take the Goodman–Thibodeau submarket construct as the null hypothesis. We have

$$y = (1 - \alpha_1 - \alpha_2)\boldsymbol{Xb} + \alpha_1(\widehat{\boldsymbol{Z_1}\gamma_1}) + \alpha_2(\widehat{\boldsymbol{Z_2}\gamma_2}) + \varepsilon, \tag{12}$$

where y is the (log of) the actual transaction price, $\boldsymbol{Xb}$ is the GT regression, $\boldsymbol{Z}_1\gamma_1$ are the predicted values of the zip code (ZC) regressions, and $\boldsymbol{Z}_2\gamma_2$ are the predicted values of the census tract (CT) regressions.

α_1 and α_2 are jointly distributed $F_{2,\sum(mi-ki)}$, with $m - k$ degrees of freedom in each of the i submarkets. The test is $H_0:\ \alpha_1 = \alpha_2 = 0$ vs. $H_1:\ \alpha_1 \neq 0$ or $\alpha_2 \neq 0$. If F is significant we reject H_0, which assumes that the alternative housing market construc-

Table 7
Estimation sample spatial disaggregation test statistics

	Dallas County	Zip code districts	Census tracts	Goodman–Thibodeau
# Submarkets	1	55	82	90
Parsimonious specification				
SSE	2626.1	1160.5	1058.8	1023.7
MSE	0.1022	0.0452	0.0412	0.0398
Submarket F test		177.2	128.7	120.9
J test		2762.8	1395.3	922.8
Expanded specification				
SSE	2043.5	909.3	831.1	806.5
MSE	0.0795	0.0354	0.0323	0.0314
Submarket F test		25.9	18.1	16.7
J test		1398.1	1272.0	887.0

Table 8
Impacts of combining estimators for prediction sample

Market boundaries	Mean squared prediction error (MSPE)	Weight for combined estimator	Percent reduction in MSPE (%)
Parsimonious specification			
Goodman–Thibodeau (GT)	0.0471	0.5158	7.43
Census tracts (CT)	0.0486	0.3950	10.29
Zip codes (ZC)	0.0527	0.0892	17.27
Combined	0.0436		
Expanded specification			
Goodman–Thibodeau (GT)	0.0420	0.5014	10.48
Census tracts (CT)	0.0416	0.3951	9.62
Zip codes (ZC)	0.0435	0.1035	13.56
Combined	0.0376		

tions do not provide additional information. We compute similar test statistics with ZC as the null and with CT as the null. For the Goodman–Thibodeau submarket construct to dominate, we must fail to reject the GT null (i.e., the first J test must be insignificant), but we must reject similar hypotheses with ZC and CT as the null (both J tests must be significant significant). The non-nested test statistics appear in Table 7. All the F statistics comprising the J test are statistically significant, indicating none of the three housing market constructions dominate the alternatives.[10]

The J test also provides an indirect demonstration of the benefits of combining estimators (Fair and Shiller, 1990, 1989). Re-examining Eq. (12), note that the alternative estimators provide weights α_1 and α_2, which serve to reduce the mean squared error if α_1 and α_2 are jointly significant. Table 8 shows how the three estimates are combined to reduce the variance yet further. The best of the three, the GT estimator still reduces its MSPE by 7.43% when combined linearly with the CT and the ZC estimates. While the weights of 0.52, 0.39, and 0.09 for the GT, CT, and ZC estimates respectively, represents a result that is derived from the J test, they demonstrate the possibility of optimally combining estimators to achieve additional improvements in prediction accuracy.

8. Conclusions

This paper refines the characterization, measurement, and impact of housing submarkets. It derives, rather than imposes, a set of housing submarkets for an entire metropolitan area. It then compares the derived set of submarkets to others that may be imposed, either at the zip code or the census tract level.

[10] A reviewer has suggested that these methods could be augmented by taking spatial dependence into account, as do Goodman and Dubin (1989). Although they find little impact of spatial dependence corrections, further examination of this issue may be productive.

