Investigating the effects of service and management on multifamily rents: a multilevel linear model approach

Qiong Peng & Gerrit J. Knaap

Journal of Housing and the Built Environment

ISSN 1566-4910

J Hous and the Built Environ
DOI 10.1007/s10901-020-09786-w

ONLINE FIRST

Journal of Housing and the Built Environment
https://doi.org/10.1007/s10901-020-09786-w

ARTICLE

Investigating the effects of service and management on multifamily rents: a multilevel linear model approach

Qiong Peng[1,2] · Gerrit J. Knaap[1,2]

Received: 22 November 2019 / Accepted: 25 September 2020

Abstract

Unlike the large body of research on the determinants of single family prices and rents, the determinants of multifamily rents has received much less exploration. Using a recent and comprehensive micro-level dataset of multifamily housing units in Montgomery County, Maryland, USA, we applied a multilevel linear model with random coefficient to explore the determinants of multifamily rents, including the effects of service and management attributes. The findings are as follows: (1) first we find that a multilevel linear model is better suited to address datasets that include multiple apartment units in a smaller set of facilities, (2) for certain datasets—including ours–a random coefficients model outperforms both an OLS and random intercept model and (3) the effects of service and management variables on multifamily rents vary across types of service and management. Pet allowance, availability of short-term leasing options, and storage service availability increase rents significantly. Conversely, offering units to property employees and services to those with a disability decrease rents significantly.

Keywords Housing · Service and management · Hedonic model · Multilevel linear model · Hierarchical linear model · Montgomery county · Multifamily rent

1 Introduction

Multifamily housing[1] plays a vital role in the U.S. real estate marketplace. As a primary rental housing type, multifamily housing accommodated 17.8 million multifamily renter households (14.5% of total households) in the United States in 2013.[2] Even more, multifamily housing is

[1] The definition of a multifamily house varies by organization. The standard industry definition of multifamily housing is a structure with five or more units. According to this definition, multifamily housing is also generally considered to be renter-occupied housing, while owner-occupied condominiums are usually not considered to be multifamily housing units even though they may be located in multifamily structures.
[2] The number of households in multifamily rental housing comes from Multifamilybz.com. The definition of multifamily housing for this calculation is a structure with five or more units that is renter-occupied.

✉ Qiong Peng
xqpengumd@gmail.com

1 Department of Urban Studies and Planning, University of Maryland, College Park, MD 20742, USA

2 National Center for Smart Growth, University of Maryland, 1219D Preinkert Field House, College Park, MD 20742, USA

Published online: 01 October 2020

a common option for low- and medium-income households. Unlike single-family housing, multifamily housing provides unique services and management features. Multifamily housing services could include pet allowance, maid service, storage service, special services for those with disabilities, and more. Management features could include an option for a short-term leases, renovations, offering units for property management company employees, and more. As a focus of multifamily housing research and practice, investigating the effects of service and management on multifamily rents has important implications for optimizing management, setting rents, attracting potential tenants, and promoting housing affordability. However, there is little research on the effects of service and management on rents.

In addition, most multifamily housing hedonic analyses ignore nature of hierarchical data, a commonly referenced data type in the housing literature, and just employ an OLS approach. Applying the OLS approach with hierarchical data will underestimate standard errors, and could lead to false inferences (Krull and MacKinnon 2001). In contrast, multilevel linear modelling (MLM) approach addresses the multilevel nature of data and can account for both spatial autocorrelation and heteroscedasticity of residuals (Djurdjevic et al. 2008). However, there are few multifamily housing hedonic analyses that utilize the MLM approach. The purpose of this paper is to estimate the market values (implicit prices) of service and management variables within a hedonic framework. That is, we use transaction rents as the dependent variable; include indicators of service and management, structure characteristics, amenities, and locational characteristics; and conduct multilevel linear regression analyses.

This study addresses the following research questions: (1) Should a multilevel linear model be used in multifamily housing hedonic analysis? (2) If so, random intercept model or random coefficient model should be selected in our case? (3) Does provision of service and management impact multifamily rents?

We proceed as follows. We first review previous research regarding the evaluation of effects of service and management on multifamily rents and the application of multilevel linear model in housing and real estate. After introducing the dataset and modeling approaches, we compare the results of the traditional OLS model, the random intercept model, and the random coefficient model and, in particular, explain the results of the random coefficient model. The final section replicates the key findings and offers implications for the housing and real estate fields.

2 Literature review

We start the literature review by visiting the scopes of apartment amenities, services, and management. "The apartment units contain a set of various amenities. These range from a swimming pool to covered parking to a modern kitchen"(Sirmans et al. 1989). "Likewise, a set of various services is provided to the apartments, ranging from maid service to restrictions on residents or pets"(Sirmans et al. 1989). Regarding the scope of management, it ranges from managers holding professional designations to management company size to managed units (Sirmans and Sirmans 1992, Benjamin and Lusht 1993, Appelbaum and Glasser 1982). In this analysis, we follow the literature.

In the last three decades, there has been a growing body of research focused on assessing multifamily housing characteristics and locational characteristics on multifamily rents (for reviews, see Jud et al. (1996) and Zietz (2003)). These studies use a hedonic approach to estimate implicit prices of complex attributes, as well as locational characteristics. However, there are few existing studies that assess the effects of service and management on rents, despite the fact that service and management play a critical role in optimizing management and setting rents. The lack of related studies can be attributed to the absence of data related to micro-level service and management in multifamily housing analysis. There are exceptions. Sirmans and Sirmans (1992) use professional designations as a proxy for quality of service provided by the property management and find that the holding of such designations has a positive effect on monthly rent. Security service and management—such as gated access restrictions and 24-h security—enhanced rental rates and rents in garden-style apartments (WG Hardin III and Cheng 2003) and high-rise apartments (Benjamin et al. 1997). Multifamily properties owned and managed by real estate investment trusts (REIT) have higher effective rents than non-REIT-owned properties (William Hardin et al. 2009). Regarding the relationship between property management and rent, Benjamin et al. (2007) examined Atlanta apartments and found that larger-scale owners and local property managers earn higher effective rents (Benjamin et al. 2007). Pet allowance was found to have a positive effect on rent (WG Hardin III and Cheng 2003; Sirmans et al. 1989). The effects of short-term renovation on apartment rents is small but not significant (Mejia and Potter 2015). The availability of maid service has no significant effect on rent (Sirmans et al. 1989), whereas age restrictions have a positive effect on the price of condominiums (Guntermann and Norrbin 1987).

