Journal of Housing and the Built Environment (2019) 34:91–109
https://doi.org/10.1007/s10901-018-9618-1

ARTICLE

The effect of young, single person households on apartment prices: an instrument variable approach

Tanja Tyvimaa[1] · **Md. Kamruzzaman**[2]

Received: 15 September 2017 / Accepted: 7 June 2018 / Published online: 20 June 2018

Abstract
There is an increasing evidence in the literature that the average household size is declining in many western countries with people living alone more often and maintaining a single or solo household. This trend can affect the housing market in many ways including an increase in demand for housing and property prices. However, little knowledge exists on what effects singles can have on property values. This study, using a two stage least square (2SLS) regression model, estimates the impact of young (aged 34 years or less), single person households on apartment prices in Helsinki, Finland. Results show that a 1% increase of young, single person household in a neighbourhood leads to a 0.51% increase of apartment prices. This means, for example, that if young, single person household increases by 10% in an area, the average value (€196,348.10) of an apartment increases by €10,061. The finding implies, *ceteris paribus*, that the apartment market will continue to experience an upsurge if the current trend of decline in household sizes continue.

Keywords Single person household · Apartment prices · Simultaneity bias · Instrument variable · Helsinki

1 Introduction

The number of single person households has been gradually increasing since the last few decades in many developed countries. For example, the average household sizes in Australia declined from 3.4 persons in 1966 to 2.6 persons in 2006 (Australian Bureau of Statistics 2012). The result of which is that many countries are now experiencing a higher share of single person household. As an example, 42% of all households were single person households in 2014 in Finland. A similar proportion is also reported in other countries within the

✉ Md. Kamruzzaman
md.kamruzzaman@monash.edu

Tanja Tyvimaa
t.tyvimaa@qut.edu.au

1 School of Civil Engineering and Built Environment, Queensland University of Technology (QUT), 2 George Street, Brisbane, QLD 4001, Australia

2 Monash Art Design & Architecture (MADA), Monash University, 900 Dandenong Road, Caulfield East, VIC 3145, Australia

European Union such as in Germany and in Denmark (Eurostat 2016). The increase of singles is not a trend only in Europe but also in the US, where the share of single person households increased from 6 to 28% between 1930 and 2010, and the share of married-couple households decreased from 79 to 49% over the same period (Deka 2014). Presently, the average household size in the US is similar (2.58 persons) to that found in Australia, but slightly larger than Europe (2.4) (US Census Bureau 2012). In Australia, about one quarter of households are reported to be single person households with an increasing trend from 8% in 1946 to 24% in 2011 (de Vaus and Qu 2015). Some Asian countries also corroborate with this trend. For example, one-third of high-educated females are single in Singapore and nearly 40% of Taiwanese women over 30 years of age living alone (Turkki 2013). All these increasing trends for singles and decreasing trend in household sizes likely to affect the housing and property market in several ways. For example, the demand for small apartment units may increase and so do their prices at the expense of demand for large apartments and houses.

The nature of challenges to own a property is different for singles from other types of households (e.g. couples). Singles have a relatively lower purchasing power given that they are the lone earner in their households (Quintano and D'Agostino 2006). However, their expense levels are comparable to those couples because singles enjoy the same basic living condition to that larger sized households, such as satellite television, internet, and other household appliances (washing machine or coffee maker). As a result, singles often faced with a reduced level of disposable income which limits their choice set to smaller apartments when it comes to property ownership. Consequently, the demand for smaller apartments likely to go up with increasing trend of single person household. Although the number of singles has increased over the years, little knowledge exists on their property ownership behaviour and the consequent demand in housing market. Kohler and van der Merwe (2015) have highlighted that a decline in household size could be one of the factors that can potentially affect housing price growth in the long-run. Along with this argument, this current study hypothesises that an increasing number of young, single person households is likely to increase apartment prices. This study aims to test this hypothesis using a large property transaction database gathered from Helsinki, Finland.

The rest of the paper is organised in four sections. The following section defines and characterises singles based on the literature. The data and methods used in this research are discussed in Sect. 3. The results are presented in Sect. 4. Finally, Sect. 5 discusses the findings in policy terms and concludes this research together with direction for future research on this topic.

2 Singles: definitions and residential preferences

Singles have been defined in many ways in the literature. DePaulo and Morris (2006) have classified singles into legally single and socially single. Legally singles are adults who are divorced, widowed or never married. Socially single is a term used in everyday life to describe individuals' life stages. Wulff (2001) has used the United Nation's definition that a single is a person who makes provision for his or her own food and other essentials for living without combining with any other person and lives in an independent dwelling. In this study, a single means a household of one adult living financially independently.

Singles have many common features across the world and residential location has been identified as an important one to describe singles. Hall et al. (1997) have identified the patterns of geographical distribution and characteristics of single person households in England, Wales and France over the period of 1982–1990. They found that the working age singles were more attracted to the cities whereas older single person households were concentrated in rural areas. Areas close to the central business districts (CBD) contained the highest proportion of single person households in Australia, as the 25–34 age cohort seeks the lifestyle, amenities, and access to jobs in inner city areas (Reynolds et al. 2004). Similar findings have been echoed in another Australian study in case of apartment choices (Wulff 2001). An American study by Scott and Horner (2008) found that single person households in Louisville, KY, lived with greater accessibility to commercial centres and services such as banks, post offices and cleaners. A similar pattern was also reported for Finnish singles who preferred inner city living (Kauppinen et al. 2014).

Rented homes are often a preferred option for singles. Deka (2014) has reported that 54% of the single men and 45% of the single women in the US lived in rented dwellings. Similar findings have been reported for Dutch single person households with preferences for rented dwellings and multi-storey buildings (Faessen 2002). Nevertheless, studies in London and Paris found that singles were moving away from rented homes and owner occupation was becoming more common over the period of 1981–1991 (Hall and Ogden 2003; Ogden and Schnoebelen 2005).

Single person households exhibit a higher level of residential mobility than other households. Bogue et al. (2009) have reported that American singles were twice as mobile as those who were married and living with their spouse. Also, divorced, and separated persons (married but the spouse was absent) had higher mobility rate than married couples. A similar pattern have also been reported in the Netherlands, where single person households were 22–31% more willing to move than other households (Faessen 2002).

Higher education and social class are also found to be common features of singles. Hall and Ogden (2003) have demonstrated that the number of singles under the age of 60 was increasing in the inner city area of London, and many of them were in a higher social class. A similar trend was found by Ogden and Schnoebelen (2005) in Paris. Klinenberg (2012) found that higher education and a middle class income are positively associated with American singles. This does not necessarily mean that single person households should be treated as a homogeneous group, rather the characteristics significantly vary depending on age, gender, social class and geography (Ogden and Schnoebelen 2005). For example, a study conducted in England and Wales found that female singles were better educated, had higher professional status and financial condition than male singles (Chandler et al. 2004).

All the above findings suggest that the housing demand for singles should not be assessed in isolation rather any modelling effort to understand the impact of singles on apartment prices must take into account the various characteristics of singles including age, gender, home ownership pattern, income, and locational factors (Wulff 2001). Although the above studies have examined single person households, they have focused on the socio-economic characteristics of singles and their housing conditions, showing similar choice patterns all over the world. Surprisingly, very little or no research has focused on the apartment prices and how the increasing number of singles affects apartment prices.

