Exploring the spatial variation in quality-adjusted rental prices and identifying hot spots in Berlin's residential property market

Meulen, Philipp an de; Mitze, Timo Friedel

Exploring the spatial variation in quality-adjusted rental prices and identifying hot spots in Berlin's residential property market

Philipp an de Meulen & Timo Mitze

Routledge
Taylor & Francis Group

Exploring the spatial variation in quality-adjusted rental prices and identifying hot spots in Berlin's residential property market

Philipp an de Meulen[a] and Timo Mitze[b]*

[a]Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI), Essen, Germany; [b]University of Southern Denmark, RWI and RCEA, Sonderborg, Denmark

In this work, we use residual values obtained from an estimated hedonic pricing model to assess the role of district-level neighbourhood effects for the spatial variation in quality-adjusted rental prices in Berlin between 2008 and 2013. By doing so, we also aim at identifying hot and cold spots of quality-adjusted rental prices for apartments across the residential locations within the city of Berlin. The resulting pattern of 'residual' rental prices with a growing concentration of hot spots in central districts of Berlin can be interpreted as the tenants' valorization of apartments in geographic proximity to the city centre compared to similar properties in Berlin's periphery once we control for the properties' physical characteristics. The observed temporal evolution of the rental price distribution between 2008 and 2013 thereby hints at an ongoing gentrification process in Germany's capital associated with the current housing market boom. This visual impression is also confirmed by the application of quantile regressions for a correlation analysis between quality-adjusted rental price values and Berlin district-level characteristics obtained from the last census in 2011. Among other factors, we find that districts' net in-migration balances are positively correlated with quality-adjusted rental price levels for higher quantiles of the distribution, thereby potentially proxying the price dynamics of underlying gentrification processes. Using statistical tests from the explanatory spatial data analysis (ESDA) toolbox, we finally pinpoint particular hot spots of the city's residential property market associated with a significant spatial clustering of similar rental price values around individual observations.

Keywords: housing market; hedonic pricing model; quantile regressions; ESDA; Berlin

Introduction

Berlin's housing market is currently booming. Within the last five years, rental prices have grown by more than 40% on average, with local peaks of almost doubling rental price levels in some inner-city districts.[1] It seems that the property market boom, which has swept over most of Europe in the first decade of this millennium, has finally reached Germany and its capital. While German real estate prices actually fell in the period 2001–07, by the end of the global financial and economic crisis they started to increase considerably – most strongly among metropolitan areas such as Munich, Hamburg, Berlin etc. Although Berlin's mean housing price level per square metre, by now, still does not match with the premium price levels faced in other European capitals such as Inner London or Paris, the regional price dynamics seem to be similar (Deloitte, 2014).

---

*Corresponding author. Email: tmitze@sam.sdu.dk

It appears to be another common pattern within many major cities that, in periods of boom, inner-city house price appreciation is associated with large and amplified spatial differences and proceeds with considerable time lags between the initial growth in hot spots and the adjustments taking place in other parts of the city (if such lagged price convergence takes place at all). The spread of house price increases over a city is often argued to be related to the phenomenon of gentrification. McDonald and McMillen (2011) define gentrification as private and commercial developments, which typically appear in the central city and include movements to the neighbourhood of households with larger incomes than the current residents. In a recent paper, Guerrieri, Hartley, and Hurst (2013) provide evidence for the important role of gentrification in explaining differences in house price movements during periods of appreciation. Considering the history of housing price booms among major metropolitan areas in the United States since the 1980s, the authors find that the largest relative price appreciations took place among neighbourhoods which are cheap themselves but which border to rich neighbourhoods. To explain the data trends, the authors have developed a theory of 'endogenous' gentrification: high-income households move to neighbourhoods with cheap housing prices but which border rich neighbourhoods in order to live close to other high-income households. This drives up prices in the vicinity of the initially rich neighbourhoods in a disproportionate way.

The strong rise of real estate prices and associated rental price levels in the recent past has led to a heated debate in Germany with regard to the proper functioning of the market mechanisms in the real estate sector. The recent price dynamics in the aftermath of the global financial and economic crisis thereby eventually led to recent changes in legislation, implementing an automatic 'rent price break' for congested urban areas.[2] In a European dimension, Germany is indeed a special case since the rental market makes up the vast majority of the overall market for real estate properties. According to Ball (2012), only Switzerland has a higher rental share than Germany. For example, in 2012, the average home-ownership rate (that is, the percentage share of persons living in owner-occupied dwellings in total population) was 53% in Germany compared with 62% in France, 70% in the UK and even 83% in Spain according to data from Eurostat's SILC database (Eurostat, 2014).

Besides this overall characteristic of the German real estate market, there are significant intra-country differences. As Ball (2012) points out, the North and East of Germany have on average lower home-ownership rates than the South and West. According to the author, this difference is partly due to specific state-level and city politics. For instance, Hamburg and Berlin have traditionally emphasized the importance of rental housing for urban planning and development. For the year 2011, Figure 1 shows the regional differences in the ownership rate among German federal states (left part) and the spatial variation at the German NUTS-3 level (right part). Here, with 15.6%, Berlin shows to have a particularly low home-ownership rate highlighting the importance of the rental market within the city.

Based on these insights, it is the aim of this paper to further investigate the spatio-temporal pattern of rental price differences in inner-urban Berlin during the period of the real estate price boom experienced after 2007. Given the importance of the rental market in Berlin, our focus lies on an analysis of rental prices and their spatio-temporal evolution after 2007. To do so, we use internet ads for property rental offers placed on the platform of Immobilienscout24 in 2008 and 2013.[3] In addition to the rental offer, the ads capture the most important physical characteristics of each apartment. This serves as important information to decompose rental prices into their underlying price-determining quality components. By regressing rental offers on their physical
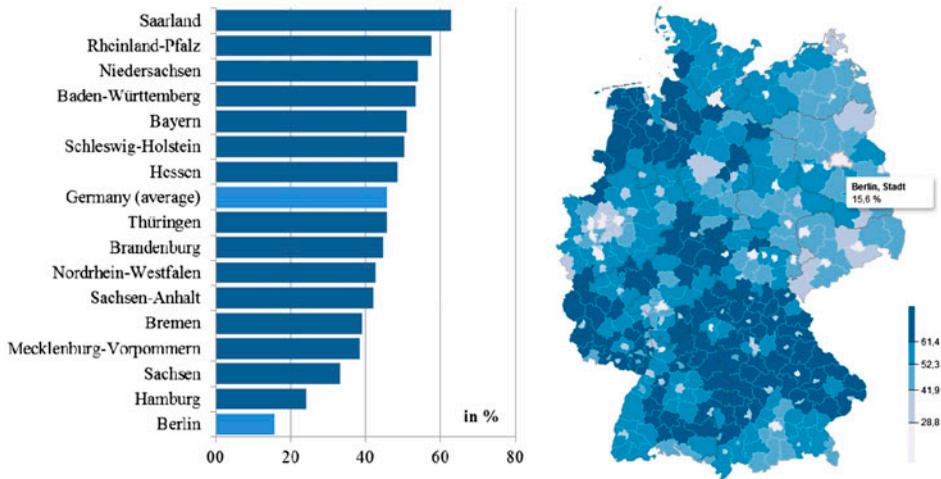
Figure 1.   Home-ownership rates across German federal states (left) and NUTS-3 districts (right).
Source: Statistisches Bundesamt (2013); Census data are for 2011.

characteristics by means of a hedonic pricing model (HPM), we can thus adjust for price differences which owe to differences in apartments' qualities. In what follows, we can consider the HPM regression residuals rather than the original rental offers as variable of interest and interpret them as quality-adjusted or 'excess' rental prices. The estimated remaining differences in rental prices may then be attributed to locational and neighbourhood effects, i.e. to the influences resulting from gentrification processes we are up to identify.