Three main approaches are common within the hedonic framework. First, OLS is used widely and broadly in the existing housing hedonic literature. Most OLS models use semi-log or double-log specifications because heteroscedasticity is substantially reduced or eliminated in semi- or double-log models (Clapp and Salavei 2010). Although the OLS approach offers many advantages, somewhat restrictive assumptions—specifically, independence and homoscedasticity of the residuals–are required. The OLS approach cannot address the hierarchical structure of certain data and thus underestimates standard errors, which can lead to false inferences (Krull and MacKinnon 2001). In addition, given the fact that the OLS approach fails to account for spatial autocorrelation, the estimates of the OLS model are no longer reliable. Second, to account for the hierarchical data structure and spatial autocorrelation there is a growing body of research that uses multilevel linear models. Those studies have focused on single-family housing (Brunauer et al. 2013; Chasco and Gallo 2013; Djurdjevic et al. 2008; Gelfand et al. 2007; Giuliano et al. 2010; Glaesener and Caruso 2015; Goodman and Thibodeau 1998; Jones and Bullen 1993; Orford 2002; Shin et al. 2011; Treg 2010; Uyar and Brown 2007). However, there are only few studies that apply multilevel linear models (MLM) for multifamily hedonic analysis, even dozens or even hundreds of units can be nested within an apartment complex.

In addition previous multifamily housing studies have often used random intercept models, not random coefficient models. Random coefficient models, in some cases, could be more appropriate than the random intercept models. Third, some studies use the OLS approach and adjust standard errors for clustering to address the multilevel nature of the data, instead of the MLM approach. However, according to Cheah (2009), modeling the clustering of data using multilevel methods is a better approach than adjusting the standard errors of the OLS estimates. More specifically, the MLM approach can account for both spatial autocorrelation and heteroscedasticity of residuals, whereas the OLS with clustering standard errors is unable to do that.

Overall, the gaps in assessing the effects of service and management on multifamily rents within a hedonic framework are: (1) Few studies examine the effects of service and management on multifamily rents, despite the fact that service and management are critical factors in optimizing management and in setting rents; (2) Most multifamily hedonic analyses ignore the hierarchical data structure, which may result in false inferences; (3) Even if we conduct the MLM approach in multifamily hedonic analyses, random coefficient models are unintentionally ignored. However, the random coefficients models, in some cases, are more appropriate than the random intercept models. This study aims to fill the gaps in the literature by (1) testing if a multilevel linear model can be used to provide more credible estimates of the determinants of multifamily rents; and (2) estimating the effects of service and management factors on rents based on an up-to-date comprehensive unit-level multifamily dataset.

3 Data

Our empirical dataset comes from four data sources. The first data source is the Montgomery County rental housing survey conducted by the Department of Housing and Community Affairs (DHCA) in 2018. The survey includes 79.5% of 930 complexes throughout Montgomery County, Maryland. The rental housing survey has a hierarchical structure with two levels. The lower level is the unit level, which has three variables–unit rent, number of bedrooms, and a dummy variable indicating if a unit is occupied by a property company employee. The unit rent is a transaction rent for which the payment is made to property managers or landlords. We exclude observations that are vacant. The number of bedrooms indicates how many bedrooms are in a unit. Employees of property companies may get a discounted rent if they live in property owned or managed by their companies. The dummy variable indicating if a unit is occupied by an employee is used to control this situation. The upper level, the complex level, addresses information regarding service and management, amenities, and structure characteristics. Using longitudinal and latitudinal information, we geocode the observed complexes to get locational characteristics, such as proximity to shopping centers, public schools, light rail transit stations, bus stops, highway routes, highway exits, and crime using information from the second data source, Data Montgomery. To indicate the performance of public schools, we requested access to the 2013 Maryland State Education Indicators, the third data source. These indicators offer information on average reading scores for fifth-graders for each public elementary school in Montgomery County. Our fourth data source is Open Data DC-a public website providing data about Washington D.C. In order to control for the distance of each complex from the central business district (CBD) of Washington, D.C.

After we cleaned the raw data by dropping duplicated entries, vacant units, and outliers, our final dataset included 73094 units from 740 complexes. Figure 1 shows the locations of multifamily complexes in our study. Information about our observations is grouped into five categories: unit-level attributes, service and management information, complex structure characteristics, complex amenities, and locational characteristics. Table 1 lists variables included in the analysis as they correspond with each category. For each observation, variables associated with unit-level attributes include transaction rent, number of bedrooms, and a dummy variable indicating if the unit is employee-occupied. It should be noted that we have a unit-level service and management variable indicating if a unit is employee-occupied. Complex-level service and management variables contain the following: a dummy variable indicating if the complex allows pets in the building; a dummy variable indicating if renovations had been made in the prior 12 months; a dummy variable indicating whether the complex offers short-term leasing options; a dummy variable indicating if the complex offers a storage service; and a dummy variable indicating if the complex provides special services for those with disabilities.