The first conclusion can be stated as ‘‘smaller is better.’’ The ZC, the CT, and the GT submarkets all perform better than pooled estimates by any prediction criteria that one wishes to use. Indeed, given the often arcane formulation of zip codes, it is surprising how well they characterize submarkets. Moreover, they are the easiest submarket indicators to use—everyone knows his or her zip code.

We also apply a method that we piloted in an earlier paper, for an entire metropolitan area. It certainly compares with the zip code and the census tract measures (although it does not statistically dominate either), and it appears to provide these results with the benefit of reduced variance, particularly for the parsimonious specifications that are often used in property valuation. To the extent that variance is a tangible cost in prediction (whether for property taxes or for the characterization of risk in mortgage-based securities), reduction of this variance is a substantive and important benefit. Moreover, the GT method is easily implemented and programmed, and it can be easily *updated* by intuitive criteria, whereas the others provide no transparent method for updating.

Acknowledgments

We thank David Lenze for some particularly thoughtful comments and we gratefully acknowledge the comments and suggestions of three anonymous referees. Remaining errors, of course, are our own.

References

Bourassa, S.C., Hamelink, F., Hoesli, M., MacGregor, B.D., 1999. Defining residential submarkets. Journal of Housing Economics 8, 160–183.

Bryk, A.S., Raudenbush, S.W., 1992. Hierarchical Linear Models: Applications and Data Analysis Methods. Sage, Newbury Park, CA.

Brasington, D.M., 2000. Demand and supply of public school quality in metropolitan areas: the role of private schools. Journal of Regional Science 40, 583–605.

Brasington, D.M., 2001. Capitalization and community size. Journal of Urban Economics 50, 385–395.

Dale-Johnson, D., 1983. An alternative approach to housing market segmentation using hedonic price data. Journal of Urban Economics 11, 311–332.

Davidson, R., MacKinnon, J., 1981. Several tests for model specification in the presence of alternative hypotheses. Econometrica 49, 781–793.

Davidson, R., MacKinnon, J., 1993. Estimation and Inference in Econometrics. Oxford University Press, NewYork. pp. 384–387.

Dubin, R.A., Goodman, A.C., 1982. Valuation of neighborhood characteristics through hedonic prices. Population and Environment 5, 166–181.

Fair, R.C., Shiller, R.J., 1990. Comparing information in forecasts from econometric models. American Economic Review 80, 375–389.

Fair, R.C., Shiller, R.J., 1989. Informational content of ex ante forecasts. Review of Economics and Statistics 71, 325–331.

Goetzmann, W.N., Spiegel, M., 1997. A spatial model of housing returns and neighborhood substitutability. Journal of Real Estate Finance and Economics 14, 11–31.

Goodman, A.C., 1977. A comparison of block group and census tract data in a hedonic housing price model. Land Economics 53, 483–487.

YJHEC 1146 ARTICLE IN PRESS No. of pages: 21
DISK / 29/7/03 / Hilda(CE)/ Suresh(TE) DTD 4.3.1 / SPS

Goodman, A.C., 1981. Housing submarkets within urban areas: definitions and evidence. Journal of Regional Science 21, 175–185.

Goodman, A.C., Dubin, R.A., 1990. Non-nested tests and sample stratification: theory and a hedonic example. Review of Economics and Statistics 72, 168–173.

Goodman, A.C., Thibodeau, T.G., 1998. Housing market segmentation. Journal of Housing Economics 7, 121–143.

Greene, W.H., 2003. Econometric Analysis, fifth ed. Prentice-Hall, Upper Saddle River, NJ. pp. 152–154.

Rosen, S., 1974. Hedonic prices and implicit markets: product differentiation in pure competition. Journal of Political Economy 82, 34–55.

Schnare, A.B., Struyk, R.J., 1976. Segmentation in urban housing markets. Journal of Urban Economics 3, 146–166.

Straszheim, M.R., 1975. An Econometric Analysis of the Urban Housing Market. National Bureau of Economic Research, New York.