

# Different preferences, but the same approach: the practice of the sales comparison in the Berlin housing rental and sale market

Dongwoo Hyun<sup>1</sup> and Steffen Heinig<sup>2</sup>

<sup>1</sup>) Corresponding author, School of Urban Planning and Real Estate, Dankook University, South Korea, d.hyun@dankook.ac.kr

<sup>2</sup>) Co-author, Department of the Built Environment, Liverpool John Moores University, UK, s.heinig@ljmu.ac.uk

## **Abstract**

While determining the value of various housing attributes can vary between sales and rental markets, market participants may use the sales comparison approach as a common strategy. Due to the intransparent and illiquid nature of the housing market, the method provides a simple way for market participants in the two different markets to agree on prices. We show that a strong dependence tendency of market participants on market price information is commonly observed in both the sales and the rental market. Using spatial econometrics for 17,500 apartment rents and sales data respectively in Berlin, Germany, we find that the dependence tendency in the rental market is significant, and it is comparable with that in the sales market. In particular, the results are robust within the restricted market setting of increased regulation for rent. These findings are robust over various spatial specifications.

Keywords: House price, Sales price and rent, Apartment, Hedonic price model, Spatial dependence

## 1. Introduction

While house prices and rents are fundamentally linked, evaluating the two may need different considerations of utility-bearing characteristics of housing. Homeowners and renters have other preferences in housing attributes, as found in empirical studies (e.g., Hanink et al. 2012). For example, homeowners tend to pay little attention to home-accommodated facilities, but renters are willing to make more positive payments (Sirmans et al. 1989). While determining the value of various housing attributes can vary significantly between sales and rental markets, we can find a similar pattern in the transaction price determination process. Due to the intransparent and illiquid nature of housing transactions, it is not a trivial task for general market participants to evaluate a property given a bundle of characteristics appropriately and then reach an agreement upon a transaction price level (Hyun and Milcheva 2018). Therefore, the most practical way for the housing transaction amateurs with little experience may use transaction price history in the same local housing market, and benchmark prices of comparable properties recently transacted in the market. In other words, they may adopt a practice that resembles the sales comparison approach, which is widely employed in real estate appraisals.

This paper would like to examine if the practice is commonly found in the sales and rental markets using spatial econometrics. We compare spatial dependence in sales prices and rents in the same housing market. The spatial relation measures a correlation between a given house price and a spatially weighted average price of its nearby houses. The basic idea embedded in the method is that the transaction price history of spatially and temporally neighbouring properties, which share locational characteristics and market conditions, has a crucial role in housing market participants setting their transaction price. The effect is based on a distance-decay function that closer ones have more impact than further ones (Tobler 1970). In this regard, the spatial application closely resembles the practice of comparable sales approaches employed in real estate appraisals that housing sellers and buyers are influenced by other transactions that provide information on a price determination mechanism with non-identical bundles of housing attributes in the market (Can and Megbolugbe 1997; Small and Steimetz 2012; Thanos et al. 2016).

The spatial application has been widely applied in a hedonic framework in recent years, but the main focus is on econometric concerns in most empirical research; for example, model misspecification and inefficiency by ordinary least square (OLS) estimation due to omitted (locational and spatial related-) variables and (spatial) autocorrelation. Evidence is available in empirical research for the outperformance of the spatial hedonic modelling over non-spatial one, with improved statistics such as Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC) and log-likelihood (e.g., Kim et al. 2003; Brasington and Hite 2005; Militino et al. 2004; Gelfand et al. 2004; Sun et al. 2005; Dubé and Legros 2014). Other

studies have focused on empirical econometric issues, such as spatial processes, forms of spatial structures, measurement unit constraints, and which distance best fits the data. However, the estimated spatial coefficient *per se* has been paid little attention.

Spatial dependence on house prices can be explained by sharing locational attributes and spatial externalities in locational similarity (Can and Megbolugbe 1997). In addition to such fundamentals, a few tries have examined the spatial effects as the effects of the behavioural tendency of housing sellers and buyers (e.g., Thanos et al. 2016; Hyun and Milcheva 2018). In line with the recent findings, we estimate spatial dependence as a proxy for the market participants' dependence tendency in the transaction price determination process. A symmetric effect of spatial dependence between sales prices and rents may suggest the practice as a common approach in the two different markets. In contrast, an asymmetric effect may suggest the spatial effect indicating different valuing on (omitted) locational factors in the markets.