3 Methodology

3.1 Study area

To test the hypothesis, this research focuses on apartment purchase behaviour of young singles in the context of Helsinki, Finland. Finland possesses an interesting case for this research because of its highest share of single person household in Europe (42%) with an average household size of 2.09 persons (Official Statistics of Finland 2016). It is also forecasted that the household size of Finland will continue to shrink to 1.92 persons in 2025 (Juntto 2008). Within Finland, Helsinki was chosen for three reasons.

First, the average household size in Helsinki is even smaller (1.95 persons) than that of in Finland with a total of 620,715 population in 2014 (Official Statistics of Finland 2016). In Helsinki, the proportion of single person households increased from 40% in 1985 to 48% in 2014. The largest proportion of single person households belongs to two age groups (25–34 years and 75+ years), each group consists of about 64% of the total population in these respective cohorts. The smallest proportion of single person households (36%) is in the 35–44 age group. Households living in multi-storey buildings are dominated by singles, 53% being the single person households. Finnish welfare policy guarantees equal opportunity to everyone to access public services, such as health care and schools, as well as access to public transportation.

Second, recent apartment price statistics from Finland indicates that the resale prices of smaller sized apartments (studio and one bedroom) have increased more than the prices of larger apartment between 2011 and 2015. This is particularly the case in the city of Helsinki with an increase of up to 20% (Statistics Finland 2017). In addition, smaller sized apartments were sold in a faster rate than the larger units reflecting their demand in the property market. Recent studies also show that apartments located in central Helsinki became a preferred option for singles than those larger apartments located further away from city centre (Kauppalehti 2016).

Third, the restrictive urban planning, housing policy, and construction regulations have led to a homogenous supply of housing in Helsinki (Puustinen and Kangasoja 2009). These unique qualities make the city of Helsinki to be an ideal case study to investigate the impacts of singles on apartment prices.

3.2 Data

Research has identified that property prices are determined based on a range of factors including structural features of the properties such as size and number of bedrooms (Pashardes and Savva 2009), lot area (Chiang et al. 2015), building age (Goodman and Thibodeau 1995); the quality of neighbourhoods such as income (Fingleton 2006), education (Gibbons 2003); environmental conditions such as crime (Ceccato et al. 2011) and noise (Marmolejo Duarte and González Tamez 2009); and locational attributes such as access to public transport services (Brandt and Maennig 2012; Efthymiou and Antoniou 2014; Mulley 2013), distance to CBD, parks and open spaces, shopping centres, and water features (Daams et al. 2016; Kam et al. 2016; McCord et al. 2012; Parent and vom Hofe 2013). As a result, this study collated/derived three types of datasets to investigate the impacts of single households on property prices: (a) property transaction data together with the attributes of sold properties—property characteristics; (b) socio-economic profile of people living in an area where a property transaction was made—neighbourhood characteristics; and (c) spatial dataset to derive locational characteristics of the properties sold—geographic characteristics.

3.2.1 Property transaction data

The property transaction data contained residential sale information of all properties sold in the city of Helsinki from January 2005 to September 2012. The data were obtained via KVKL Hintaseurantapalvelu online service. This online database is managed by The Central Federation of Finnish Real Estate Agencies. The Federation represents approximately 80% of the professional real estate agencies in Finland and they upload their transaction data to this online database in every month (www.kvkl.fi/kvkl-hintaseurantapalvelu.html). This online database is similar to the Multiple Listing Service (MLS) database in the USA (www.mls.com) or CoreLogic RP Data service in Australia (www.corelogic.com.au). The database contains detailed information of each sold property including sales prices, listing and transaction dates, size and number of rooms, position (level) of the property if in a multi-story building, condition of the property, geographic reference (e.g. X-coordinate and Y-coordinate), and legal characteristics.

The original dataset contained transaction details of 71,508 new and existing housing units. This research used only resales of existing units located in multi-storey buildings. Transactions which have all the variables of interest were kept and several steps were taken to clean the data: (a) exclusion of new sales resulted in 64,260 records; (b) resales of apartments located in a multistorey building were only kept (56,730 records); (c) apartments with more than six rooms were removed which resulted in 56,645 records; (d) buildings with floors less than or equal to zero (i.e. all floors are under the ground) were deleted which returned 56,286 records; (e) buildings with less than two floors were removed which resulted in 56,168 records; (g) unspecified characteristics of lots were deleted which resulted in 55,004 records; (h) unspecified condition of properties were removed which returned 51,243 records; (i) missing socio-demographic information of census grids (e.g. number of singles living) within which a property is located was excluded which returned 51,174 records; (j) if the sales frequency was found to be less than 10 per census grid (discussed in Sect. 3.1.2), these properties/grids were not included in the analysis for the sake of statistical modelling confidence which resulted in 50,758 transactions; and (k) if a census grid is located in isolation in geographic terms (i.e. there is no geographic neighbour either due to the presence of small islands in Helsinki or the removal of census grids in the previous step), these properties were also excluded which resulted in 50,253 transactions as the analytical sample for this research. The database, however, does not contain any distinction of the sold properties between investment sales and owner-occupation sales. The location of the sold properties is shown in Fig. 1. It shows that the transactions are spatially distributed across Helsinki with a higher concentration in the CBD area.

3.2.2 Socio-economic profile

The socio-economic profile data of people living in an area, where a property transaction was made, were obtained from the census data as provided by the Statistics Finland. The database is referred to as "The Grid Database 2012" and is geographically defined by a 250 × 250 m geocoded grid. Therefore, the grids do not follow any natural/administrative boundaries, properties or buildings. The size of one grid includes approximately two blocks of buildings. Figure 2 shows the location of these census grids in Helsinki together with the grid locations where the property transactions were made.