In a second step, we take the residuals of the cross-sectional regressions for the two sample years 2008 and 2013 in order to visualize (location-dependent) rental price distributions for the urban area of Berlin. This mapping approach enables us to identify geographical neighbourhoods associated with rather positive and negative locational characteristics. A visual comparison of the maps further allows us to describe the temporal evolution of spatial price differences. Moreover, using data at the district level in Berlin obtained from the last census in 2011, this enables us to give a first statistical indication with regard to the importance of location and neighbourhood effects for quality-adjusted rental prices and shed some light on the potential correlation between gentrification and the spatial distribution of 'excess' rental price levels. We use quantile regressions to investigate this potential correlation and make use of small-scale migration trends at the district level as well as control for further socio-economic characteristics (such as the share of the district population with a migration background, the average household size, the share of households with singles in all households) and institutional differences among Berlin districts (e.g. the share of buildings in communal ownership). As Ball (2012) points out, institutional factors may be a distinct driver of local real estate markets and Germany has very specific characteristics here as well. A related empirical regression approach has, for instance, been conducted by Case and Mayer (1996): The authors correlate house prices with socio-economic indicators and further town-specific amenities for the metropolitan area of Boston. Their empirical findings suggest that these variables have a notable effect on the pattern of price changes across towns between 1982 and 1992. Moreira de Aguiar, Simoes, and Braz Golgher (2014) analyse the determinants of apartments' prices

in Belo Horizonte, MG, Brazil, by means of of hierarchical models, spatial models and a hierarchical–spatial approach. Their results indicate that local variables explain more than 75% of the prices' remaining variability after controlling for the dwellings' characteristics.

In a third and final step, we use tools from the field of exploratory spatial data analysis (ESDA) to extend our regression approach based on broad locational characteristics at the Berlin district level and measure the spatial correlation between the quality-adjusted rental prices to identify particular rental price clusters. In this regard, high positive spatial correlation may be interpreted to stem from neighbourhood effects potentially associated with gentrification processes. Also, we aim to test for the persistence of historically driven neighbourhood effects in Berlin's residual rental prices, potentially being due to the administrative division of the city prior to German re-unification. As for the visual mapping approach, comparing the results for 2008 and 2013 provides us with useful information how this pattern has changed over the period of booming house prices.

Our findings suggest that the spatio-temporal pattern of rental price differences in Berlin can be brought in line with the notion of gentrification during the period of booming house prices since 2007. Based on the quantile regression approach, we find that the district's net migration balance is positively correlated with higher price levels for the upper tail of the quality-adjusted rental price distribution. With regard to the ESDA analysis, especially rental price increases in districts close to the famous city centre may not fully be attributed to quality improvements in the stock of apartments but rather reflect valorizing image effects of contiguous central neighbourhoods. We find that the tendency of emerging hot spots in and around the city centre has increased in Berlin between 2008 and 2013. While we could identify three separate, polycentric hot spots in 2008, in 2013 high price apartments largely concentrate in the centre of Berlin. Presumably, the centre of Berlin appears to be more attractive for tenants and house owners due to positive neighbourhood effects such as an attractive business environment, good physical and social infrastructure connections, rich cultural offers as well as a snobbish image.

The remainder of the paper is structured as follows. The next section presents the strategy to investigate the spatio-temporal pattern of quality-adjusted rental prices in inner-urban Berlin. The third section presents the results before the fourth section draws the main conclusions from our analysis.

Data and method

In order to provide new insights into the spatial variation and recent spatio-temporal dynamics of the inner-urban rental market in Berlin, we refer to internet ads for property rental offers provided by the platform of Immobilienscout24, which capture the most important physical characteristics of each apartment (such as area of living space, number of rooms, age, equipment etc.). The information on apartments' physical characteristics is important. It enables us to decompose rental prices into a part, which is determined by its physical quality, and a residual term. With apartments being heterogeneous in terms of quality, this decomposition serves the purpose to make apartments comparable goods. Adjusting for differences in quality, the residual information then includes only the price determining factors which are related to an apartment's location.

In what follows, we estimate HPMs to measure the correlation between apartments' physical characteristics and their rental prices (e.g. Gillingham, 1975; Goodman, 1978; Thibodeau, 1989, 1992, 1996). The residuals from these regressions then serve as our variables of interest, which we interpret as apartments' quality-adjusted rental prices.

To visualize their spatial distribution within the metropolitan area of Berlin in the two periods 2008 and 2013, we estimate a cross-sectional HPM equation according to the following stylized equation for both sample years:

$$\log(\text{Rental Price}_i) = f(\text{Structural Hedonic Variables}_i) + \varepsilon_i \qquad (1)$$

where $i = 1, \ldots, N$ is the cross-sectional dimension of the data, with $N = 91,907$ and 79,466 for 2008 and 2013, respectively; 'log(Rental Price$_i$)' denotes the logarithmic transformation of the rental price offer per square metre for apartment i as dependent variable of the HPM specification.[4] The set of physical hedonic variables consists of the property's age, the logarithm of its living space area measured in square metres, the number of rooms, the period of advertisement[5] (both in absolute and squared terms) and the floor the apartment is located in. Moreover, we include a set of binary variables, which provide information whether the apartment is still under construction or not as well as information on the availability of a cellar, an elevator, a garden, a balcony, and a built-in kitchen. A further set of dummy variables controls for an apartment's category, grouped into high- and low-quality and an apartment's overall condition, separated between good and bad.[6] The variable 'high-quality object' is therefore equal to one, whenever the apartment is of high quality, and zero otherwise. Analogously, the binary variable 'bad condition' equals one, basically when there is need of renovation, and is zero otherwise. Finally, we include a set of advertisement duration dummies for all possible periods of advertisement as well as a binary variable which equals 1, if the apartment was already advertised in January 2007, the very first month of the Immobilienscout24 records.[7] Table 1 reports descriptive statistics for each variable in the HPM specification for the sample years 2008 and 2013.