The empirical analysis is conducted on the Berlin housing market, Germany. Overall the housing market can be described as a renting market with an ownership rate below the level of other European countries<sup>1</sup>. Due to its historical development, Berlin has been exempt from increased homeownership rates, as more multifamily houses were constructed to support a functioning housing market. The German housing market also provides a unique setting for leases due to various rent increase regulations. According to the civil code of Germany (Bürgerliches Gesetzbuch, BGB)<sup>2</sup>, the comparative rent common for the location (CRCL) can be demanded by the landlord only if the current rent has been unchanged for the last 15 months. The law further describes that the rent can only be raised by 20% within three years; however, only until the level of the CRCL at that time. The latter is known as the cap, which is focused on our empirical analysis. This essentially leads to the fact that rents of existing contracts lag behind the current market developments since new leases can be arranged for rents of up to 10% above the CRCL<sup>3</sup>.

The regulation may put a dominant role of the current contracted price level on a new or renewed price over other impacts such as market trends and landlords' strategy for a transaction. However, the restricted market condition may not influence the fundamental dependency tendency of landlords and tenants; transaction prices of comparable properties are still the most accessible and practical reference for them. Therefore, housing leases are still likely to occur at a price similar to nearby equivalent properties.

---

<sup>1</sup> According to the Federal Office of Statistics, the national homeownership rate is 45.6% in 2020, in comparison to 64.4% of the UK in 2014 and 65.0% of France in 2014.

<sup>2</sup> The civil code describes the rent increase in paragraphs 557 to 560 BGB. Next to step (paragraph 557a BGB) and index (paragraph 557b BGB), the most common form of rent increase is regulated in paragraph 558 BGB. There has been an additional cap on rent increases, known as the "Mietpreisbremse", in 2016 and the rent cap act from 2020 in Berlin, but the dataset of this study pre-dates the two most recent legislations.

<sup>3</sup> There are some exceptions, depending on the year of the construction of the building.

That means there may be a spatial dependence between a given rent (possibly controlled by the regulation) and nearby rents (possibly controlled by the regulation). As evidence for the argument, we can observe comparable spatial dependence in rents with that in sales prices not controlled by the increased regulation.

We employ a spatial autoregressive model to investigate spatial dependence in housing rents and sales prices in the Berlin apartment market. Using a dataset of 17,500 observations on apartment leases and sales between 2011 and 2015 in the market, respectively, the empirical results show significant and relatively symmetric effects of spatial dependence in rents and sales prices, implying that the sales comparison approach would be a common practice for market participants in both markets. The results are robust over various temporal specifications, providing evidence that the rent increase regulations would not directly impact the dependence tendency of market participants in the rental market. As well, the results are robust over various spatial specifications.

The remainder of this paper is organised as follows. In the next section, we review the related literature, followed by a description of the empirical strategy and methodology applied. Section 4 offers background information about the Berlin housing market and its policies. Section 5 describes the data in more detail. Finally, we present our empirical results and conclude with a summary of our key findings.

## **2. Literature review**

Empirical studies on housing values are extensively conducted in a hedonic framework, and many papers find various housing value determinants concerning property-specific characteristics, locational and neighbourhood attributes. Within the same empirical framework, the estimation of which factors impact shows different patterns according to a transaction type (i.e., sale or lease), implying other preferences between housing owners and renters. For example, while proximity to stations is usually positively correlated to both the sales price and rent (John 1996; McMillen and McDonald 2004; Gibbons and Machin 2005), housing equipment and facilities are factors that renters are more willing to make positive payments for (Sirmans et al. 1989). There has been less attention on housing rents than sales prices in empirical research. Academic interest in rents is limited to house price to rent ratios (e.g., Hill and Syed 2016) and rent controls (e.g., Glaeser and Luttmer 2003). In contrast to the sales market, market participants' behavioural tendency is seldomly examined in the housing rental market.

Some studies compare the renting market with the sales market in the Berlin housing market. Kholodilin and Mense (2012), which use a similar dataset to the one used in this study for a shorter time frame, find that while rents and sales prices are influenced differently by some property-specific

characteristics such as the flat's area, floor level and type of flat, they show similar patterns in spatial and temporal fixed effects respectively. Kholodili et al. (2017) compare the capitalisation of energy efficiency in selling prices and rents. The results suggest that the energy efficiency is well capitalised in both selling prices and rents for apartments but, the value of future energy cost savings tends to be more reflected in selling prices than rents. Hahn et al. (2022) investigate the short-term effects of the rent freeze policy on the supply-side of the market. Applying the difference-in-difference, the empirical results suggest that the policy drops advertised rents significantly. The market also experiences a significant drop in advertised properties for rent. Using a hedonic price model, empirical studies on the Berlin rental market mainly focus on how property-associated factors affect rents. This study, however, focuses on the behavioural tendency of renters in transactions, controlling for other hedonic variables.