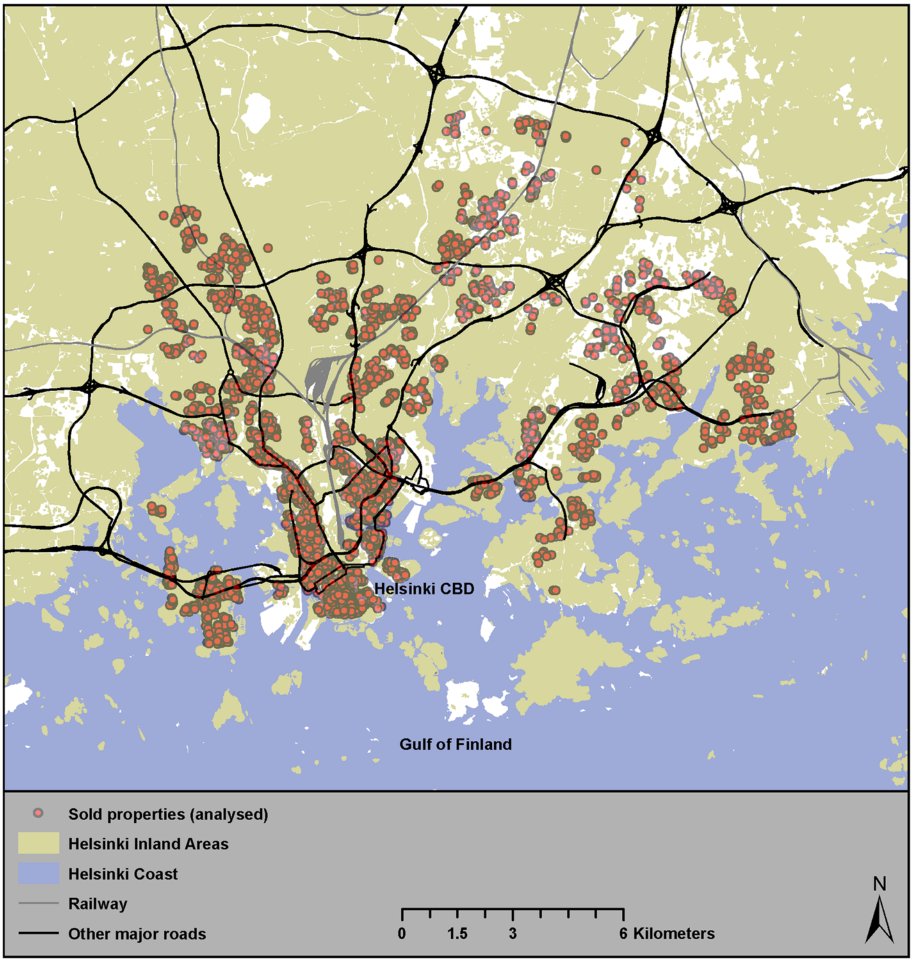

Fig. 1 Location of the resale apartments in Helsinki

The database contains a range of socio-demographic information of people living within the grids including average age, gender, average household income, household size and household structure (including proportion of young single households), employment status, and education level. The dataset also contains some urban form variables including dwelling density, proportion of commercial and residential buildings, average occupancy of the buildings etc.

In order to test the hypothesis of this study, the information contained within The Grid Database was needed to associate with the property database. Note, however, that the property database does not contain any information that can be used to associate with the census grid information. Nevertheless, both databases contained geographic reference attributes. As a result, the Spatial Join operation was performed in ArcGIS 10.3 to link the attributes from the two databases together. Later on, a summary statistics was calculated to examine how many apartments were sold within each grid. As

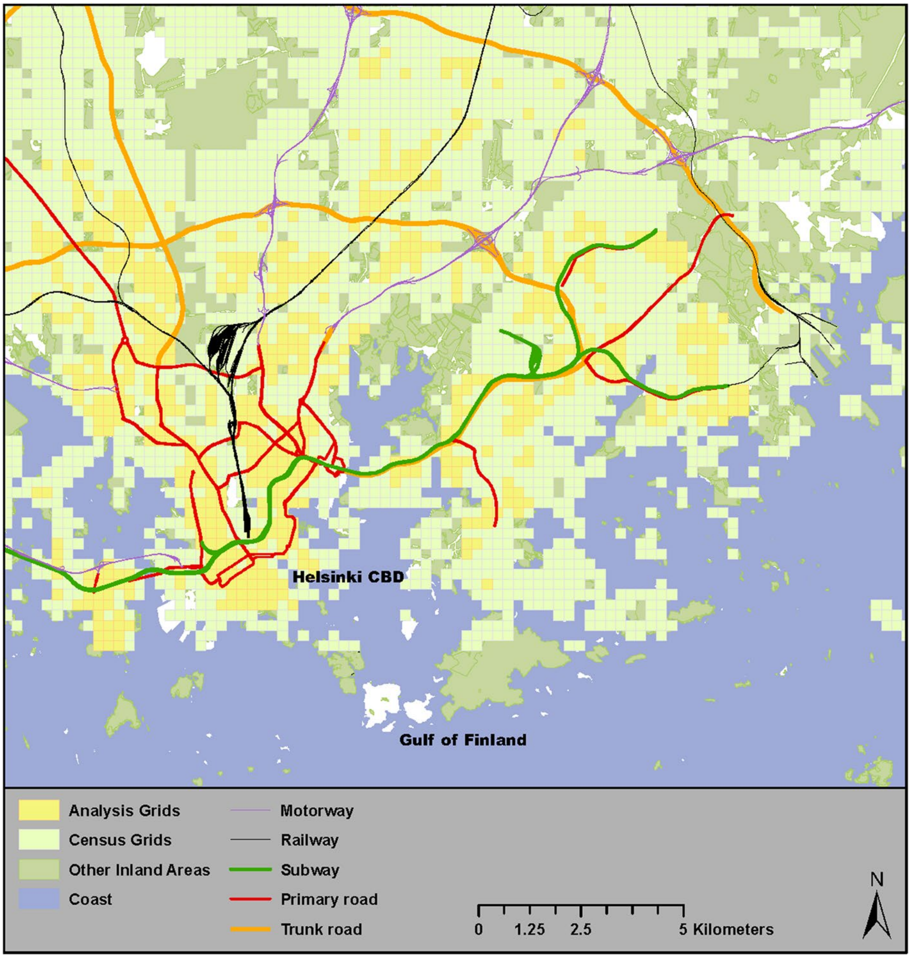

Fig. 2 Census grids in Helsinki and the location of the property transaction grids

indicated previously, if the sales frequency was found to be less than 10 per grid, these properties/grids were not included.

3.2.3 Spatial dataset to derive locational characteristics

The objective of this paper is to assess the impacts of singles on apartment prices based on an estimate of hedonic-type regression model that incorporates distances from the sold properties to different opportunities and services. As a result, this research conducted spatial analyses using the ArcGIS software to derive distances from each sold properties to their closest facilities. In particular, five distances were used in this research: distance to CBD, distance to motorways, distance to major public transport stops, distance to coast, and distance to hospital. Prior research has shown that these spatial attributes significantly affect property prices in Helsinki (www.maanmittauslaitos.fi/en/kartat; www.helsinki.

Table 1 Summary statistics

Variables	Mean (%)	Std. Dev.	Min	Max
Property characteristics				
Property prices (Euro) (2012 real prices)	196,348.10	136,708.80	13,581.00	3,630,000.00
Natural log of property prices	12.04	0.50	9.52	15.10
Floor on which the unit is located	3.06	1.75	1.00	25.00
Number of floors in building	5.02	1.96	2.00	26.00
Number of rooms in unit	2.15	0.99	1.00	6.00
Size of the unit (square metre)	55.63	26.62	8.50	1555.00
Presence of sauna				
No (%)	89.14	–	–	–
Yes (%)	10.86	–	–	–
Property condition				
Good (%)	52.23	–	–	–
Average (%)	41.01	–	–	–
Poor (%)	6.66	–	–	–
Lot				
Rental (%)	26.05	–	–	–
Owned (%)	73.95	–	–	–
Census block characteristics within which the property is located				
% of young single persons' hh (– 34 years)	19.21	9.21	0.00	50.41
% of young couple's hh without children	7.47	3.20	0.00	31.05
% of females	53.82	3.17	34.83	68.18
% of people educated	77.08	9.53	35.83	100.00
% employed	92.84	3.58	67.90	100.00
% of pensioner's hh	22.25	8.83	0.76	82.15
Average age of residents (year)	40.91	4.22	24.00	71.00
Average income of hh (Euro)	48,219.49	19,830.47	17,061.00	377,647.00
% of buildings designated as non-residential	19.60	16.23	0.00	88.89
Dwelling density (number of dwelling/grid)	572.81	432.84	18.00	2257.00
Occupancy rate of dwelling (square metre/person)	33.92	4.53	23.00	80.90
% of owner occupied dwelling	53.72	16.23	0.00	100.00
Geographic characteristics of the property				
Distance to CBD (m)	5188.79	3491.65	208.95	13,374.61
Distance to motorway (m)	2145.73	1314.63	20.34	8492.42
Distance to major public transport (train, metro, tram) (m)	461.83	510.93	1.00	4859.20
Distance to coast (m)	1265.25	1273.42	6.86	6798.85
Distance to hospital (m)	1478.73	1128.42	1.74	6323.00
Year dummy				
2005 (%)	12.31	–	–	–
2006 (%)	12.72	–	–	–
2007 (%)	13.63	–	–	–
2008 (%)	11.19	–	–	–
2009 (%)	12.39	–	–	–
2010 (%)	12.60	–	–	–
2011 (%)	13.47	–	–	–