Controlling for all of these physical characteristics, rental prices may nevertheless still differ in a systematic way as a result of different valuations of the geographical locations of the apartments, where the latter location factor typically exerts direct and indirect effects. Direct effects can be associated to the characteristics of the neighbourhood of an apartment. Among these 'neighbourhood effects' one can find accessibility of certain amenities, e.g. public services and facilities (Can, 1992). Following the monocentric model in Urban Economics, location affects the price as it determines the distance to the Central Business District. Indirect effects of location refer to price externalities associated with an apartment's vicinity, which measure underlying socioeconomic factors to the occupants in the neighbourhood. Can and Megbolugbe (1997) mention price mark-ups in high-income neighbourhoods paid for the snobbish image of a particular location.

Since location effects can thus be seen as a suitable proxy for mapping gentrification processes linked to the spatial neighbourhood of the individual apartment, it is exactly this type of residential property market signal that we want to investigate in the following. To do so, in a first step, we adhere to the empirical approach made in Bauer, Fertig, and Vorell (2011) and use the obtained residuals (ε_i) of the HPM equation (1) to draw a cartographic map of Berlin in order to visualize price differences which are not attributed to physical characteristics of the individual apartment but can be linked to the (direct and indirect) location effects.[8] This mapping approach also helps to identify geographical neighbourhoods and to group them by location characteristics, i.e. into concentrations of positive and negative residual values of the estimated HPM specification, respectively. To this end, an 'excess' rental price value for apartment i can be associated with positive neighbourhood effects, which may be subject to a gentrification process.

Table 1. Descriptive statistics for variables used in the Hedonic Price Model (HPM).

Variable	2008				2013			
	Mean	SD	Minimum	Maximum	Mean	SD	Minimum	Maximum
Rent per square metre (€)	5.93	1.70	2.26	45	8.36	2.75	0.05	224.45
Advertisement duration (months)	2.60	2.59	1	24	2.62	4.90	1	84
Squared advertisement duration	13.51	38.56	1	576	30.85	240.89	1	7056
Advertised in January 2007	0.01	0.08	0	1	0.00	0.01	0	1
Age (years)	57.91	36.72	−1	138	61.76	40.28	−1	143
Area of living space (m^2)	73.43	30.94	7.58	491	79.53	39.06	10	472
Floor	2.29	2.11	0	67	2.51	2.48	−1	56
Number of rooms	2.54	1.02	1	10	2.60	1.09	1	10
Cellar	0.14	0.35	0	1	0.56	0.50	0	1
Elevator	0.26	0.44	0	1	0.35	0.48	0	1
Garden	0.08	0.27	0	1	0.13	0.34	0	1
Balcony	0.67	0.47	0	1	0.70	0.46	0	1
Built-in kitchen	0.37	0.48	0	1	0.46	0.50	0	1
High-quality object	0.09	0.28	0	1	0.13	0.34	0	1
Bad condition	0.38	0.48	0	1	0.31	0.46	0	1
Under construction	0.00	0.02	0	1	0.01	0.10	0	1

Note: SD is the standard deviation for each variable.
Source: Authors' own calculations based on data from Immobilienscout24.

As outlined above, the empirical validity of this visual data inspection is then checked with the help of a stylized regression approach and ESDA tools. For the regression approach, we use quantile regressions as a method for fitting a regression line through the conditional quantiles of a variable's distribution (for instance, see Koenker & Bassett, 1978).[9] As Koenker and Hallock (2001) point out, the quantile regression approach estimates conditional quantile functions, in which quantiles of the conditional distribution of the regressand are expressed as functions of observed covariates. By doing so, the approach allows to focus on the lower and upper tails of a given distribution, which may be of particular relevance for the analysis of our HPM residuals, where a regression exercise built upon the conditional mean framework may disregard important 'excess' variation in the data.

Specifically, we regress ε_i from equation (1) against a set of demographic and housing characteristics at Berlin's district level in order to correlate 'excess' rental price values with potential drivers of neighbourhood effects such as the district's net migration balance, the average household size and the share of population with migration background, the share of builds in communal ownership. The district characteristics are extracted from the last German census in 2011. We use regression residuals from 2013 as dependent variable in order to guarantee the predetermined nature of the census information. Definitions and descriptive statistics for the census data on Berlin districts are given in Table 2.

Table 2. Definitions and descriptive statistics for district-level data in Berlin according to the Census 2011.

Variable	Description	Source	Mean	SD	Minimum	Maximum
Net migration balance	Share of gross in-migration minus gross out-migration for the district as a percentage share of the district's total population	Data for Census 2011 were obtained from Amt für Statistik Berlin-Brandenburg (2014)	1.30	0.47	–0.08	1.75
Population with migration background	Percentage share of residents with a migration background in the total population of the district	Data for Census 2011 were obtained from Amt für Statistik Berlin-Brandenburg (2014)	24.45	11.62	6.8	43.4
Average household size	Average household size per district (persons)	Data for Census 2011 were obtained from Amt für Statistik Berlin-Brandenburg (2014)	1.83	0.072	1.71	1.93
Share of single households	Percentage share of households consisting of singles in total households	Data for Census 2011 were obtained from Amt für Statistik Berlin-Brandenburg (2014)	49.57	4.64	41.2	55.9
Buildings in communal ownership	Percentage share of buildings in communal ownership in the total stock of buildings per district	Data for Census 2011 were obtained from Amt für Statistik Berlin-Brandenburg (2014)	5.49	4.28	0.2	20.3
Owner-occupied dwellings	Percentage share of owner-occupied apartments in the total stock of apartments per district	Data for Census 2011 were obtained from Amt für Statistik Berlin-Brandenburg (2014)	14.89	6.49	5.6	26.6
Buildings constructed after 2000	Percentage share of buildings built after the year 2000 in the total stock of buildings per district	Data for Census 2011 were obtained from Amt für Statistik Berlin-Brandenburg (2014)	7.97	5.88	2.4	19.8

We use simultaneous quantile regressions for the 5th, 10th, 25th, 50th, 75th, 90th and 95th quantiles as a robust regression technique accounting for non-normal residuals and heteroskedasticity (Koenker & Hallock, 2001). Standard errors are bootstrapped based on the entire variance–covariance matrix of the simultaneous regression approach. For the *m*th quantile (Q_m), the estimation equation has the following form:

$$Q_m(\varepsilon_{i,s}) = \beta_m'\mathbf{X}_s + u_{i,s} \tag{2}$$

where $\mathbf{X}_s$ is a vector of district characteristics for the $s = 1, \ldots, 12$ districts in Berlin (Figure A1); and β_m is a vector of regression coefficients, which may vary across quantiles and measures the degree of correlation between quality-adjusted rental price values and district characteristics. Since the intra-urban variation for the set of location characteristics is limited to data for the 12 Berlin districts, the regression approach can only serve as a broad proxy for such location effects impacting on the apartments' rental price levels. Hence, as a second complementary pillar of the explorative statistical analysis, we employ ESDA tools to test for the degree of spatial dependence in our property rental price data by means of computing global and local Moran's *I* statistics for each sample year. The global Moran's *I* statistic (Moran, 1950) is defined as:

$$I = \frac{N}{\sum\limits_{i=1}^{N}\sum\limits_{j=1}^{N} w_{ij}} \cdot \frac{\sum\limits_{i=1}^{N}\sum\limits_{j=1}^{N} w_{ij}(\varepsilon_i - \bar{\varepsilon})(\varepsilon_j - \bar{\varepsilon})}{\sum\limits_{i=1}^{N}(\varepsilon_i - \bar{\varepsilon})^2} \tag{3}$$