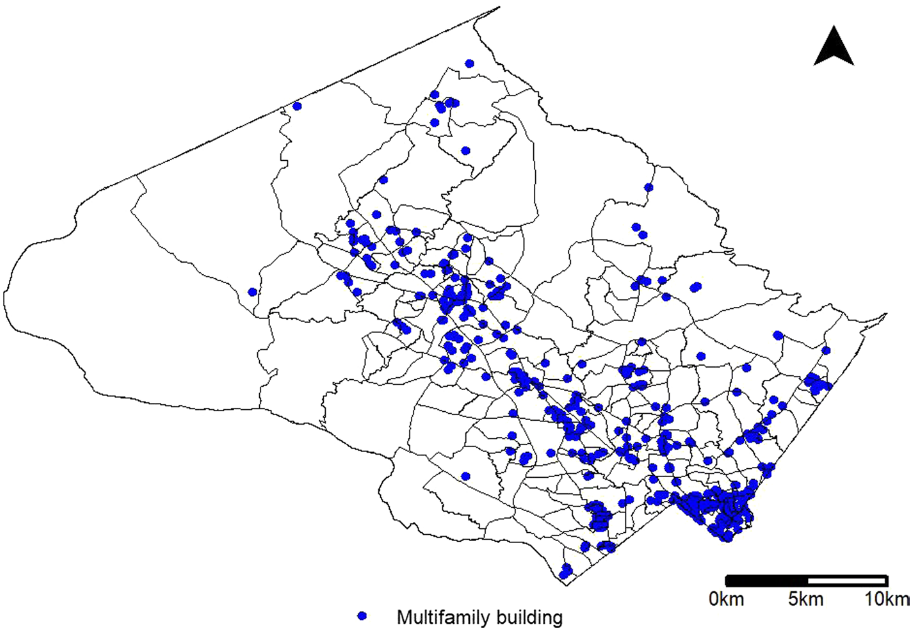

Fig. 1 Location of multifamily complexes in the study

The complex-level structure attributes include: a continuous variable indicating the age of the structure; a categorical variable indicating the structure type for the complex (for which the base group is "townhouse"); and the total number of units for the complex. The complex-level amenities include: a dummy variable representing if the complex has a gym center; a dummy variable indicating if the complex provides a common laundry room; a dummy variable indicating if the complex has a parking lot or garage; and a dummy variable indicating if the complex has a swimming pool. Variables for locational characteristics include school quality, which uses average reading scores for fifth-graders for each school to indicate school quality; crime rate, which describes the correlation of the number of properties to violent crimes reported to law enforcement agencies per 100,000 total population at the census tract level in 2017; a dummy variable indicating if a complex is within one mile from its nearest light rail station; the distance between the complex and the nearest light rail transit station; the distance from highways measured in two bands up to 0.5 mile; distances from highway exits to complexes, measured in 0.25 mile bands, up to two miles; the distance between the complex and the nearest shopping center; and the distance between the complex and the CBD of Washington, D.C. All distances were measured in Euclidean terms.

The dataset we construct presents four merits and one possible weakness compared with previous multifamily housing studies. First, the dataset includes unit-level attributes for multifamily housing, which have seldom been used in previous studies. The unit-level data usually contains more detailed information, and therefore, the estimation of implicit price could be more accurate based on that data. Second, our dataset has a good measure of market rents, reported by the property manager for each unit on a legally required survey. As the indicator of the market value of rental units, this measure is much better than those collected from self-selected occupants and aggregated over varying geographies. Such

Table 1 Definition of variables

Category	Variable Name	Definition
1. Unit-level attributes	Rent	Monthly rent of unit
	Bedrooms	Number of bedrooms of a unit
	EmployeeOccupied	If a unit is occupied, by property employee, assign 1. Otherwise,0
2. Complex-level characteristics	**Pets**	If a complex allows pets, assign 1. Otherwise, 0
2.1 Service and management	**ShortLease**	If a complex provides a short-term lease contract (fewer than 12 months) for tenants, assign 1. Otherwise, 0
	Renovation	If a complex has implemented a renovation in the last 12 months, assign 1. Otherwise, 0
	Storage	If a complex provides storage service, assign 1. Otherwise, 0
	Disability	If a complex provides service for those with a disability, such as by offering specially designed rooms for those with a disability, assign 1. Otherwise,0
2.2 Complex amenities	Gym	If there is a gym center, assign 1. Otherwise, 0
	Laundry	If there is a laundry room, assign 1. Otherwise, 0
	Parking	If a complex provides a parking lot or garage, assign 1. Otherwise, 0
	Swimming	If a complex has a swimming pool, assign 1. Otherwise, 0
2.3 Structure attributes	Age	Age of building structure
	Structure	Structure type of building. If it is a garden apartment with 1–4 stories, assign 1. If it is a midrise apartment with 5–8 stories, assign 2. If it is a high-rise apartment with 9 + stories, assign 3. If it is a townhouse, assign 0
	ComplexSize	Number of units in a complex

Table 1 (continued)

Category	Variable Name	Definition
2.4 Locational characteristics	School quality	Use average fifth-grade reading score per school as indicator of the quality of elementary schools
	Crime	Crime rate at the census tract level (2017)
	Metro_1 mile	If a complex is within one mile from a light rail transit station, assign 1. Otherwise, 0
	DisMetro	Distance between a complex and its nearest light rail transit station
	Bus_0.25 mile	< 0.25 mile from bus stop (dummy)
	Highway_0.25mile	< 0.25 mile from highway (dummy)
	Highway_0.5 mile	0.25–0.5 mile from highway (dummy)
	Exit_0.25 mile	< 0.25 mile from highway exit (dummy)
	Exit_0.5 mile	0.25–0.5 mile from highway exit (dummy)
	Exit_0.75 mile	0.5–0.75 mile from highway exit (dummy)
	Exit_1 mile	0.75–1 mile from highway exit (dummy)
	Exit_1.25 mile	1–1.25 miles from highway exit (dummy)
	Exit_1.5 mile	1.25–1.5 miles from highway exit (dummy)
	Exit_1.75 mile	1.5–1.75 miles from highway exit (dummy)
	Exit_2 mile	1.75–2 miles from highway exit (dummy)
	DisShopping	Distance between complex and its nearest shopping center
	DisCBD	Distance between complex and the CBD of Washington, D.C

Names in bold are service and management variables alternative indicators of rent are not often market-determined and, therefore, are subject to inconsistencies inherent to non-market valuation techniques. Third, the dataset includes detailed service and management information, which makes the dataset unique to the literature. Both researchers and practitioners are consistently interested in the effects of various services and management factors on multifamily rents but are not able to access these variables. Fourth, the dataset is based on the latest, population-scale multifamily surveys and has a large sample size. The dataset has a possible weakness, however. There are only two unit-level explanatory variables: the number of bedrooms and a dummy variable indicating if a unit is occupied by a property company employee. This may lead us to an omitted-variable bias problem. For example, the dataset is missing the size of the unit in terms of square footage. While our dataset offers information on the number of bedrooms for each unit, this may not be an efficient proxy for square footage; but, we believe it is an acceptable proxy for unit size.

The summary statistics for each variable are reported in Table 2. The average transaction rent is \$1,573 per month. The average number of bedrooms is 1.57, while 0.8% of units are employee-occupied. Regarding complex-level service and management characteristics, 77.5% of units are nested in complexes that allow pets in their complexes; 46.2% of units are nested in complexes that offer short-term lease options for tenants; 8.2% of units are nested in complexes that have made observable renovations in the prior 12 months; 40.2% of units are nested in complexes that provide storage services; and 13.4% of units are nested in complexes that offer service for those with disabilities. Regarding structure attributes, the average age of the complex structures is 37.6 years. The average complex size is 308 units. Furthermore, complex-level amenities include 60.6% of units nested in complexes that have gym centers, while 55.5% of units are nested in complexes that have laundry rooms, 88.3% of units are nested in complexes that have parking lots or parking garages, and 64.8% of units are nested in complexes that have swimming pools. With regards to locational characteristics, we can see that the average fifth-grade reading score is 91.7 (out of 100), while the average crime rate is 9,751 incidents per 100,000 people a year. The average distance to the nearest rail transit station is 1.984 miles while the average distance to the nearest shopping center is 0.353 miles. The average distance to the CBD of Washington, D.C. is 12.15 miles.