Table 1 (continued)

Variables	Mean (%)	Std. Dev.	Min	Max
2012 (%)	11.69	–	–	–
Instrument				
Spatially lagged % of young single persons' hh (– 34 years)	15.07	7.26	0.00	33.91
N	50,253			

fi). In this research, the required spatial datasets representing each of these features were downloaded from the website of OpenStreetMap (www.openstreetmap.org).

Table 1 shows a summary of all the variables derived and used in this research as discussed above. Sales were relatively evenly distributed over the 8-year study periods (about 12% in each year). 52% of the sold units had a good condition at the time of sale. Sales prices ranged from 13,581 to 3,630,000 Euros, with a mean of 196,348 Euros. Two room (one bedroom) units were the most common sales. The size of the units ranged from 8.5 to 1555 square meters, with an average of 55.6 square meters. A private sauna, a preferred amenity in the Finnish housing market, was found to exist in 11% of the units sold. Young, single person households represented 19% of all population in the studied Grid blocks. The average age of citizens is 41 years and slightly more than half (54%) are females.

3.3 Analytical strategy

The study examines how an increasing number of young, single person households affect apartment prices based on an estimate of two types of regression analyses: ordinary least square (OLS) regression, and instrument variable approach within a two stage least square (2SLS) regression in order to encounter possible endogeneity bias from reverse causality. In both models, the final sale prices (Euros in 2012 real prices) of apartments were used as outcome variables. An initial investigation of the distribution of this variable shows that the variable is highly skewed to the right which violates the normality assumption of regression models. As a result, a natural log transformation was made of this variable which showed that the transformed variable is approximately normally distributed. Therefore, the transformed variable was used as outcome variables for the regression analyses.

In the OLS model, the outcome variable was regressed based on percentage of young (– 34 years of age) single person households located within the grid. The distribution of this focused, impact variable was found to be approximately normally distributed and thus meets the regression assumption. This type of hedonic model is a well-known method for the analysis of property values, the model treats property values as a function of a bundle of various characteristics (Armstrong and Rodríguez 2006; Cebula 2009; Cho et al. 2006; Huang and Yin 2015; Hui et al. 2007). The characteristics included to our model are structural characteristics of a building and unit, neighbourhood characteristics, and geographic characteristics as outlined previously. The value of an apartment is the sum of the implicit prices for the characteristics included to the hedonic model. Also time of the sale is controlled by the year dummy variables in order to account for time-dependent factor that may have influenced property prices in a particular year (e.g. interest rates). The final model, however, included only the statistically significant factors ($p < 0.1$) upon refinement of an initial model that included all the explanatory factors. Operationally, this means estimating a basic model as outlined below:

$$\ln(P_i) = \propto_0 + \propto_1 Singles_{ig} + \propto_2 Structure_i + \propto_3 Neighbourhood_{ig} + \propto_4 Geography_i + \propto_5 Year_i + \varepsilon_0 \quad (1)$$

where, P_i refers to the sale price of an apartment i; $Singles_{ig}$ represents the proportion of young single household within census grid g in which apartment i is located; $Structure_i$ represents the structural characteristics of an apartment i (age, floor, size); $Neighbourhood_{ig}$ represents the characteristics of neighbourhood (census grids) g where the sold apartment i is located (education, income);$Geography_i$ represents the geographical characteristics of the apartment i (distance to CBD); $Year_i$ represents the year of sale of an apartment i; and ε_0 denotes the residual error.

The problem with the OLS estimation using apartment prices as an outcome variable and percentage of young, single household as an independent variable is that they are hypothetically linked in two-way. For example, more singles can increase the prices of smaller apartments because of their higher demand. On other hand, singles will be attracted to those areas where the apartment prices are cheaper given their limited disposable income base as discussed earlier. This means that a bi-directional causal relationship could emerge between the apartment prices and the percentage of singles living in an area (simultaneity bias). In other words, both apartment prices and singles in Eq. 1 influence each other at the same time, and as a result, $Singles_{ig}$ in Eq. 1 might not be a truly exogenous variable but possesses the quality of an endogenous variable. If $Singles_{ig}$ is not an exogenous variable, it will be related to the error term which violates the basic assumption upon which OLS is based (Webster 2013).

Research has highlighted that an instrumental variables (IV) technique is more appropriate to address the endogeneity problem (Agnello and Schuknecht 2011; Benefield et al. 2014). The basic procedure underlying this technique involves two-steps. First, this research estimated an auxiliary instrumenting equation (Eq. 2) where percentage of young, single household is regressed on all the exogenous variables and the candidate set of instruments, in this case spatially lagged values of the young, single households. Second, the research estimated an OLS model where the predicted values of the percentage of singles obtained from the first-stage were used (instead of the original values) to regress apartment prices (Eq. 3). Operationally, these can be formulated as:

$$\begin{aligned}\widehat{Singles}_{ig} = \beta_0 + \beta_1 SpatillyLaggedSingles_{ig} + \beta_2 Structure_i \\ + \beta_3 Neighbourhood_{ig} + \beta_4 Geography_i + \beta_5 Year_i + \varepsilon_1\end{aligned} \quad (2)$$

$$\ln(P_i) = \gamma_0 + \gamma_1 \widehat{Singles}_i + \gamma_2 Structure_i + \gamma_3 Neighbourhood_i + \gamma_4 Geography_i + \gamma_5 Year_i + \varepsilon_2 \quad (3)$$

where, $\widehat{Singles}_{ig}$ is the predicted proportion of young singles within census grid g of the sold apartment i; and $SpatiallyLaggedSingles_{ig}$ represents the average of the percent of young singles in the neighbouring census grids of grid g where the apartment i is located and was used as an instrument variable.

An IV estimation is based on choosing a variable which is exogenous and therefore unrelated to the error term, but which at the same time may serve as a proxy for singles. It must also be an exogenous variable that does not appear in Eq. 1 (Mukherjee et al. 2013). Therefore, the two main requirements for using IV are:

1. The instrument $SpatiallyLaggedSingles_{ig}$ must be correlated with the endogenous explanatory variable $Singles_{ig}$, conditional on the other covariates. If this correlation is strong, then the instrument is said to have a strong first stage. A weak correlation may provide misleading inferences about parameter estimates and standard errors.
2. The instrument $SpatiallyLaggedSingles_{ig}$ cannot be correlated with the error term ε_0 in the explanatory equation, conditional on the other covariates. In other words, the instrument cannot suffer from the same problem as the original predicting variable $Singles_{ig}$. If this condition is met, then the instrument is said to satisfy the exclusion restriction.