The formulation of Moran's *I* thereby closely resembles the design of classical correlation coefficients based on the calculation of a cross-product for two variables, except that - in this case - besides the untransformed residual vector of the HPM equation (1), the variable's spatial lag ($\mathbf{W}\varepsilon$) is used for the bivariate correlation analysis. The latter spatial lag, in turn, can be defined for cross section *i* as:

$$\sum\limits_{j=1}^{N} w_{ij} \times (\varepsilon_j - \bar{\varepsilon}),$$

where $\bar{\varepsilon}$ is the variable's cross-sectional average; and w_{ij} is a measure of the spatial connectivity between two cross-sectional units *i* and *j* as element in the $N \times N$ spatial weighting matrix $\mathbf{W}$ with $i \neq j$. Moran's *I* can take values between –1 and 1, where a value of $I > 0$ indicates positive spatial autocorrelation, while values < 0 hint at negative spatial autocorrelation for the sample of cross-sectional units. Since this global measure of spatial autocorrelation is, however, not able to identify certain clusters of 'excess' rental prices, we also apply the local version of Moran's *I* in order to detect neighbourhoods with above and below average property prices levels, respectively. The local Moran's I_i for cross-section *i* can be defined as:

$$I_i = \frac{\sum\limits_{j=1}^{N} w_{ij}(\varepsilon_i - \bar{\varepsilon})(\varepsilon_j - \bar{\varepsilon})}{\frac{1}{N}\sum\limits_{i=1}^{N}(\varepsilon_i - \bar{\varepsilon})} . \tag{4}$$

The local Moran's I_i values are related to the global Moran's *I* since:

$$I = \sum_{i=1}^{N} I_i.$$

As Anselin (1995) points out, local indicators for spatial association (LISA) such as I_i can be interpreted as indicators of local pockets of non-stationarity or hot spots.[10] Moreover, these indicators may be used to assess the influence of individual locations on the magnitude of the global Moran's I statistic and thus to identify outliers.

To apply the global and local version of the Moran's I test of spatial dependence, we need to define a measure for spatial association among the individual apartments. Since we are dealing with point data of real estate properties, we operationalize the design of the spatial weighting matrix **W** by constructing Thiessen polygons (Voronoi Tessellation) for each house location (for a similar approach, see, for instance, Conway, Li, Wolch, Kahle, & Jerrett, 2010). Matrix elements of **W** are both constructed as binary entries based on first-order contiguities, where the resulting matrix has been additionally row-standardized, as well as based on inverted distances between each pair of apartments. The latter inverse distance-based indicator is used as default specification for the construction of **W**.

Empirical results

We start with the estimation of the HPM as outlined in Bauer *et al.* (2013). Table 3 presents the regression output of the HPM specification according to equation (1). As Table 3 shows, we obtain meaningful coefficients for the set of physical housing characteristics for both sample years. For instance, the variable log(Area of living space) is statistically significant and positively correlated with log(Rental price per square metre). That is, a 1% increase in apartment size leads to a 0.08–0.1% increase in the per square metre rental price offer. Likewise, having a cellar, an elevator, a garden, and a built-in kitchen are, everything else equal, associated with a higher price level, while other characteristics (floor, balcony, number of rooms) show a negative correlation with the rental price level if controlled for other physical characteristics. A larger number of rooms, e.g., is tantamount to smaller rooms on average when controlled for the living space. Not surprisingly, this reduces a property's attractiveness. As expected, a 'high-quality object' is associated with a higher rental price, while a 'bad condition' implies a deduction on the rental price of the property.

The residuals of the HPM regression for the years 2008 and 2013 are visualized in Figure 2 based on the apartments' exact geographical location in Berlin. The solid lines in the figure mark the outer geographical boundary of Berlin and the borders of the city's 12 administrative districts (Bezirke). Additionally, the dark grey polygons in the figure visualize the city's stock of buildings in order to provide a visual check whether the Immobilienscout24 data are rather equally distributed among the different residential locations in Berlin or not.[11]

To visualize hot and cold spots of HPM quality-adjusted rental prices, we use a simple categorisation of the residual values, where 'red' data points indicate moderate to strong positive residual values, 'orange' data points indicate the range of slight positive to slight negative residual values around the average Berlin rental price level, while 'yellow' points indicate moderate to strong negative residual values. A first observation from Figure 2 is that for both sample years our dataset of Berlin properties covers most parts of the city's current stock of buildings and thus does not show to

Table 3. Hedonic regression results for 'Rental price per square metre' of apartments in Berlin, 2008 and 2013.

Dependent variable: log(Rental price per square metre)	2008	2013
Advertisement duration	0.127***	0.0351***
	(0.00855)	(0.00530)
Squared advertisement duration	−0.00292***	−0.000185**
	(0.000533)	(0.0000636)
Advertised in January 2007	0.00555	−0.247
	(0.0126)	(0.139)
Age	0.000626***	0.00126***
	(0.0000268)	(0.0000272)
log(Area of living space)	0.0864***	0.158***
	(0.00584)	(0.00533)
Floor	−0.00748***	−0.0103***
	(0.000407)	(0.000444)
Number of rooms	−0.0266***	−0.0460***
	(0.00190)	(0.00189)
Cellar	0.0164***	0.0143***
	(0.00240)	(0.00192)
Elevator	0.0695***	0.103***
	(0.00205)	(0.00224)
Garden	0.0806***	0.0407***
	(0.00323)	(0.00289)
Balcony	−0.0138***	−0.0329***
	(0.00168)	(0.00203)
Built-in kitchen	0.111***	0.138***
	(0.00185)	(0.00205)
High-quality object	0.149***	0.114***
	(0.00331)	(0.00309)
Bad condition	−0.0235***	−0.130***
	(0.00161)	(0.00216)
Under construction	0.588***	0.125***
	(0.0415)	(0.00764)
Adjusted R^2	0.984	0.986
Number of observations	91,907	79,466

Note: Regressions include advertisement duration dummies, whose coefficient results are available upon request. Standard errors are given in parentheses.
*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

suffer from a particular location bias. Second, with regard to concentrations of positive residual values of the HPM specification (red crosses), these can be found in and close to the city centre including districts such as Berlin-Mitte, Tiergarten and Prenzlauer Berg.[12] With the exception of the south-western part of the city (likely due to the proximity to the Wannsee), in tendency, rental prices seem to be penalized in their value with increasing distance to the city centre.