Meanwhile, the application of spatial econometrics into housing studies has been widely examined in recent years and dominantly on sales data. The main goal of spatial application in early studies is to handle omitted variable problems in a hedonic framework. As it is almost impossible to control location-related factors thoroughly, which should be considered in housing values, a correlation among properties spatially located nearby is also included in a regression equation as an alternative way. Early studies show that the additional inclusion of a spatially lagged variable reduces the impact of locational and neighbourhood related variables in a hedonic framework (Dubin and Sung 1990; Can 1992). Also, including the spatial variable instead of multiple locational variables yields similar regression results to a non-spatial model including multiple location variables.

Following the early studies, there have been several attempts to account for spatial dependence for various property types, markets and sample sizes. A common finding is a significant spatial dependence in housing data. For example, spatial dependence is significantly measured in a single-family housing market from 577 observations in Columbus, Ohio, US (Can 1990); from 27,440 observations in metropolitan areas in Ohio, US (Brasington 1999); from 1,377 observations in Stockholm, Sweden (Wilhelmsson 2002); from 607 observations in Seoul, South Korea (Kim et al. 2003); from 44,255 observations in metropolitan areas in Ohio, US (Brasington and Hite 2005). Militino et al. (2004) also find the spatial relation in an apartment price from 293 observations in Pamplona, Spain.

Unidirectional causality in the temporal dimension in the spatial relation in housing data has been another crucial issue; for housing data pooled over time, past transactions can impact a given current one, but non-exited future transactions cannot. Various spatio-temporal models have been applied to handle the problem; Pace et al. (1998) introduce the filtering process in an attempt to consider spatial, temporal, and spatio-temporal dependence in a hedonic equation respectively; Sun et al. (2005) develop a model based

on a Bayesian approach; Smith and Wu (2009) and Dubé and Legros (2014) includes a spatial-temporal variable in a regression equation using the Hadamard product between an independent spatial and temporal weight matrix.

The academic focus has recently moved on the spatial coefficient estimated per se by a few empirical pieces of research. These studies examine spatial dependence on house prices as a behavioural tendency of housing sellers and buyers. Thanos et al. (2016) investigate the effects of information decay from events/ sales, which is discounted according to proximity in time and space on the market participants' strategic behaviours. They measure spatial dependence in asking prices of houses not yet sold as instruments for sellers' future market expectations. Another spatial application within a tinier timeframe in their study encapsulates sellers, buyers and intermediaries' actions into spatial correlation in asking prices, such as searching, competing and bargaining before reaching a transaction. Hyun and Milcheva (2018) propose market participants' behavioural tendencies, such as sellers' loss aversion or herding of buyers, as the leading cause of spatial dependence in apartment transaction prices by showing a higher spatial spillover effect for a housing boom than a bust period.

Overall, empirical studies present evidence on the diagnosis of spatial dependence in house prices and the outperformance of a spatial model over a non-spatial one supported by improved statistics such as AIC and BIC. Also, a few studies that examined the estimated spatial coefficients per se only focus on the spatial effects in sales markets. This study would like to provide additional evidence on the recent try to consider the spatial dependence on housing values as a market participant's behavioural tendency. In particular, this study examines if the behavioural tendency is observed in the rental market and how comparable it is to the sales market.