Conceptually, the selection of the spatially lagged values of the young, single households as an instrument is justified based on the first law of geography: "everything is related to everything else, but near things are more related than distant things" (Tobler 1970, p. 236). This means that singles are likely to be higher in a grid if it is surrounded by grids (neighbours) with a higher proportion of singles.[1] Our investigation shows that a strong correlation (0.78) exists between $Singles_{ig}$ and $SpatiallyLaggedSingles_{ig}$. In contrast, it is unlikely that apartment price in a grid affects the number of singles to be lived in surrounding grids. A number of studies used spatially lagged values of independent variables to account for the possible endogeneity bias in the literature (see for example, Irwin et al. 2014; Kajuth et al. 2016; Kim and Goldsmith 2009). Nevertheless, the research tested the validity of instrument variable approach and the quality of the instrument $SpatiallyLaggedSingles_{ig}$ used in this research.

The 2SLS method, therefore, isolates the impact of singles on apartment values and thus the true effect of singles on apartment values can be established (see for example, Habibov and Afandi 2016; Lokshin and Ravallion 2008). The results of the 2SLS estimation are interpreted in the same way like the OLS coefficients.

The Part and Partial correlations and the Multicollinearity analysis (variance inflation factor—VIF) results using all the explanatory variables entered in the regression models show that three variables had a high level of VIF (% of pensioner's household, % of people educated, and distance to CBD). These variables were gradually removed from the equation by testing multicollinearity after each exclusion. Importantly, the inclusion of the distance to CBD variable changes the sign of the key exposure variable (% of young singles' household) due to a significantly higher level of negative correlation (− 0.60). None of the explanatory variables included in the final model had a multicollinearity problem. The highest VIF value was found to be 4.25 which is lower than the rule-of-thumb value of 5 (Hair et al. 1995; Rogerson 2001). This is again confirmed in the collinearity diagnostics analysis that there was no serious problems with multicollinearity.

All models were estimated using Stata software (version 13). Given that the sold properties are clustered within each grid, the *vce*(*cluster clustvar*) option was applied in order to account for the clustering effect within each CCD. This technique allows to obtain a robust variance estimate after adjusting for within-cluster correlation—i.e. a high value property in a grid is likely to increase the value of neighbouring properties (Kamruzzaman et al. 2015). The Grid code was used as the clustering variable in the models.

[1] The neighbouring grids were given a weight of 1 if they are located either in the edge or corner of a grid, otherwise zero (0). The weights were row standardised in the spatial weight matrix so that weights across rows sum to one.

Table 2 Regression analyses results (std. error adjusted for 760 clusters in census grid)

Explanatory variable	Outcome variable: sale price [Euros in 2012 real prices (natural log transformed)]					
	OLS estimate			2SLS estimate		
	Coef.	*t*	*p* > *t*	Coef.	*z*	*p* > *z*
Property characteristics						
Floor on which unit located	0.014702	14.340	0.001	0.014535	14.060	0.000
Number of floors in building	−0.004689	−2.220	0.027	−0.004399	−1.920	0.055
Number of rooms in unit	0.089189	1.930	0.053	0.090731	1.960	0.050
Size of the unit (square metre)	0.009678	4.630	0.001	0.009721	4.630	0.000
Sauna: yes (ref: no)	0.141414	10.440	0.001	0.151048	10.250	0.000
Condition (ref: good)						
Average	−0.109612	−35.450	0.001	−0.110269	−34.250	0.000
Poor	−0.188005	−36.760	0.001	−0.188830	−37.500	0.000
Building located on owner occupied lot (ref: leased)	0.080162	8.890	0.001	0.069649	6.670	0.000
Census block characteristics within which the property is located						
% young single persons' hh (−34 years)	0.002242	2.570	0.010	0.005111	2.550	0.011
% of young couple's hh without children	0.008509	4.730	0.001	0.008841	4.410	0.000
% of females	0.004382	2.680	0.008	0.003440	1.820	0.069
% employed	0.013367	6.010	0.001	0.013050	5.250	0.000
Average age of residents	−0.003753	−1.760	0.080	–	–	–
Average income of resident	0.000003	2.500	0.013	0.000004	2.890	0.004
% of buildings designated as non-residential	0.001348	4.540	0.001	0.001144	3.730	
Dwelling density (number/grid)	0.000092	3.380	0.001	0.000071	2.390	
Occupancy rate of dwelling (m^2/person)	0.013910	2.980	0.003	0.013741	3.080	0.002
% of owner occupied dwelling	−0.001860	−4.980	0.001	−0.002086	−5.470	0.000
Geographic characteristics of the property						
Distance to motorway	−0.000020	−5.390	0.001	−0.000020	−5.380	0.000
Distance to major public transport (metro, tram) (m)	−0.000073	−7.030	0.001	−0.000067	−6.240	0.000
Distance to coast (m)	−0.000061	−13.380	0.001	−0.000059	−12.100	0.000
Distance to hospital (m)	−0.000019	−4.390	0.001	−0.000016	−3.390	0.001
Year dummy (ref: 2005)						
2006	0.091864	16.210	0.001	0.092746	16.440	0.000
2007	0.161198	27.910	0.001	0.161848	28.040	0.000
2008	0.171611	28.910	0.001	0.171645	29.000	0.000
2009	0.199333	31.140	0.001	0.199505	31.220	0.000
2010	0.303744	45.250	0.001	0.303904	45.520	0.000
2011	0.341268	49.060	0.001	0.341690	49.060	0.000
2012	0.381518	54.920	0.001	0.381870	54.680	0.000
Constant	9.202624	51.870	0.001	9.073554	43.990	0.000
N	50,253			50,253		
Adjusted R^2	0.8690			0.8678		
F	651.88		0.001			
Wald chi^2				19,327.57		0.001