4 Models

As a valuation technique, hedonic price modeling has been used broadly and has been noted to generate fruitful results in housing literature since Lancaster (1966) and Rosen (1974) began to explore the determinants of housing prices. Most hedonic models cited in the housing literature are one-level OLS regression models; however, residential location decisions are inherently hierarchical. The search process for a house begins with choosing a town or city to live in, followed by a neighborhood, and finally, a house, given the neighborhood and the town (Quigley 1985). This search process could happen within a multifamily context. Moreover, it is reasonable to assume that individual units in the same complex will have correlated responses on rent since they share the same location, amenities, services, and management. In other words, the multifamily dataset is hierarchical in structure—individual units are nested within complexes. The one-level OLS method assumes that all observations are independent. It is apparent that the one-level OLS based on multifamily hierarchical dataset violates that assumption, and the one-level OLS will

Table 2 Summary statistics

Sub-category	Variable name	Mean	Std.dev	Min	Max
1. Unit-level attributes	Rent	$1573	$458.41	$50	$3673
	Bedrooms	1.57	0.729	0	5
	EmployeeOccupied	0.008	0.089	0	1
2. Complex-level characteristics	**Pets**	0.775	0.417	0	1
2.1 Service and management	**ShortLease**	0.462	0.498	0	1
	Renovation	0.082	0.274	0	1
	Storage	0.402	0.490	0	1
	Disability	0.134	0.340	0	1
2.2 Complex amenities	Gym	0.606	0.488	0	1
	Laundry	0.555	0.496	0	1
	Parking	0.883	0.321	0	1
	Swimming	0.648	0.477	0	1
2.3 Structure attributes	Age	37.6	21.005	1	133
	Structure	Category variable Number of units in garden apartment with 1-4 stories: 43145 Number of units in midrise apartment with 5–8 stories: 3066Number of units in high-rise apartment with 9 + stories: 24176 Number of units townhouse: 2707			
	ComplexSize	308	220.2	1	1067
2.4 Locational characteristics	School quality	91.7	6.102	73.0	100.0
	Crime	9751	12,028	953	66,161
	Metro_1 mile	0.317	0.465	0	1
	DisMetro (in mile)	1.9842	2.098	0.021	13.768
	Bus_0.25 mile	0.579	0.493	0	1
	Highway_0.25mile	0.057	0.231	0	1
	Highway_0.5 mile	0.0868	0.281	0	1
	Exit_0.25 mile	0.0304	0.171	0	1
	Exit_0.5 mile	0.0974	0.296	0	1
	Exit_0.75 mile	0.0691	0.253	0	1
	Exit_1 mile	0.0732	0.260	0	1
	Exit_1.25 mile	0.166	0.372	0	1
	Exit_1.5 mile	0.118	0.322	0	1
	Exit_1.75 mile	0.178	0.382	0	1
	Exit_2 mile	0.0639	0.244	0	1
	DisShopping (in mile)	0.3536	0.283	0.013	4.489
	DisCBD (in mile)	12.15	5.467	5.19	30.73

Variables in bold are service and management variables. 73094 units in 740 complexes

yield biased results. To account for the hierarchical search process and multifamily hierarchical data structure, it is more appropriate to implement a multilevel linear model (MLM) (also known as a hierarchical linear model, or a mixed model) in the multifamily hedonic analysis. The MLM approach has the following benefits compared with a one-level OLS when the dataset is hierarchical.

1 One-level OLS will underestimate the standard error, and thus overestimate test statistics and the statistical significance of the parameters; this can result in spuriously significant effects (Krull and MacKinnon 2001). Column 1 of Table 3 reports the results of the one-level OLS regression. We can see that nearly all coefficients of explanatory variables are statistically significant, which differs from the results of the MLMs in Column 2 and Column 3. Although we can correct standard errors of one-level OLS through clustering method in some extents, Cheah (2009) argues that modeling hierarchical data using multilevel methods is better than fixing the standard errors of the OLS estimate.
2 The MLM approach can account for both spatial correlation and heterogeneity of residuals (Djurdjevic et al. 2008). Housing prices or rents are more likely to be similar within submarkets than across submarkets, which implies spatial dependence in the error terms of the hedonic equation. In our case, rents of units are much more likely to be correlated within the same apartment complex (submarket), given the similarities in rents within an apartment complex. Random coefficient models allow intercepts and implicit prices of characteristics to vary across apartment complexes. Thus, the random coefficient models can account for both spatial correlation and heterogeneity.
3 The MLM specification can help in examining the variability of the coefficient across groups and can be useful in examining cross-level interaction.

The MLMs in our case have two levels.[3] Level 1 is the unit level while Level 2 is the complex level. The MLM can take many forms depending on which predictors are included at each level, and whether the model accounts for a random intercept, a random coefficient, or both. We start with the simplest MLM model: no explanatory variables at any level. Raudenbush and Bryk (2002) refer to it as the "unconditional model."

4.1 Unconditional model

The first level of the unconditional two-level model is:

$$y_{ij} = \beta_j^0 + \varepsilon_{ij}, \varepsilon_{ij} \sim N(0, \sigma^2) \tag{1}$$

where the ij subscript refers to the ith individual unit in the jth complex; y_{ij} is the natural log of rent of unit i in complex j. β_j^0 is the varying intercept across complex j. ε_{ij} is a stochastic error term following a normal distribution with zero mean and variance of σ^2.

The second level of the two-level model assumes the coefficient β_{0j} can be expressed as:

$$\beta_j^0 = \gamma^0 + \mu_j, \mu_j \sim N(0, \tau^2) \tag{2}$$

[3] Three-level MLM could be an alternative: Level 1 (unit level), Level 2(complex level), and Level 3 (neighborhood level). Level 3 could include a few explanatory variables. Two of them are school quality and census tract crime rate. The school district boundaries do not entirely contain boundaries of certain census tracts, which may complicate the analysis when we applied a three-level MLM. For simplicity and feasibility, this study utilizes a two-level MLM.