### **3. Empirical methodology**

A hedonic price model is basically used to estimate house prices and rents in this study. The approach values goods for utility-bearing attributes or characteristics (Lancaster, 1966; Rosen, 1974). The consumer theory of Lancaster (1966) suggests that consumer choice arises from specific characteristics of goods, not goods per se, as the attributes give rise to utility. Any composited goods can be defined in their characteristics, and each attribute commands a price. Hence a price for a composite good can be estimated as a sum of its attributes' prices. In other words, hedonic prices are defined as the implicit prices of each characteristic (Rosen, 1974). In this framework, house prices and rents are estimated by property-specific structural attributes and attributes reflecting the effects of location and neighbourhood. It is assumed that

the attractiveness of the residential real estate commodities is fully capitalised into prices. As well, macroeconomic shocks and time-invariant location characteristics significantly impact housing values, and these are usually controlled for by adding time and location dummies, respectively. A baseline hedonic price model (BHM) then takes the following form:

$$P = \alpha + \beta H + \gamma T + \delta L + \varepsilon \quad (1)$$

where  $P$  is a vector of house prices.  $H$  is a matrix of variables accounting for the property specific and locational characteristics.  $T$  is a matrix of time dummy variables, indicating time fixed effects to control for market conditions that are common to the study area and temporal heterogeneity (Wooldridge 2010).  $L$  is a matrix of location dummy variables, indicating locational fixed effects to control for regional heterogeneity.  $\varepsilon$  is a vector containing error terms which are independent and identically distributed with a zero mean and a variance  $\sigma^2$ .  $\beta, \gamma$  and  $\delta$  are coefficients associated with the variables  $H, T$  and  $L$  respectively. A semi-logarithmic specification is commonly chosen in hedonic house price models. It allows for nonlinearity and has a simple and intuitive interpretation; the estimated coefficients suggest the percentage impact of changes in attribute values on house prices (Malpezzi 2003; Sirmans et al. 2005).

In practice, however, most general housing sellers and buyers who have little experience in housing transactions may not quickly assess the value of such property fundamentals. Alternatively, more advanced market participants are met by those with certain price expectations based on this relatively simple approach. Therefore, the former tend to rely on market information such as the transaction price history of comparable properties in a local market. They see how bundles of housing characteristics are capitalised in the market from other transaction prices and use them as a benchmark to agree upon a transaction price. We try to estimate the practice that resembles the sale comparison approach by measuring a spatial correlation between a given house price and its nearby properties' values which are likely to be used as the benchmark. To consider the spatial relation, we extend the BHM by a spatially lagged house price to form the spatial autoregressive model (SAR) as follows:

$$P = \alpha + \beta H + \gamma T + \delta L + \rho SP + \varepsilon \quad (2)$$

where  $S$  is an exogenous non-negative matrix and takes the following form:

$$S = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ S_{21} & 0 & 0 & \dots & 0 \\ S_{31} & S_{32} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ S_{n1} & S_{n2} & S_{n3} & \dots & 0 \end{pmatrix} \quad (3)$$

Each element in the matrix corresponds to a distance between each pair of properties in the data. The distance measurement is based on the Euclidean distance, using each property's longitude and latitude. The distance in the matrix is defined as an inverse function to ensure the distance-decay relation that other transactions influence housing values in spatial proximity but are more influenced by closer properties. A non-negative value is given only to elements which define the spatial relationship within a critical distance threshold, beyond which properties are assumed to have no direct connection. For their transactions, market participants would not directly refer to properties and their values located over the distance threshold. Each element also reflects temporal relations between the property transactions, and the inverse function and critical distance threshold are also applied to the temporal proximity. Therefore, more recently transacted properties have more significant impacts on a given property than older ones. The relations are not defined for the pairs in which a temporal gap is over a particular range.

One of the crucial points in the application is that the spatial relations should be measured in unidirectional temporal causality<sup>4</sup>; market participants are exposed to nearby property transactions which have already finished but not to ones which will be done in the future. To ensure such nature, all the elements are chronically ordered and given zero values to the temporally violated spatial relations, forming the matrix with non-zero values only for the lower triangular elements. The critical distance threshold rests with the researcher to choose, usually following the goodness-of-fit criteria. We test various specifications to find the most fitted one, but the focus is on the specification of 36 months for the temporal threshold to reflect the rent increase regulations in the Berlin housing market. Finally, the matrix is normalised so that each row sums to one to avoid over-connection and spatial unit root issues (Dubé and Legros 2014).

Multiplying the matrix,  $S$ , by the vector of house prices,  $P$ , then provides a vector of weighted average transaction prices of spatiotemporally nearby properties. A coefficient of the interaction term,  $\rho$ , represents a dynamic effect of changes in transaction prices in spatial and temporal proximity. Assuming

---

<sup>4</sup> Empirical studies report that violation of unidirectional temporal causality in the spatial effects for pooled over time housing data result in biased coefficients and imprecise variance estimation (see Dubé and Legros 2014; Thanos et al. 2016; Hyun and Milcheva 2018).

that housing sellers and buyers hugely depend on recent outcomes for other nearby transactions to determine their transaction prices,  $\rho$  could be considered as the practice of the sales comparison approach of market participants. Therefore, the main interest in this study is if we can find a significant spatial coefficient in rents and how comparable the coefficient is with that in sales prices. The main aim here is to demonstrate it but not to identify the channel that goes beyond the scope of this research.