Noteworthy, two decades after German re-unification we do not observe any particular East–West divide in 'excess' rental prices.[13] Rather, locations in the north-western and south-eastern part of the city are characterized by below average rental price levels. For the latter south-east corridor from Berlin's city centre to the periphery, rental prices have been marked below the citywide mean in 2008 - even at moderate distances to the city centre (such as the districts Friedrichshain-Kreuzberg and northern parts of Neukölln). The latter two districts are shown in greater detail in the 'zoomed' map
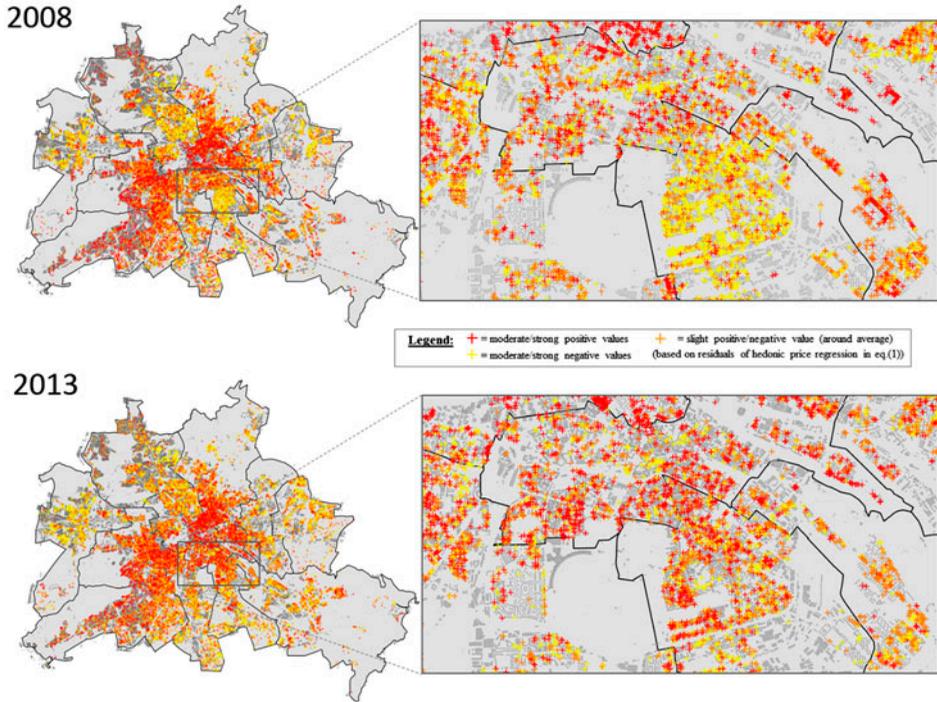
Figure 2.   Quality-adjusted rental price differences for apartments in Berlin (2008 and 2013).
Source: Authors' own figure based on data from Immobilienscout24, Geofabrik and Geoportal Berlin-Brandenburg.

sections in Figure 2 (upper right graph in Figure 2 for 2008). If we compare the change in the (residual) rental price distribution between 2008 and 2013, the general spatial distribution seems to be stable over time. However, the above mentioned south-central districts with a moderate distance to the city centre have considerably caught up in their rental price valorization with the city centre (see change in the 'zoomed' map sections between 2008 and 2013). This tendency supports the above mentioned hypothesis of an ongoing gentrification process in Berlin, which consecutively upgraded apartments in moderate geographical proximity to the city centre.

Given the limited supply of housing space in the inner city, offered properties in the spatial neighbourhood of the close-to-centre districts also grow in value and become more and more attractive to tenants. This demand pressure may result in a positive price spiral for these apartments compared to similar properties in greater distance to the centre. Compared to this catch-up tendency of locations with moderate distance to the city centre, properties in the southern and northern periphery seem to have further lost their attractiveness (especially in Steglitz-Zehlendorf and the Southern part of Tempelhof-Schöneberg). Among the few exceptions of rising rental prices in the northern periphery is Tegel (as part of the district Reinickendorf), which has benefited from a specific 'shock', namely the announced shut-down of the airport Tegel in course of the construction of the new international airport Berlin-Brandenburg.

Table 4. Quantile regression results for correlation among HPM residuals and location factors.

Dependent variable: Residuals from hedonic price regression for 2013	Q_{05}	Q_{10}	Q_{25}	Q_{50}	Q_{75}	Q_{90}	Q_{95}
Net migration balance	-1.024*	0.571	-0.849***	-0.765***	1.105***	1.767***	2.048***
	(0.5959)	(0.4023)	(0.2641)	(0.2372)	(0.2216)	(0.3373)	(0.4272)
Population with migration background	-0.002**	-0.002***	-0.003***	-0.001***	0.001***	0.005***	0.010***
	(0.0009)	(0.0006)	(0.0004)	(0.0004)	(0.0004)	(0.0006)	(0.0007)
Average household size	0.348**	0.217**	0.339***	0.273***	-0.219**	-0.715***	-1.156***
	(0.1371)	(0.1003)	(0.0675)	(0.0693)	(0.0849)	(0.1036)	(0.1115)
Share of single households	-0.004	-0.006***	0.001	0.005***	0.002	-0.002	-0.008***
	(0.0027)	(0.0020)	(0.013)	(0.0014)	(0.0017)	(0.0021)	(0.0024)
Buildings in communal ownership	0.001	-0.001	-0.002***	-0.001**	0.001**	0.004***	0.007***
	(0.0009)	(0.0079)	(0.0006)	(0.0004)	(0.0005)	(0.0007)	(0.0011)
Owner-occupied dwellings	-0.004***	-0.003***	-0.004***	-0.003***	0.004***	0.009***	0.014***
	(0.0012)	(0.0012)	(0.0008)	(0.0005)	(0.0007)	(0.0008)	(0.0015)
Buildings constructed after 2000	-0.001	-0.003***	-0.003***	-0.0003	0.001	0.004***	0.009***
	(0.0017)	(0.0009)	(0.0008)	(0.0006)	(0.0007)	(0.0010)	(0.0012)
Number of observations	79,466	79,466	79,466	79,466	79,466	79,466	79,466
Pseudo-R^2	0.024	0.016	0.006	0.001	0.007	0.014	0.018

Note: Bootstrapped standard errors in parentheses.
*p<0.10; **p<0.05; ***p<0.01.

Turning to the results of the simultaneous quantile regressions, Table 4 reports the conditional correlation coefficients for average district characteristics in 2011 and quality-adjusted rent price values from the HPM for 2013. The results show that most of the neighbourhood proxies have a dual impact on quality-adjusted rental price levels conditional on the chosen quantile of the regression. That is, the district's net migration balance as share of its total population is positively correlated with quality-adjusted rental price levels for higher quantiles of the 'excess' rental price distribution (Q_{75}, Q_{90}, Q_{95}). This indicates that especially for the upper tail of the distribution, net in-migration trends can be seen as a potential proxy for gentrification processes and may drive quality-adjusted rental prices upwards. At the lower tail of the distribution, however, a higher net in-migration balance actually reduces the adjusted rental price level of the apartment across districts. In fact, the observed pattern may thus reflect the dual role of migration in an urban city economy: On the one hand, it reflects the above mentioned gentrification process following a demand in high-end neighbourhoods. At the same time, it appears that residents which cannot afford this rental price adjustment are crowded out into low-end housing neighbourhoods, where a higher in-migration balance is seen as a bad signal for the real estate market.