Table 3 Regression Results

	Model 1 One level standard linear hedonic model (OLS)		Model 2 Random intercept model		Model 3 Random coefficient model	
Fixed eff	Coef	Std.error	Coef	Std.error	Coef	Std.error
Level 1 (unit level)						
Intercept	8.5959***	0.0995	7.9334***	0.8262	7.6866***	0.8212
Bedrooms	0.1822 ***	0.0013	0.1806***	0.0010	0.1842***	0.0040
EmployeeOccupied	− 0.0496***	0.0104	− 0.0763***	0.0071	− 0.0963***	0.0243
Level 2 (complex level)						
Pets	− 0.0041	0.0026	0.0450*	0.0198	0.0441*	0.0196
ShortLease	0.0944***	0.0021	0.0635**	0.0229	0.0664**	0.0224
Renovation	0.0118**	0.0036	0.0031	0.0322	− 0.0039	0.0317
Storage	0.0988 ***	0.0023	0.0843***	0.0227	0.0857***	0.0223
Disability	− 0.0555***	0.0028	− 0.0685*	0.0321	− 0.0692*	0.0313
Gym	0.0278***	0.0027	0.0637*	0.0312	0.0338	0.0304
Laundry	− 0.0270***	0.0022	0.0152	0.0186	0.0029	0.0184
Parking	0.0257***	0.0033	− 0.0310	0.0215	− 0.0329	0.0215
Swimming	0.1020***	0.0027	0.1320***	0.0304	0.1407***	0.0296
Log(Age)	− 0.0662***	0.0014	− 0.0535***	0.0150	− 0.0390**	0.0146
ComplexSize	0.0366***	0.0013	0.0164*	0.0075	0.0218**	0.0075
Log(School)	0.2333***	0.0163	0.2122	0.1360	0.2312^	0.1348
Log(Crime)	− 0.0353***	0.0014	− 0.0337*	0.0144	− 0.0316*	0.0141
Metro_1 mile	0.0696***	0.0042	0.1287**	0.0394	0.1315***	0.0387
Metro_1 mile:DisMetro	− 0.1224***	0.0031	− 0.1471***	0.0250	− 0.1492***	0.0250
Bus_0.25 mile	0.0009	0.0022	− 0.0197	0.0206	− 0.0164	0.0203
Highway_0.25mile	− 0.0866***	0.0097	− 0.0811	0.1101	− 0.0757	0.1071
Highway_0.5 mile	− 0.0575***	0.0073	− 0.0286	0.0890	− 0.0340	0.0864
Exit_0.25 mile	0.0893***	0.0119	0.0157	0.1375	0.0012	0.1347

Table 3 (continued)

	Model 1 One level standard linear hedonic model (OLS)		Model 2 Random intercept model		Model 3 Random coefficient model	
Exit_0.5 mile	0.1089***	0.0084	0.0881	0.0978	0.1121	0.0952
Exit_0.75 mile	0.0652***	0.0050	0.0430	0.0535	0.0584	0.0529
Exit_1 mile	0.1411***	0.0048	0.1259**	0.0432	0.1256**	0.0427
Exit_1.25 mile	0.0648***	0.0038	0.0955*	0.0377	0.0900*	0.0374
Exit_1.5 mile	0.1351***	0.0040	0.1576***	0.0354	0.1557***	0.0350
Exit_1.75 mile	0.0537***	0.0034	0.0822*	0.0334	0.0891**	0.0329
Exit_2 mile	0.1195***	0.0045	0.1513***	0.0376	0.1576***	0.0370
Log(DisShopping)	− 0.0116***	0.0013	− 0.0236^	0.0123	− 0.0230^	0.0122
Log(DisCBD)	− 0.2226***	0.0042	− 0.1479***	0.0370	− 0.1408***	0.0367
Control structure type	Yes		Yes		Yes	
Random eff						
Level 1 (unit level)						
σ^2(Residual)	–		0.170		0.163	
Level 2 (complex level)						
τ^2(intercept)	–		0.223		0.225	
τ_1^2(Bedrooms)	–		–		0.079	
τ_2^2(EmployeeOccupied)	–		–		0.318	
AIC	5595.3		− 48535		− 53007	
Log likelihood	− 2762.6		24303		26545	
Likelihood Ratio test	–		–		4482.2 (Model 2 is the base)	
Number of Obs	73094		73094 at unit level; 740 at complex level		73094 at unit level; 740 at complex level	
R^2	0.423		Conditional R^2: 0.768		Conditional R^2: 0.846	

logarithm of rent for each unit as dependent variable

***$p=0.001$(two-tailed); **$p=0.01$(two-tailed); *$p=0.05$(two-tailed); ^$p=0.10$(two-tailed) where γ^0 indicates average intercept across complexes; μ_j is a complex- specific effect on the intercept. μ_j is assumed to have a variance of τ^2 and uncorrelated with ε_{ij}. After substituting Eq. (2) into Eq. (1), we have a combined equation:

$$y_{ij} = \gamma^0 + \mu_j + \varepsilon_{ij}, Cov(\varepsilon_{ij}, \mu_j) = 0 \tag{3}$$

while the simple (intercept-only) multilevel model does not include independent variables in the regression, it includes important information regarding how variations in y_{ij} are partitioned between variance among the individual units and variance among the complexes. Regarding the estimation method for MLMs, the maximum likelihood estimation (MLE), the restricted maximum likelihood (REML), or the Bayesian can be used instead of standard OLS, because the random errors for the combined Eq. (3) are neither independent nor have a constant variance. MLE is an iterative methodology in which the algorithm searches for parameter values that will maximize the likelihood of the observed data (Raudenbush and Bryk 2002). REML goes a step further and corrects the estimate of the variance by taking an appropriate degree of freedom into account, a process which has proven more accurate than MLE in estimating variance parameters (Kreft and De Leeuw 1998). In this study, we apply REML; the difference in value for MLE and REML estimates becomes very small as the number of Level 2 clusters increases (Snijders & Bosker, 1999), as is the case in our study of 740 complexes. To fit multilevel linear models in this study, we utilize an R library, nlme (Team, 2013), and apply REML.

The "intraclass correlation coefficient" (Raudenbush and Bryk 2002) is a measure of the proportion of variations in the outcome variable that are attributable to differences at the group level in the MLM literature. The intraclass correlation coefficient (ICC) is computed as follows:

$$ICC = \tau^2/(\tau^2 + \sigma^2) \tag{4}$$

For this case, the ICC is 0.594, which suggests that 59.4% of the variance in rents is due to differences at the complex level. This further justifies the idea that the multilevel linear model approach should be applied in this study.