#### **4. Market description and data**

##### **4.1. The Berlin housing market**

The Berlin housing market is second to none on a national and European level. Due to its historical development after the Second World War, when the city was divided, the housing market was mainly intended as a renting market. Unlike other European countries such as France or the United Kingdom, Germany is characterised by a polycentric outlay. Therefore, it is not surprising that even after 30 years of the fall of the Berlin Wall, Berlin, as the capital, does not offer the highest rent or sale prices. Cities such as Munich or Hamburg have higher rent and price levels in comparison. The German housing market is strongly influenced by national and federal legislation. Since the homeownership rate is comparably low in Germany, the civil code has a series of tenant-friendly laws which protect the tenant during the leasehold. The most common intervention in the local housing markets is observed during the rent increase demand from the landlord, following the legislation in the civil code in paragraph 558 BGB. Initial rents and rents for existing leases have to follow the CRCL. It is stated that rents for new leases cannot exceed 10% of the CRCL.

The German law controls the market through a cap for existing leases, which regulates the market in two ways. Firstly, it is only possible to increase the rent by a maximum of 15% within three years to the level of the CRCL. Secondly, the rent increase demand has to be justified by providing evidence. This evidence can either be a rent value in the rent index, a value from a rent database, a valuation of a chartered surveyor or proof through three suitable rent comparables (paragraph 558a BGB). Even if the demand for a rent increase is based on a different proof (paragraph 558a (3) BGB) and an officially qualified rent index exists in the location, it is mandatory to accompany the claim. Significantly, the qualified rent index (paragraph 558d BGB) is of utmost importance, as it is not clearly defined how stakeholders of tenants and landlords have to derive from the rent index. The law states that it has been created with scientific standards (paragraph 558d (1) BGB). It is further defined that the index needs to be updated every two and regenerated every four years (paragraph 558d (2) BGB). This explains the ever-existing difference between the current

market level and the rent level of the CRCL.

Selling apartments, on the other hand, is less regulated. Prices can be derived between seller and purchaser at their discretion. The government controls the price only through taxation, where the revenue office might question the agreed price if it diverges too much from the property's expected value. This can be described as a relatively soft market control and probably has its origin because the tax authority<sup>5</sup> is interested in earning its fair share.

#### 4.2. Data

This study uses two datasets of asking rents and prices from various online trading platforms. The datasets span over five years, from 2011 to 2015, on a monthly base. Jones Lang LaSalle GmbH provides the data. As contractual data is hard to come by and actual sell prices are not publicly available, we need to rely on the accuracy of this data. Especially, transactional prices of apartments and houses are kept confidential, as they are the base for taxation. The data is collected and managed by the committee of valuation experts. To our knowledge, there is no such dataset available in the market. It can be assumed that there might be a difference between the asking and the contractual rent or price. However, since the data represents a period that sees a price increase in both market segments, prices are unlikely to differ much since tenants and buyers do not have much bargaining power due to a limited supply within the market.

As the data comes from online trading platforms such as Immobilienscout24, ImmoWeb or ImmoWelt, each observation carries the information about *the location of the apartment (District, Zip code, Street), when the property was advertised, the age of the building, the number of rooms, the size of the apartment, in which floor it is located, rent and running costs, whether the apartment is refurbished and when, if the property is barrier-free and if the building construction is publicly subsidised.*

We further have the information about the potential rental income and the sales price per sqm for the sales dataset. Also, we are able to extract further information from the provided description of each advertisement. A total of 47 different possible features are mentioned in these descriptions. However, many of those are only mentioned a couple of times. Therefore, we decide to narrow the number of variables down to those mentioned in Table 1. Descriptive statistics of the continuous variables are given in Table 2.

---

<sup>5</sup> Currently, Berlin's real estate transfer tax is 6% of the purchase price. Another tax that may influence the housing transaction's selling process is the so-called speculative tax, which is applied if a property is sold with a profit during 10 years. However, the tax depends on the personal tax bracket and is not generalizable.