A similar self-selection process into different rental price regimes is also observed for the average household size (low-end neighbourhoods are positively correlated with household size, while high-end neighbourhood are negatively correlated with household size proxying differences in the disposable income per household), for the share of population with migration background and the share of owner-occupied dwellings. Other variables such as the share of households with singles in all households is negatively correlated with the quality-adjusted rental price level over all included quantiles of the distribution. With regard to institutional factors, the share of buildings in communal ownership in the total district's housing stock is found to be positively correlated with the 'excess' rental price level for quantiles at the upper end of the distribution, while we do not find statistically significant parameter estimates for quantiles at the lower end of the distribution. Although these inter-district variations in Berlin are thus able to partly explain the spatial heterogeneity of 'excess' rental prices and may support the argument of gentrification as one driver of this pattern, the results nevertheless have to be interpreted carefully given that the district level can only proxy broader trends and cannot account for intra-district differences such as observed for Neukölln in Figure 2. Further, given the limited number of control factors, we cannot fully account for all potential unobserved district factors, which makes it hardly possible to interpret the correlations as causal impacts.

Thus, as an alternative approach to explain the visual differences among 'excess' rental price levels in Figure 2, we employ ESDA tools. Thereby, the impression of an expanding inner-city core of 'excess' rental prices is also supported by the application of

Table 5. Results from global Moran's *I* statistic for alternative spatial weighting schemes.

Year	Type of **W**	Inverse distances	First-order contiguity
2008	Moran's *I*	0.413	0.497
	(*p*-value)	(0.00)	(0.00)
2013	Moran's *I*	0.426	0.386
	(*p*-value)	(0.00)	(0.00)

the global and local Moran's *I* tests. Table 5 reports the results of the global Moran's *I* statistic for both sample years and alternative operationalizations for the spatial weighting matrix **W**. As the table shows, for both sample years the null hypothesis of no spatial association among the observations is clearly rejected for reasonable significance levels. The obtained test statistics thus indicate that positive and negative (residual) price values are clustered together in space. While the test results also appear to be robust with respect to the chosen operationalization of **W** from a qualitative point of view, it is worth noticing the quantitative difference in the reported Moran's *I* test statistics in Table 5.

If we measure the degree of spatial autocorrelation based on inverse distances, it appears that spatial dependence of quality-adjusted rental prices have grown over the period 2008–13. However, the opposite seems to be true if we use the row-standardized first-order contiguity weighting matrix to measure spatial linkages. To explain the basic differences in the two concepts of spatial association, the row-standardized first-order contiguity matrix is a more restrictive (local) presentation of spatial relations, while using inverse distances basically implies that each apartment in Berlin is somehow linked to all other apartments within the urban area (although declining in distance). Hence, if we rather consider Berlin as *one* integrated residential property market, weighting schemes based on inverse distances can be seen as the more appropriate estimate for the presence of location effects in Berlin. Then, the results in Table 5 lead us to conclude that spatial autocorrelation has moderately grown over the considered period 2008–13, reflecting the ongoing polarization of the city's residential property market.

Moving from a global measure of spatial association to Anselin's (1995) LISA approach, Figure 3 plots the results of the local Moran's *I* indicator for the two sample years 2008 and 2013.[14] Red surfaces in the figure (based on Thiessen polygons) indicate a statistically significant local Moran's *I* value for observation *i*, which is associated with high residual price values for apartment *i* and high residual price values in the spatial neighbourhood of *i* (based on the HPM function). This means that red surfaces indicate local clusters of apartments associated with quality-adjusted rental price levels above average. Likewise, yellow surfaces indicate a statistically significant local Moran's *I* index characterized by the simultaneous presence of low (negative) residual values for apartment *i* and its spatial neighbourhood. While grey areas show statistically insignificant results from the local Moran's *I* indicator for *i*, light and dark orange surfaces indicate a mix of high residual values for *i* together with a low value in the spatial neighbourhood and vice versa.

As Figure 3 shows, the sample year 2008 was characterized by several clusters of high–high (HH) rental price level combinations (red areas in Figure 2) hinting at a polycentric urban structure with regard to residential properties. The main concentration of above-average priced apartments can be found in the south-western part of Berlin, surrounding the Wannsee. A second concentration of HH combinations – potentially belonging to the same location band as the Wannsee area – can be found in the city centre. Moreover, local HH clusters can also be found in the northern part of the city (i.e. Tegel) and the south-east end (Treptow-Köpenick).