4.2 Random intercept model

After determining that the multilevel linear model shouldbe used, it is reasonable to include variables for both levels. The random intercept model allows the intercept to vary across complexes, and includes all our explanatory variables,[4] but it makes the coefficients of the explanatory variables fixed. Then, at the first level

$$y_{ij} = \beta_j^0 + \sum_{k=1} \beta^k x_{ij}^k + \varepsilon_{ij}, \varepsilon_{ij} \sim N(0, \sigma^2) \tag{5}$$

where x_{ij}^k is the kth independent variables at the first level; β^k is the coefficient of x_{ij}^k. In this study, we only have two explanatory variables at the unit level: the number of bedrooms and a dummy variable indicating if a unit is occupied by a property company employee.

[4] We tried to include an explanatory variable indicating the distance to open space, and an explanatory variable indicating the distance to an elementary school in the model; however, VIF values of these explanatory variables are larger than 10. This suggests that there may be a multicollinearity problem if they are included in the model. As such, we do not include these two explanatory variables in the models.

This means that k = 2 in this case. Notice that β^k does not vary across complex j. At the second level

$$\beta_j^0 = \gamma^0 + \sum_{l=1} \omega^{0l} z_j^{0l} + \mu_j, \mu_j \sim N(0, \tau^2) \tag{6}$$

where l indicates the number of complex-level independent variables; ω^{0l} is the coefficient indicating the fixed effect of the lth independent variable,z_j^{0l}. After substituting Eq. (6) into Eq. (5), the random intercept model in combined form is:

$$y_{ij} = \gamma^0 + \sum_{k=1} \beta^k x_{ij}^k + \sum_{l=1} \omega^{0l} z_j^{0l} + (\mu_j + \varepsilon_{ij}), Cov(\varepsilon_{ij}, \mu_j) = 0 \tag{7}$$

The random components of the model have two parts:ε_{ij} and μ_j; others are fixed components.

4.3 Random coefficient model

There is no reason why the coefficients of the first-level independent variables must remain constant. For example, there may exist a unique *Bedrooms* effect on rents for complex j in our case. Likewise, there may exist a unique Employee-Occupied effect for complex j. The random coefficient model considers that the coefficients of unit-level independent variables – as well as the intercept – can vary across complexes. At the first level,

$$y_{ij} = \beta_j^0 + \sum_{k=1} \beta_j^k x_{ij}^k + \varepsilon_{ij}, \varepsilon_{ij} \sim N(0, \sigma^2) \tag{8}$$

where β_j^k is the random coefficient of unit-level independent variables,x_{ij}^k. At the second level, the random intercept (β_{0j}) and the random coefficients (β_j^k) can be expressed as the following, respectively:

$$\beta_j^0 = \gamma^0 + \sum_{l=1} \omega^{0l} z_j^{0l} + \mu_j, \mu_j \sim N(0, \tau^2) \tag{9}$$

$$\beta_j^k = \gamma^k + \mu_j^k, \mu_j^k \sim N(0, \tau_k^2) \tag{10}$$

After substituting Eq. (9) and Eq. (10) into Eq. (8), we have the combined form of the random coefficient model:

$$y_{ij} = \gamma^0 + \sum_{k=1} \gamma^k x_{ij}^k + \sum_{l=1} \omega^{0l} z_j^{0l} + \left(\sum_{k=1} \mu_j^k x_{ij}^k + \mu_j + \varepsilon_{ij}), Cov(\varepsilon_{ij}, \mu_j^k \right) = 0 \tag{11}$$

The random components of the model have three parts: $\sum_{k=1} \mu_j^k x_{ij}^k$, ε_{ij} and μ_j; others are fixed components.

5 Results

We start by reporting one-level OLS hedonic regression results (see Column 1 in Table 3). Nearly all the coefficients are statistically significant at 99% and most of them have expected signs. In addition, the magnitude and direction of the estimated coefficients match our intuition at a glance. For example, renovations increase rents significantly; proximity to highway routes decreases rent significantly, while proximity to highway exits increases rent significantly. However, one-level OLS regression could lead to a substantial underestimation of the standard errors when the data structure is hierarchical; as such, the results could be spuriously significant (Krull and MacKinnon 2001). Based on the suspect standard error estimation, the inference would not be reliable.

To account for the hierarchical data structure, we try two primary types of multilevel linear models: the random intercept model and the random coefficient model. The only difference between the two types of models is that the random coefficient model allows the intercept and the coefficients of the unit-level independent variables to vary across complexes, while the random intercept model only allows the intercept to vary across complexes. We run Eq. (7) and report results for the random intercept model in Column 2 of Table 3. We also run Eq. (11) and report results for the random coefficient model in Column 3 of Table 3. With these two candidates' models, the question that arises is, which model fits the dataset better – the random intercept model or the random coefficient model?

The AIC (− 53007) of the random coefficient model is smaller than that of the random intercept model (− 48535), which indicates that the random coefficient model fits the dataset better than the random intercept model. The likelihood ratio statistic (4482.2) justifies the conclusion that the random coefficient model is a better fit than the random intercept model. Additionally, the conditional R^2 of the random coefficient model is 84.6% – larger than that of the random intercept model (76.8%)—which further affirms that the random coefficient model fits the dataset better than the random intercept model.[5]

Next, we report the results of the random coefficient model (Column 3 of Table 3) in detail. The natural logarithm of unit rent is the dependent variable. The multilevel linear model accounts for a respectable 84.6% of the variation in log-rents. The results show that the number of bedrooms positively and significantly impacts unit rent. These findings echo those of previous studies (Babawale et al. 2012, Guntermann and Norrbin 1987, I.Hoch and Waddell 1993, Sirmans et al. 1989, Wilson and Frew 2007). More specifically, the addition of one bedroom increases by 18.42% on average in this study.

Next, we turn to the service and management variables of particular interest in this study. For the unit-level service and management variable a dummy variable indicating that a unit is property employee-occupied is used, the results show that rent is 9.63% less than a unit that is not employee-occupied. As for the effects of complex-level service and management, pet allowance increases multifamily rent by a magnitude of 5.63%, which echoes previous studies (WG Hardin III and Cheng 2003; Sirmans et al. 1989). Permitting pets in

[5] R^2 is frequently used as an indicator for comparing models in terms of model fitness. However, Nakagawa and Schielzeth (2013) suggest that using R^2 from traditional OLS linear model for MLMs yields misleading results and therefore should not be used. There are multiple ideas regarding how to compute R^2 for MLMs For example, pseudo-R^2, marginal R^2, and conditional R2 could be used, but there is no consensus on this. In this study, we also use marginal R^2 and conditional R^2 to compare MLMs in terms of modeling fitting. Marginal R^2 describes the proportion of variance explained by the fixed factor(s) alone, while conditional R^2 describes the proportion of variance explained by both the fixed and random factors (Johnson 2014, Nakagawa and Schielzeth 2013). It is worth noting that marginal R^2 and conditional R^2 are not comparable with R^2 in the traditional OLS context.