– Coefficient is not statistically significant at the 0.1 level

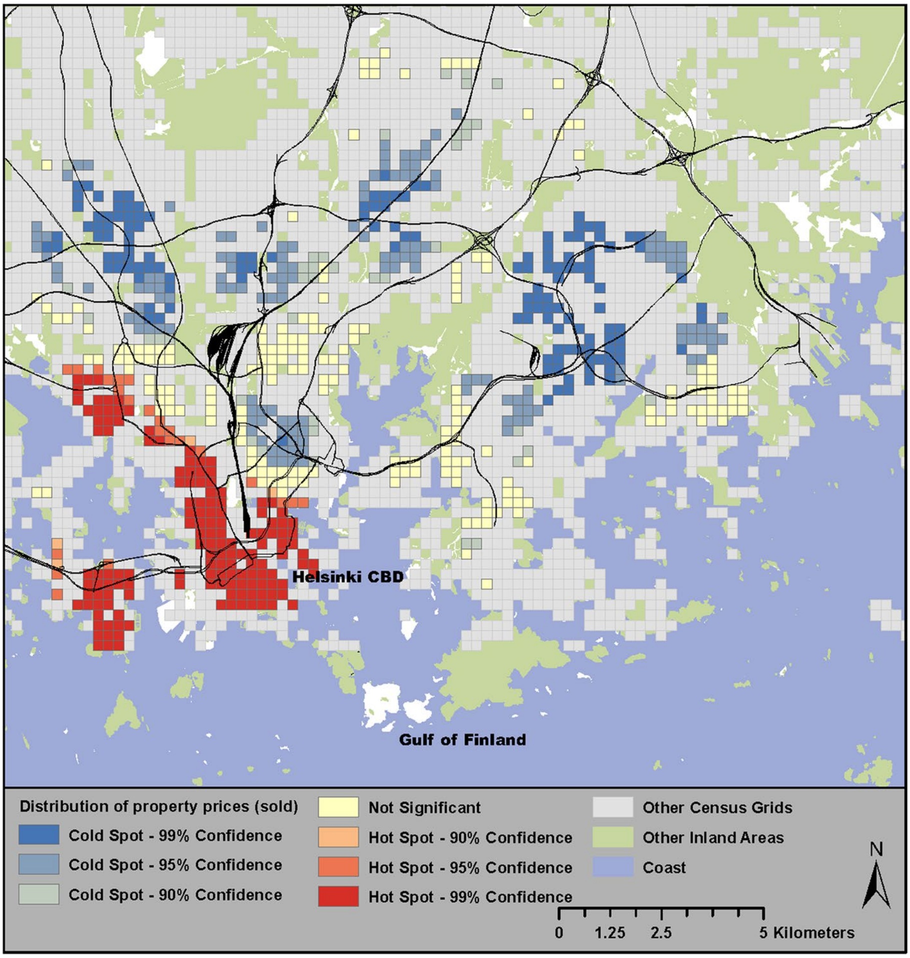

Fig. 3 Hot spot analysis result showing the spatial distribution of apartment prices in Helsinki

4 Results

4.1 Results of from the OLS model estimates

The first three outcome columns in Table 2 show the results obtained from the OLS model. Overall, the estimated model was found to be statistically significant with a very good explanatory power ($R^2 = 87\%$). The model shows that the proportion of young, single person households significantly affected apartment prices. However, the positive sign associated with the coefficient signifies that the relationship between singles and apartment prices is positive. This means that an increasing number of singles increases apartment prices. The coefficient value of 0.002242 suggests that one percent increase in singles household in a grid increases apartment prices by 0.22% [i.e.

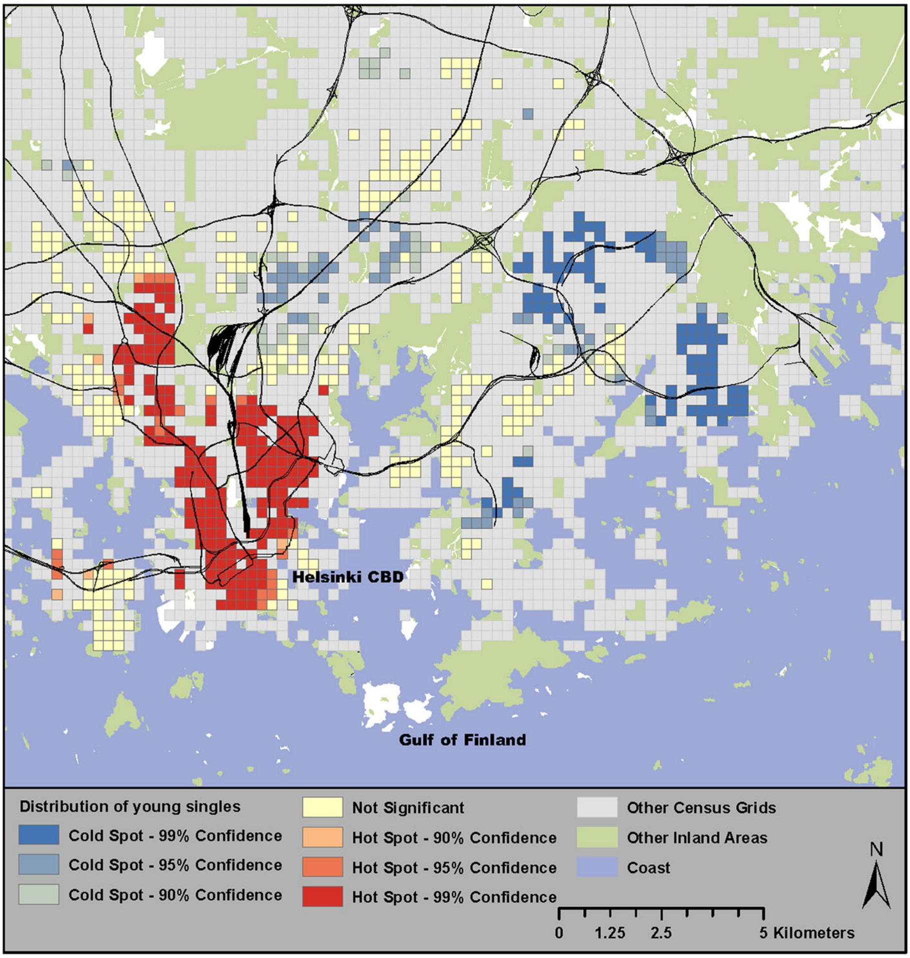

Fig. 4 Hot spot analysis result showing the spatial distribution of young, single person households in Helsinki

(exp(0.002242) – 1)*100] given that the outcome variable was log transformed. However, this finding is intuitive based on a general observation of the distribution of apartment prices in Fig. 3 and the distribution of single households in Fig. 4. Figure 3 shows that apartment prices are significantly higher in the CBD and western parts in Helsinki and significantly lower in eastern part of Helsinki. A similar spatial pattern was observed for the distribution of young, single person household in Helsinki. In addition, a spatial autocorrelation (Global Moran's I) test was conducted using both variables to measure whether the patterns of these two factors are clustered, dispersed, or random in nature. The test results showed that both factors possess a positive Moran's I index (1.087 and 0.318 for singles and property prices respectively) and statistically significant at the 0.001 level which suggest that both factors have a tendency towards clustering. The question is whether the estimated coefficient holds after adjusting for possible endogeneity bias.

4.2 Results from the 2SLS model estimates

The findings from the 2SLS model corroborate to the findings from the OLS model in general as shown in the last three outcome columns in Table 2. The explanatory power of the model is also strong ($R^2=87\%$). The directions of the coefficients are identical and the magnitudes of the coefficients are in an acceptable level of agreement between the models (except that average age of people was not statistically significant in the 2SLS model). However, both the Wu-Hausman and Durbin tests reject the null hypothesis that the percentage of young, single person households is an exogenous variable (as used in the OLS model) and thus confirmed the presence of endogeneity (Durbin test $chi^2=70.40$, $p=0.001$ and; Wu-Hausman F test $=70.45$, $p=0.001$). This means that the coefficients of singles as found in the OLS model is biased. In this research, the spatially lagged of single household was used as an instrument to address the endogeneity issue. Note that finding a suitable instrument is always a difficult task (Spearing et al. 2012). In this case, a strength test result confirmed that the applied instrument was strong (with $R^2=0.73$, adjusted $R^2=0.73$, partial $R^2=0.15$ and $F=8993.38$, $p=0.001$). As expected, the results from the 2SLS model shows a relatively higher coefficient estimate for the single household. It shows that 1 % increase of young, single person households increases the apartment prices by 0.51%.

The impacts of property characteristics were found to be in line with common wisdom and support previous empirical research on this topic. Apartments located in upper levels had been sold in higher prices. However, taller buildings in general reduced the price of apartment. The size of apartments (measured either in terms of the number of room or in square metres) positively affected their prices. The presence of a sauna was found to have a profound effect on apartment prices (increased value by 16%). In contrast, the price fell significantly when the physical condition of the apartments worsen.