Until 2013 the picture has changed, though, foremost with respect to the location of clusters indicating above-average priced apartments. Compared to the multiple peak rental price map for 2008, the HH clusters now tend to concentrate in the inner districts of Berlin, while high price clusters in the periphery appear to have vanished. Moreover, it seems that the Wannsee area and the city centre now join a common band of high priced apartments, which gradually moves towards the centre. As it was already shown in Figure 2, especially regions in the south-central part of the city have experienced a
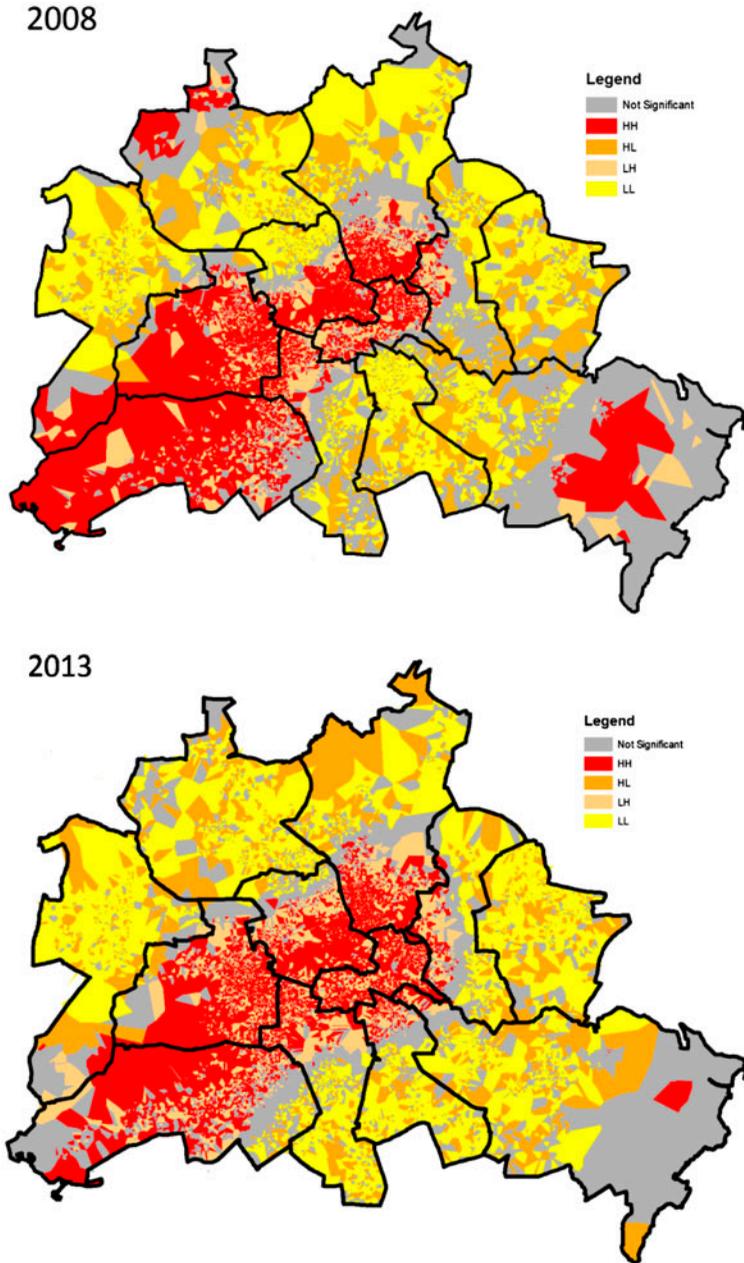
Figure 3.   Cluster Analysis of Berlin property data by means of local Moran's *I* (2008 and 2013). HH = combination of high residual values for apartment *i* and its spatial neighbourhood (together with a statistically significant local Moran's *I* test); LL = a combination of low residual values for apartment *i* and its spatial neighbourhood (together with a statistically significant local Moran's *I* test); LH = a mix of low residual value for apartment *i* with high value in spatial neighbourhood; HL = a mix of high residual value for apartment *i* with low value in spatial neighbourhood; grey areas indicate insignificant local Moran's *I* test statistics for observation *i*. Computations are based on spatial weighting matrix with inverse distances between apartments.

locational valorization and are now characterized by statistically significant local Moran's I values associated with high residual values for apartment i and its spatial neighbourhood (particular Friedrichshain-Kreuzberg and northern parts of Neukölln). The results from the local Moran's I map thus support the hypothesis of an ongoing gentrification process in Berlin, which is associated with on average rising rental prices in Germany's capital.

Conclusions

This paper has investigated the current rental price dynamics in the city of Berlin with a special focus on the role played by the spatial variation in quality-adjusted rental price levels proxying the importance of location and neighbourhood effects. Gaining insights into the spatio-temporal dynamics of such location effects can help to shed light on the question whether we observe ongoing gentrification processes in Germany's capital or not. To this end, we have estimated a HPM to identify 'hot' and 'cold' spots of quality-adjusted rental price levels for the metropolitan area of Berlin in 2008 and 2013. Afterwards, our empirical analysis consisted of three steps: 1) a visual inspection of the quality-adjusted rental price levels, 2) an explorative quantile regression approach and 3) the use of ESDA statistical tools applied to the residual values of the HPM. Our empirical results show that the underlying rental price levels for residential properties are indeed subject to differences in socio-economic and institutional factors at the district level and exhibit positive spatial association. If we view Berlin as a contiguous housing market, the degree of spatial dependence has increased over the time period considered.

The results of the explorative quantile regressions have shown that most of the neighbourhood proxies follow a dual regime over the distribution of quality-adjusted rental price levels. For instance, the district's net migration balance is positively correlated with quality-adjusted rental price levels for higher quantiles of the distribution (Q75, Q90, Q95), while, at the lower tail, a higher net in-migration balance actually reduces the adjusted rental price level of apartments across districts. This indicates that especially for the upper tail of the distribution, net in-migration trends can be seen as a potential indicator for the outcome of gentrification processes driving quality-adjusted rental prices upwards. A similar self-selection process into different rental price regimes is also observed for the average household size, for the share of population with migration background and the share of owner-occupied dwellings in the total housing stock of a district.

Based on the computation of the local Moran's I indicator, we were then finally able to pinpoint particular 'hot spots' of the city's residential property market associated with a significant spatial clustering of similar rental price values around individual observations. While we could identify at least three 'hot spots' for our reference year 2008, spread over the city, five years later, contiguous neighbourhoods of distinct high price locations increasingly tend to concentrate in the inner districts of Berlin with a circularly growing central 'hot' spot location. Taken together, the empirical results of the explorative quantile regression approach and the visualization of the local Moran's I statistic thus support the impression of an ongoing concentration trend, which may be associated with the concept of 'endogenous gentrification' as described in Guerrieri *et al.* (2013): high-income households gradually move to neighbourhoods with cheap housing prices but which border rich neighbourhoods in order to live close to other high-income households. The pace of this process appears to be fuelled by the current property market boom in Germany's capital.

Of course, we are aware that a visual inspection, a first explorative regression approach as well as a spatial data analysis of the spatio-temporal dynamics of metropolitan property rental prices can only serve as a starting point for an in-depth investigation of this issue. Therefore, future research on the role of neighbourhood characteristics as a source for local (price) externalities and an analysis of the pattern of 'endogenous gentrification' should clearly further move from a visual–descriptive and explorative analysis of the Berlin residential property market to a rigorous analysis based on statistical and regression-based causal evaluation tools.[15] This may also help to answer the related question whether house prices and rental prices within a metropolitan area diverge or converge over time. The latter hypothesis can then be taken to quantitatively investigate the validity of the 'rent gap hypothesis' (Smith, 1979). Shedding light on these questions may lastly also help to provide further valuable input to the heated public debate over 'excess' rental prices and their current dynamics in German metropolitan areas and may guide policy makers in shaping the proper institutional framework and the legal boundaries for the future development of the German residential property market.