a rental property has several benefits for property-owners. This offering attracts more tenants interested in renting property, and landlords have an opportunity to charge more by charging a premium for pets. The results also show that offering short lease options is statistically significant and positively associated with rent. This means that the average rent of a complex that offers short-term leasing contracts is higher than the average rent of a complex that does not provide short-term leasing contracts for tenants. The study also shows that short-term renovations do not increase rents significantly although the renovation can be seen by tenants, which echoes the findings of Mejia and Potter (2015). The observable renovations referenced here include new paint, new carpet, yard landscape, etc. in the prior 12 months. Landlords understand that these renovations are necessary repairs or replacements that make their property nicer than the cheapest generic rental units available to the public. As such, landlords make the renovations but may not increase rents. It is worth noting that we do not have information of how much expenditure or magnitude of renovations were made. Major renovations could have positive effects on rents and further studies are needed. This study also shows that storage services increase multifamily rents significantly, by approximately 8.57%. Regarding availability of disability services, the effect is significant and negative. We are not aware of any previous studies that investigate the effects of employee occupied, short-term leasing contract options and storage service offerings on multifamily rents.

Moving to complex characteristics, multifamily rents significantly decline as the structure ages. This result is consistent with the literature (Allen et al. 1995, J. Frew and Jud 2003, I. Hoch and Waddell 1993, G.D. Jud and Winkler 1991, J.F. Kain and Quigley 1970, Lin and Cheng 2016 Sirmans et al. 1989). in age reduces rent by 0.039%. With regards to the effect of size— as the complex size increases, the rent of the units, on average, increases significantly. Regarding complex-level amenities, having a swimming pool increases rent significantly; this is consistent with the results put forth by previous studies (Guntermann and Norrbin 1987, I. Hoch and Waddell 1993, G.D. Jud and Winkler 1991, Sirmans et al. 1989, B. Wilson and J. Frew, 2007). Regarding the presence of laundry rooms, the estimated coefficient is positive but not statistically significant. Some tenants may favor the common-area laundry room while others do not. The presence of a common-area laundry room implies the lack of in-unit laundry machines, which is a disamenity. The preference dichotomy makes the effect of the laundry room statistically insignificant in this study. The coefficient of parking availability is not statistically significant. In this study, 88.3% units have parking lot or garage. For those that do not have parking lot or garage, they could have no difficulty to park their vehicles along roads, which is significantly different from Eastern Asian cities and Western European cities.

Regarding locational characteristics, units in close proximity to light rail transit stations and shopping centers have higher rents than units located further from these urban amenities. For example, units that are within one mile from their nearest light rail transit stations are expected to see a rent increase of 13.15% than other units. Furthermore, we observe that the closer proximity to light rail stations results in higher rent; however, proximity to bus stops does not increase unit rent significantly. The results show that two coefficients on the dummy variables for the distance bands from highways have negative signs as hypothesized, though none are even close to being significant. Thus, disamenities such as noise—represented by distance from highways—appear to have no significant effect on multifamily rents. Seo et al. (2014) found similar results regarding the effects of proximity to highway routes on property value. Accessibility effects, on the other hand, are very significant at the 0.05 level as complexes are located upwards of 0.75 miles away but fewer than two miles away from highway exits. Units located in higher-performing public elementary school districts have higher rents than units that are not. This finding echoes the results of J.F. Kain and Quigley (1970). The crime rate of a neighborhood is also an important factor in determining rents. The coefficient of crime rate is negative and statistically significant, which means that safer neighborhoods equate to higher unit rent. This result is also consistent with the results of findings by J.F. Kain and Quigley (1970). In addition to the estimates of fixed effects, the estimates of random effects are also listed in Table 3.

In summary, the results show that the effects of service and management – such as availability of short-term leasing options, pet allowance, and storage service offerings – are significant in increasing rents after controlling unit, complex, neighborhood, and locational characteristics. Conversely, employee occupancy and services for those with disabilities decrease rents significantly.

6 Conclusion

Three main findings stand out in this study. First, we show that a multilevel linear model could, and perhaps should, be applied when conducting multifamily housing hedonic analysis. A hierarchical data structure is commonly used in the multifamily context, where hundreds or thousands of units are nested within one complex. Ignoring this type of hierarchical data structure and running only a one-level OLS model could result in spuriously significant effects and, thus, false inferences. Second, we show that, in terms of model fitness, random coefficient models are more appropriate than random intercept models under certain conditions. This is possible because the random coefficient models can account for both spatial autocorrelation and heteroscedasticity of residuals. We suggest that multifamily housing hedonic analysis should take both the random intercept model and the random coefficient model into consideration, and select appropriate one by comparing these two models. We do not suggest that the random coefficient models outperform the random intercept models in other contexts.

Finally, the results show that the effects of service and management factors on multifamily rents vary across types of service and management. Pet allowance, availability of short-term leasing options, and storage service availability increase rents significantly. Conversely, offering units to property employees and services to those with a disability decrease rents significantly.

Service and management variables are determinants of multifamily rents. Investigation into the effects of service and management variables yields important implications for multifamily housing landlords, managers, and affordable housing advocates. Multifamily landlords and managers will need to carefully consider the costs and revenues of providing parking, laundry services, pet options and more for the specific residents they seek to attract. To maximize affordability, however, affordable housing advocates and policy makers may want to minimize ancillary services, disallow pets, and encourage long term leases, especially in high quality school districts with exceptional transit service.

Acknowledgements We gratefully acknowledge helpful comments from Paul Elhorst, Ingmar Prucha, Casey Dawkins, and Andrew McMillan. We also thank Thomas Tippett and Chao Liu for their work in accessing part of the data.

References

Allen, M. T., Springer, T. M., & Waller, N. G. (1995). Implicit pricing across residential rental submarkets. *The Journal of Real Estate Finance and Economics, 11*(2), 137–151.

Appelbaum, R. P., & Glasser, T. (1982). *Concentration in ownership in Isla Vista, California*. Santa Barbara CA: UCSB Housing Office.

Babawale, G., Koleoso, H., & Otegbulu, C. (2012). A hedonic model for apartment rentals in Ikeja area of Lagos metropolis. *Mediterranean Journal of Social Sciences, 3*(3), 109–120.

Benjamin, J., Chinloy, P., & Hardin, W. (2007). Institutional-grade properties: Performance and ownership. *Journal of Real Estate Research, 29*(3), 219–240.