As expected, increasing distance from opportunities and services reduced the price of apartments. The coefficients of distances from all five geographic features were found to be very small (e.g. −0.00198% for distance to motorway variable). However, this is due to the fact that these variables were measured in metre. If, instead, these are measured in kilometre, the coefficients show some important role. For example, apartments located one kilometre away from the motorways are likely to reduce price by 2%.

The coefficients of the time dummy variables increased gradually and significantly different from that of in 2005. For example, the prices increased by 10% in 2006 whereas the increase was about 47% in 2012 compared to prices in 2005. These findings suggest a generic upward trend of apartment prices in Helsinki.

5 Conclusion and discussion

Single person households have increased in recent decades with a diminishing number of families. However, an assessment of the impact of this trend on housing market received little attention by the academic community. This paper aims to fill this gap in the literature by investigating a robust property sale database that contained 50,253 apartment transaction records over the period of 2005–2012. The data used in the analyses originated from the city of Helsinki, Finland, and it includes both the apartment final sale prices with the structural characteristics of the apartments and the census data. Two regression models, one OLS and one 2SLS, were estimated to model whether an increasing proportion of young, single person households increases or decreases apartment prices.

The results from the OLS analysis of the sale prices show that a greater proportion of young singles causes a positive effect on the values of the units. However, two types of endogeneity test results (Wu-Hausman and Durbin tests) confirmed that such an estimate is biased due to the strong presence of simultaneity between the variables. To isolate the impact of the lower prices on the singles' housing choices and rule out reverse causality, the 2SLS model is estimated with the spatially lagged of single person household as an instrumental variable. The findings from the 2SLS model generally support the OLS model but with a more pronounced effect of the singles. The findings from the 2SLS model demonstrate that an increasing proportion of young, single, person households increases the apartment prices.

The results from the both models are supportive of the hypothetical understanding as laid out by Kohler and van der Merwe (2015). They highlighted that the reduction of household sizes is an important factor to drive the long-run trends of housing market. They positioned that a decline of household sizes together with other demographic changes creates an underlying demand for new dwellings. The results from this study supports this general understanding that singles prefer to live in urban, city areas where apartment prices are higher than in suburbs as shown in Figs. 3 and 4.

This research also provides an empirical proof that traditional hedonic model is inadequate in explaining the relationship between singles and property demand, and thereby contributes to the development of proper investigation methodologies for similar research. A methodological finding of the study is that there exists a strong endogeneity between the number of young, single households and apartment prices. The study demonstrates how important it is to test endogeneity and rule it out through utilising instrument variable approach.

The main finding of the study is significant for housing market analysts and investors in order for them to understand how different factors affecting the market. Not only demand for housing assets are related to apartment prices but demographic factors including the proportion of single person households plays a significant role in the housing market. Underlying demand of housing relates to population growth (immigration and a fertility rate) but also to average household size. The study confirms that population characteristics are important factors to determine property values.

This study has some limitations in relation to the population data. The census data protects the privacy of individuals, and therefore, none of the socioeconomic characteristics can be attributed to an apartment. In addition, the data do not include variables like the race or religion. Future studies should be conducted with the data containing a detailed socioeconomic characteristics of property buyers and sellers so that these effects can be explicitly controlled for in order to increase the precision of results presented in this study and thereby to enhance our understanding of how property values vary between the socio-economic classes. This research focused only on apartment prices. Future studies should seek to investigate the impacts of singles on other types of properties, such as retirement communities and students housing that often consists of more single-person households. Although the results of this study are evident, more studies from different cities and countries are needed for an external validation of the results presented here. This research provides evidence that young, single households are driving the price of apartments in Helsinki. In a related economic sense, this increasing price of the apartments is likely to increase the supply of apartments through redevelopment or new construction in order to satisfy the demand. Further research is needed to assess the supply related effects of increasing apartment prices induced by the singles.

References

Agnello, L., & Schuknecht, L. (2011). Booms and busts in housing markets: Determinants and implications. *Journal of Housing Economics, 20,* 171–190.

Armstrong, R. J., & Rodríguez, D. A. (2006). An evaluation of the accessibility benefits of commuter rail in eastern Massachusetts using spatial hedonic price functions. *Transportation, 33,* 21–43.

Australian Bureau of Statistics. (2012). *1301.0—Year book Australia, 2012*. ABS website release, Canberra. http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/1301.0~2012~Main%20Features~Home%20page~1.

Benefield, J. D., Cain, C. L., & Johnson, K. H. (2014). A review of literature utilizing simultaneous modeling techniques for property price and time-on-maRKET. *Journal of Real Estate Literature, 22,* 149–175.

Bogue, D. J., Liegel, G., & Kozloski, M. (2009). *Immigration, internal migration, and local mobility in the US*. Books.

Brandt, S., & Maennig, W. (2012). The impact of rail access on condominium prices in Hamburg. *Transportation, 39,* 997–1017.

Cebula, R. J. (2009). The hedonic pricing model applied to the housing market of the city of Savannah and its Savannah historic landmark district. *Review of Regional Studies, 39,* 9–22.

Ceccato, V., & Wilhelmsson, M. (2011). The impact of crime on apartment prices: Evidence from Stockholm, Sweden. *Geografiska Annaler. Series B, Human Geography, 93,* 81–103.

Chandler, J., Williams, M., Maconachie, M., Collett, T., & Dodgeon, B. (2004). *Living alone: its place in household formation and change*. Sociological Research Online 9.

Chiang, Y.-H., Peng, T.-C., & Chang, C.-O. (2015). The nonlinear effect of convenience stores on residential property prices: A case study of Taipei, Taiwan. *Habitat International, 46,* 82–90.

Cho, S.-H., Bowker, J. M., & Park, W. M. (2006). Measuring the contribution of water and green space amenities to housing values: An application and comparison of spatially weighted hedonic models. *Journal of Agricultural and Resource Economics, 31,* 485–507.

Daams, M. N., Sijtsma, F. J., & van der Vlist, A. J. (2016). The effect of natural space on nearby property prices: Accounting for perceived attractiveness. *Land Economics, 92,* 389–410.

de Vaus, D., & Qu, L. (2015). Demographics of living alone. *Journal of the Home Economics Institute of Australia, 22,* 27.

Deka, D. (2014). The living, moving and travel behaviour of the growing American solo: Implications for cities. *Urban Studies, 51,* 634–654.

DePaulo, B. M., & Morris, W. L. (2006). The unrecognized stereotyping and discrimination against singles. *Current Directions in Psychological Science, 15,* 251–254.

Efthymiou, D., & Antoniou, C. (2014). Measuring the effects of transportation infrastructure location on real estate prices and rents: investigating the current impact of a planned metro line. *EURO Journal on Transportation and Logistics, 3,* 179–204.

Eurostat. (2016). Statistical Atlas: Share of single-person households. Retrieved February 11, 2016, from http://ec.europa.eu/eurostat/statistical-atlas/gis/viewer/?config=census.json.

Faessen, W. (2002). Present and preferred housing of one-personhouseholds in the Netherlands, 1998. *Journal of Housing and the Built Environment, 17,* 337–350.