Acknowledgements

Notes

1. According to the rental price development of the IMX index published by Immobilienscout24 (for details, see http://www.immobilienscout24.de/immobilienbewertung/immobilien index/berlin.html). Regarding the index construction, see Bauer *et al.* (2013).
2. For details on the new legislation, see DW (2014).
3. According to the authors' own assessment of Immobilienscout24, about half of the housing supply on the German real estate market is covered by the online platform of Immobilienscout24. For further details, see Georgi and Barkow (2010).
4. Since it is possible that the advertisement process takes more than one month, apartments may occur repeatedly in the data set. To exclude serial correlation in the error term from repeated observations, we exclude all advertisements on the same apartment except for the last announcement. Taking into account the very last advertisement appears to be a reasonable selection strategy provided that the reason of the withdrawal is the sale of the apartment. The difference between the last announced offer price and the transaction price is then likely to be small.
5. This gives the number of months that an apartment has been advertised on the platform of Immobilienscout24 before it was removed from the platform.
6. This information is based on apartment owners' own assessments.
7. Since data are not available before 2007, for those apartments it is not observable if they have been advertised before. This is also why we have decided to take 2008 rather than 2007 as the reference year.
8. A positive residual value for apartment i in the hedonic price equation can be interpreted as such that the property has an above-average (offered) rent price after controlling for its physical characteristics. Likewise, a negative residual value hints at a below-average (offered) rent price conditional on the apartment's characteristics.
9. In a similar approach, McMillen (2008) applies quantile regressions directly to the HPM model in equation (1).
10. For details on statistical tests to assess the null hypothesis of no spatial autocorrelation based on the moments of the global and local version of Moran's *I*, see, for instance, Anselin (1995).
11. Information on the stock of buildings for Berlin was obtained from the OpenStreetMap project and downloaded from Geofabrik (http://download.geofabrik.de/europe/germany/berlin.

html). The shapefile for Berlin's districts was obtained for the Geoportal Berlin-Branden-burg. Based on this information, Figure 1 was then constructed with the help of ArcGIS software.

12. A map of city districts with the associated district names can be found in Appendix A.
13. To test for persistent West–East differences in the residuals from the HPM equation, which may reflect path dependencies from the inner-city divide prior to German re-unification, we tested for the statistical significance of a binary dummy variable, which was defined as being equal to 1 for apartments located in former West Berlin, and 0 otherwise, when regressing the residual rental prices on this dummy. The estimation results, however, show that the coefficient of this dummy variable is not statistically significant from 0 for both sample years.
14. The sum of the local Moran's I indicator values is proportional to the global indicator (Anselin, 1995). Detailed test results for the local Moran's I indicator per dwelling are available from the authors upon request.
15. For an extensive long-run study on the role played by localized production and residential externalities on the urban structure in Berlin during the period of the city's division and re-unification, see, for instance, Ahlfeldt *et al.* (2014).

References

Ahlfeldt, G., Redding, S., Sturm, D., & Wolf, N. (2014). *The Economics of Density: Evidence from the Berlin Wall* (NBER Working Paper No. 20354). Cambridge, MA: National Bureau of Economic Research (NBER).

Anselin, L. (1995). Local Indicators of Spatial Association - LISA. *Geographical Analysis, 27*(2), 93–115.

Ball, M. (2012). European Housing Review 2012 (Report February 2012). RICS Research. Retrieved October 10, 2014 from http://www.rics.org/research

Bauer, T., Fertig, M., & Vorell, M. (2011). *Neighborhood Effects and Individual Unemployment* (IZA Discussion Paper No. 6040). Bonn: Institute for the Study of Labor (IZA).

Bauer, T., Feuerschütte, S., Kiefer, M., & an de Meulen, P., Micheli, M., Schmidt, T., & Wilke, L.-H. (2013). Ein hedonischer Immobilienpreisindex auf Basis von Internetdaten: 2007–2011 [A hedonic real estate price index based on internet data: 2007–2011]. *ASTA Wirtschafts- und Sozialstatistisches Archiv, 7*, 5–30.

Can, A. (1992). Specification and estimation of hedonic housing prices models. *Regional Science and Urban Economics, 22*, 453–474.

Can, A., & Megbolugbe, I. (1997). Spatial dependence and house price index construction. *Journal of Real Estate Finance and Economics, 14*, 203–222.

Case, K., & Mayer, C. (1996). Housing price dynamics within a metropolitan area. *Regional Science and Urban Economics, 26*, 387–407.

Conway, D., Li, C., Wolch, J., Kahle, C., & Jerrett, M. (2010). A spatial autocorrelation approach for examining the effects of urban greenspace on residential property values. *Journal of Real Estate Finance and Economics, 41*, 150–169.

Deloitte (2014). *Property Index – Overview of European Residential Markets. European Housing 2013* (3rd Edn, June). Retrieved July 14, 2014 from http://www2.deloitte.com/content/dam/Deloitte/de/Documents/real-estate/real-estate-property-index-deloitte.pdf

DW (2014). Merkel's cabinet approves 'brake' on rent prices. October 1. Retrieved October 7, 2014 from http://dw.de/p/1DOLn

Eurostat (2014). 'Distribution of population by tenure status, type of household and income group. Retrieved October 7, 2014 from http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home

Georgi, S., & Barkow, D. (2010). *Wohnimmobilien-Indizes: Vergleich Deutschland-Großbritannien*. Zentraler Immobilien Ausschuss e.V. Retrieved October 10, 2014 from http://www.zia-deutschland.de/index.php/tools/required/download.php?fID=107/

Gillingham, R. (1975). Place to place rent comparisons. *Annals of Economic and Social Measurement, 4*(1), 153–178.

Goodman, A. C. (1978). Hedonic prices, price indices and housing markets. *Journal of Urban Economics, 5*(4), 471–484.

Guerrieri, V., Hartley, D., & Hurst, D. (2013). Endogenous gentrification and housing price dynamics. *Journal of Public Economics, 100*, 45–60.

Koenker, R., & Bassett, G. (1978). Regression quantiles. *Econometrica, 46*(1), 33–50.

Koenker, R., & Hallock, K. (2001). Quantile regression. *Journal of Economic Perspectives, 15*(4), 143–156.

McDonald, J., & McMillen, D. (2011). *Urban Economics and Real Estate – Theory and Policy* (2nd Edn). Hoboken, New Jersey: Wiley.

McMillen, D. (2008). Changes in the distribution of house prices over time; structural characteristics, neighborhood, or coefficient? *Journal of Urban Economics, 64*, 573–589.

Moran, P. (1950). Notes on continuous stochastic phenomena. *Biometrika, 37*, 17–23.

Moreira de Aguiar, M., Simoes, R., & Braz Golgher, A. (2014). Housing market analysis using a hierarchical–spatial approach: the case of Belo Horizonte, Minas Gerais, Brazil. *Regional Studies, Regional Science, 1*(1), 116–137.

Smith, N. (1979). Toward a theory of gentrification: a back to the city movement of capital, not people. *Journal of the American Planning Association, 45*(4), 538–548.

Statistisches Bundesamt (2013). *Zensus (2011). Erste Ergebnisse des Zensus 2011 für Gebäude und Wohnungen* (Version 2). Wiesbaden: Statistisches Bundesamt.

Thibodeau, T. G. (1989). Housing price indexes from the 1974–83 SMSA Annual Housing Surveys. *Journal of the American Real Estate and Urban Association, 17*, 100–117.

Thibodeau, T. G. (1992). *Residential Real Estate Prices from the 1974–1983 Standard Metropolitan Statistical Area American Housing Survey* (Studies in Urban and Resource Economics). Mount Pleasant, MI: Blackstone.

Thibodeau, T. G. (1996). House price indices from the 1984–1992 MSA American Housing Surveys. *Journal of Housing Research, 6*(3), 439–481.

Appendix A

Figure A1. City districts of Berlin.