Benjamin, J. D., & Lusht, K. M. (1993). Search costs and apartment rents. *The Journal of Real Estate Finance and Economics, 6*(2), 189–197.

Benjamin, J., Sirmans, G., & Zietz, E. (1997). Security measures and the apartment market. *Journal of Real Estate Research, 14*(3), 347–358.

Brunauer, W., Lang, S., & Umlauf, N. (2013). Modelling house prices using multilevel structured additive regression. *Statistical Modelling, 13*(2), 29.

Chasco, C., & Gallo, J. L. (2013). The impact of objective and subjective measures of air quality and noise on house prices: A multilevel approach for downtown Madrid. *Economic Geography, 89*(2), 22.

Cheah, B. C. (2009). Clustering standard errors or modeling multilevel data. *University of Columbia*, 2-4

Clapp, J. M., & Salavei, K. (2010). Hedonic pricing with redevelopment options: A new approach to estimating depreciation effects. *Journal of Urban Economics, 67*(3), 362–377.

Djurdjevic, D., Eugster, C., & Haase, R. (2008). Estimation of hedonic models using a multilevel approach: An application for the Swiss rental market. *Swiss Journal of Economics and Statistics, 144*(4), 23.

Frew, J., & Jud, G. (2003). Estimating the value of apartment buildings. *Journal of Real Estate Research, 25*(1), 77–86.

Gelfand, A. E., Banerjee, S., Sirmans, C. F., Tu, Y., & Ong, S. E. (2007). Multilevel modeling using spatial processes: Application to the Singapore housing market. *Computational Statistics & Data Analysis, 51*(7), 13.

Giuliano, G., Gordon, P., Pan, Q., & Park, J. (2010). Accessibility and residential land values: Some tests with new measures. *Urban studies, 47*(14), 28.

Glaesener, M. L., & Caruso, G. (2015). Neighborhood green and services diversity effects on land prices: Evidence from a multilevel hedonic analysis in Luxembourg. *Landscape and Urban Planning, 143*, 100–111.

Goodman, A. C., & Thibodeau, T. G. (1998). Housing market segmentation. *Journal of Housing economics, 7*(2), 23.

Guntermann, K. L., & Norrbin, S. (1987). Explaining the variability of apartment rents. *Real Estate Economics, 15*(4), 321–340.

Hardin, W., III, & Cheng, P. (2003). Security, gated communities, and apartment rents. *Journal of Real Estate Research, 25*(2), 145–158.

Hardin, W., III, Hill, M., & Hopper, J. (2009). Ownership structure, property performance, multifamily properties, and REITs. *Journal of Real Estate Research, 31*(3), 285–306.

Hoch, I., & Waddell, P. (1993). Apartment rents: Another challenge to the monocentric model. *Geographical analysis, 25*(1), 20–34.

Johnson, P. C. (2014). Extension of Nakagawa & Schielzeth's R2GLMM to random slopes models. *Methods in ecology and evolution, 5*(9), 3.

Jones, K., & Bullen, N. (1993). A multi-level analysis of the variations in domestic property prices: Southern England, 1980–87. *Urban studies, 30*(8), 18.

Jud, G. D., Benjamin, J. D., & Sirmans, G. S. (1996). What do we know about apartments and their markets? *The Journal of Real Estate Research, 11*(3), 243–257.

Jud, G. D., & Winkler, D. T. (1991). Location and amenities in determining apartment rents: An integer programming approach. *Appraisal Journal, 59*(2), 266–275.

Kain, J. F., & Quigley, J. M. (1970). Evaluating the quality of the residential environment. *Environment and Planning A, 2*(1), 23–32.

Kreft, I. G., & De Leeuw, J. (1998). *Introducing multilevel modeling*. Sage, California

Krull, J. L., & MacKinnon, D. P. (2001). Multilevel modeling of individual and group level mediated effects. *Multivariate behavioral research, 36*(2), 29.

Lancaster, K. (1966). Change and innovation in the technology of consumption. *The American Economic Review, 56*(1/2), 14–23.

Lin, J.-J., & Cheng, Y.-C. (2016). Access to jobs and apartment rents. *Journal of Transport Geography, 55*, 121–128.

Mejia, L. C., & Potter, K. D. (2015). The value behind the value-add: Multifamily rent growth after renovations. *Journal of Housing Research, 24*(2), 211–220.

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in ecology and evolution, 4*(2), 10.

Orford, S. (2002). Valuing locational externalities: A GIS and multilevel modelling approach. *Environment and Planning B: Planning and Design, 29*(1), 13.

Quigley, J. M. (1985). Consumer choice of dwelling, neighborhood and public services. *Regional science and urban economics, 15*(1), 23.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods (Vol 1)*. California: Sage.

Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of political economy, 82*(1), 34–55.

Seo, K., Golub, A., & Kuby, M. (2014). Combined impacts of highways and light rail transit on residential property values: a spatial hedonic price modelfor Phoenix, Arizona. *Journal of Transport Geography, 41*, 53–62.

Shin, W. J., Saginor, J., & Van Zandt, S. (2011). Evaluating subdivision characteristics on single-family housing value using hierarchical linear modeling. *Journal of Real Estate Research, 33*(3), 32.

Sirmans, S., & Sirmans, C. (1992). Property manager designations and apartment rent. *Journal of Real Estate Research, 7*(1), 91–98.

Sirmans, S., Sirmans, C., & Benjamin, J. (1989). Determining apartment rent: the value of amenities, services and external factors. *Journal of Real Estate Research, 4*(2), 33–43.

Snijders, T. A., & Bosker, R. J. (1999). *An introduction to basic and advanced multilevel modeling*. California: Sage.

Team, R. D. C. (2013). nlme: Linear and nonlinear mixed effects models (Version R package version).

Treg, C. (2010). A multilevel property hedonic approach to valuing parks and open space (Doctoral dissertation, University of Vermont, Burlington, USA). Retrieved from https://scholarworks.uvm.edu/cgi/viewcontent.cgi?article=1229&context=graddis

Uyar, B., & Brown, K. (2007). Neighborhood affluence, school-achievement scores, and housing prices: Cross-classified hierarchies and HLM. *Journal of Housing Research, 16*(2), 20.

Wilson, B., & Frew, J. (2007). Apartment rents and locations in Portland, Oregon: 1992–2002. *Journal of Real Estate Research, 29*(2), 201–218.

Zietz, E. (2003). Multifamily housing: A review of theory and evidence. *Journal of Real Estate Research, 25*(2), 185–244.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.