Fingleton, B. (2006). A cross-sectional analysis of residential property prices: The effects of income, commuting, schooling, the housing stock and spatial interaction in the English regions*. *Papers in Regional Science, 85,* 339–361.

Gibbons, S. (2003). Paying for good neighbours: Estimating the value of an implied educated community. *Urban Studies, 40,* 809–833.

Goodman, A. C., & Thibodeau, T. G. (1995). Age-related heteroskedasticity in hedonic house price equations. *Journal of Housing Research, 6,* 25.

Habibov, N., & Afandi, E. (2016). Does life satisfaction determine subjective health? *Applied Research in Quality of Life, 11,* 413–428.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1995). *Multivariate data analysis* (3rd ed.). New York, NY: Macmillan.

Hall, R., & Ogden, P. E. (2003). The rise of living alone in inner London: Trends among the population of working age. *Environment and Planning A, 35,* 871–888.
Hall, R., Ogden, P. E., & Hill, C. (1997). The pattern and structure of one-person households in England and Wales and France. *Population, Space and Place, 3,* 161–181.
Huang, H., & Yin, L. (2015). Creating sustainable urban built environments: An application of hedonic house price models in Wuhan, China. *Journal of Housing and the Built Environment, 30,* 219–235.
Hui, E. C. M., Chau, C. K., Pun, L., & Law, M. Y. (2007). Measuring the neighboring and environmental effects on residential property value: Using spatial weighting matrix. *Building and Environment, 42,* 2333–2343.
Irwin, E. G., Jeanty, P. W., & Partridge, M. D. (2014). Amenity values versus land constraints: The spatial effects of natural landscape features on housing values. *Land Economics, 90,* 61–78.
Juntto, A. (2008). *Asumisen Muutos ja Tulevaisuus, Suomen Ympäristö 33*. Helsinki: Edita Prima Oy.
Kajuth, F., Knetsch, T. A., & Pinkwart, N. (2016). Assessing house prices in Germany: Evidence from a regional data set. *Journal of European Real Estate Research, 9,* 286–307.
Kam, K. J., Chuah, S. Y., Lim, T. S., & Lin Ang, F. (2016). Modelling of property market: The structural and locational attributes towards Malaysian properties. *Pacific Rim Property Research Journal, 22,* 203–216.
Kamruzzaman, M., Baker, D., & Turrell, G. (2015). Do dissonants in transit oriented development adjust commuting travel behaviour? *European Journal of Transport and Infrastructure Research, 15,* 66–77.
Kauppalehti. (2016). https://www.kauppalehti.fi/uutiset/pienten-asuntojen-hinnat-kohoavat-pk-seudulla---tutkijat-poistaisivat-rakentamisen-rajoitteita/ZvggH3bd. Accessed 26 June 2017.
Kauppinen, M., Martelin, T., Hannikainen-Ingman, K., & Virtala, E. (2014). *Yksin asuvien hyvinvointi, Mita talla hetkella tiedetaan?*. Helsinki: Terveyden ja Hyvinvoinnin Laitos.
Kim, J., & Goldsmith, P. (2009). A spatial hedonic approach to assess the impact of swine production on residential property values. *Environmental & Resource Economics, 42,* 509–534.
Klinenberg, E. (2012). *Going solo: The extraordinary rise and surprising appeal of living alone*. New York, NY: Penguin.
Kohler, M., & van der Merwe, M. (2015). Long-run trends in housing price growth. Bulletin September quarter 2015. Reserve Bank of Australia. Retrieved August 15, 2016, from http://www.rba.gov.au/publications/bulletin/2015/sep/pdf/bu-0915-3.pdf.
Lokshin, M., & Ravallion, M. (2008). Testing for an economic gradient in health status using subjective data. *Health Economics, 17,* 1237–1259.
Marmolejo Duarte, C., & González Tamez, C. (2009). Does noise have a stationary impact on residential values? *Journal of European Real Estate Research, 2,* 259–279.
McCord, M., Davis, P., Haran, M., McGreal, S., & McIlhatton, D. (2012). Spatial variation as a determinant of house price: Incorporating a geographically weighted regression approach within the Belfast housing market. *Journal of Financial Management of Property and Construction, 17,* 49–72.
Mukherjee, C., White, H., Wuyts, M., & Wuyts, M. (2013). *Econometrics and data analysis for developing countries*. Florence: Routledge.
Mulley, C. (2013). Accessibility and residential land value uplift: Identifying spatial variations in the accessibility impacts of a bus transitway. *Urban Studies, 51,* 1707–1724.
Official Statistics of Finland. (2016). PX-Web databases. Retrieved February 11, 2016, from http://pxnet2.stat.fi/PXWeb/pxweb/fi/StatFin/StatFin__asu__asas/?tablelist=true&rxid=48fbe1ef-0166-4aed-8bbb-0c1535fa5e88.
Ogden, P. E., & Schnoebelen, F. (2005). The rise of the small household: Demographic change and household structure in Paris. *Population, Space and Place, 11,* 251–268.
Parent, O., & vom Hofe, R. (2013). Understanding the impact of trails on residential property values in the presence of spatial dependence. *The Annals of Regional Science, 51,* 355–375.
Pashardes, P., & Savva, C. S. (2009). Factors affecting house prices in Cyprus: 1988–2008. *Cyprus Economic Policy Review, 3,* 3–25.
Puustinen, S., & Kangasoja, J. (2009). URBA project: Developing new urban housing concepts in the Helsinki metropolitan area. In *Second international conference on whole life urban sustainability and its assessment*. Loughborough University, UK.
Quintano, C., & D'Agostino, A. (2006). Studying inequality in income distribution of single-person households in four developed countries. *Review of Income and Wealth, 52,* 525–546.
Reynolds, M., Wulff, M., & Healy, E. (2004). Why don't small households live in small dwellings?: Disentangling a planning dilemma. *People and Place, 12,* 58.
Rogerson, P. (2001). *Statistical methods for geography*. London: SAGE Publications Ltd.

Statistics Finland. (2017). PX-Web databases. http://pxnet2.stat.fi/PXWeb/pxweb/en/StatFin/StatFin__asu__ashi/?tablelist=true&rxid=429e1165-004f-43cf-84c0-0f07ac8c7d60. Accessed 26 June 2017.

Scott, D., & Horner, M. (2008). Examining the role of urban form in shaping people's accessibility to opportunities: An exploratory spatial data analysis. *Journal of transport and land use, 1,* 89–119.

Spearing, N. M., Connelly, L. B., Nghiem, H. S., & Pobereskin, L. (2012). Research on injury compensation and health outcomes: Ignoring the problem of reverse causality led to a biased conclusion. *Journal of Clinical Epidemiology, 65,* 1219–1226.

Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography, 46,* 234–240.

Turkki, T. (2013). Miljoonien sinkkujen maat. Retrieved April 18, 2016, from http://www.sitra.fi/artikkelit/ita-aasia/miljoonien-sinkkujen-maat.

US Census Bureau. (2012). Households and families: 2010 census brief. Retrieved April 18, 2016, from https://www.census.gov/prod/cen2010/briefs/c2010br-14.pdf.

Webster, A. (2013). *Introductory regression analysis: With computer application for business and economics*. Abingdon: Routledge Ltd—M.U.A.

Wulff, M. (2001). Growth and change in one person households: Implications for the housing market. *Urban Policy and Research, 19,* 467–489.