**Table 1 Variable description**

<i>Dependent Variable</i>	
(ln) Price	(the natural logarithm of) price
<i>Independent Variable</i>	
Size	Size of an apartment unit in square metres
Rooms	Number of rooms
Floor	Floor level on which the apartment unit is located within the building
Age	(The year the price was given) – (the year of construction of the apartment building)
Age sq.	Square of the age
Parking	= 1 if either indoor or outdoor parking is assigned to the apartment, = 0 otherwise
C_heating	= 1 if central heating, = 0 otherwise
U_heating	= 1 if under floor heating, = 0 otherwise
SC_heating	= 1 if self-contained central heating, = 0 otherwise
S_heating	= 1 if stove heating, = 0 otherwise
G_heating	= 1 if gas heating, = 0 otherwise
Furnished	= 1 if furnished, = 0 otherwise
F_kitchen	= 1 if fitted kitchen, = 0 otherwise
P_floor	= 1 if parquet floor, = 0 otherwise
Attic	= 1 if attic, = 0 otherwise
Cellar	= 1 if cellar, = 0 otherwise
Conservatory	= 1 if conservatory, = 0 otherwise
Gallery	= 1 if gallery, = 0 otherwise
Storeroom	= 1 if storeroom, = 0 otherwise
F_basement	= 1 if full basement, = 0 otherwise
Garden	= 1 if garden, = 0 otherwise
Lift	= 1 if lift, = 0 otherwise
Alarm	= 1 if alarm system, = 0 otherwise
Balc_log	= 1 if either balcony or loggia, = 0 otherwise
Terrace	= 1 if terrace, = 0 otherwise
N_build	= 1 if first time use or new build, = 0 otherwise
Maisonette	= 1 if maisonette, = 0 otherwise
Villa	= 1 if the apartment is located in a villa, = 0 otherwise
Marble	= 1 if marble floor / , = 0 otherwise
Courtyard	= 1 if courtyard available, = 0 otherwise
Chimney	= 1 if chimney, = 0 otherwise
Wellness	= 1 if the apartment has either a sauna, a pool or a whirlpool, = 0 otherwise
M_bath	= 1 if marble bathroom, = 0 otherwise
R_occupied	= 1 if renter occupied (previous tenant), = 0 otherwise
Southward	= 1 if southward looking, = 0 otherwise
CBD	Distance to the closest central business district (m)
Stations	Number of stations within 300m
Location	= 1 if the apartment in the respective district, = 0 otherwise (12 districts in Berlin)
Year	= 1 if the price is given in the year, = 0 otherwise

**Table 2 Descriptive statistics of the continuous variables**

Variable	Rent (17,500 observations)				Sale (17,500 observations)			
	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
Price	602	376	85	5,800	263,227	253,551	9,000	4,500,000
(ln) Price	6.27	0.48	4.44	8.67	12.16	0.79	9.105	15.32
Size	75.66	31.21	12	410	86.66	44.47	16	555
Rooms	2.51	1.02	0	9	2.73	1.23	0	25
Floor	2.97	2.26	1	30	2.77	1.96	1	26
Age	64.09	39.85	0	322	64.09	39.85	0	322
CBD	7987.64	4321.67	336	23912	6882.72	4134.62	76	24148
Stations	0.24	0.45	0	2	0.27	0.48	0	2

Note: Price in the rent dataset is given in EUR per month. For the Sale data set, it is given as EUR per sqm. Size is measured in sqm. CBD indicates the distance to one of the three CBDs' in Berlin, either Alexanderplatz, Zoologischer Garten or Potsdamer Platz. All three variables rooms, floors and stations provide a total number.

From the original dataset representing the fully online transaction universe of the Berlin housing market for over 50,000 observations, we select 17,500 observations randomly for each transaction type, as it is the maximum size we can apply for the spatial application. Our dataset is similar to studies such as Kholodilin and Mense (2012). They use online price data for their hedonic price model between June 2011 and February 2012. While our dataset covers a longer timeframe, the descriptive statistics are similar. This underlines that our findings are representative of the market developments at the time. We would like to point out further that the German market was increasingly influenced by external factors such as the refugee crisis and further price increases towards the end of our analysis horizon. Therefore, our dataset might be able to capture the market behaviour in a seemingly functioning market.

## 5. Empirical results

The SAR is operationalised with controls for all the 54 variables for rent and sale<sup>6</sup>. Table 3 represents the summary results of the SAR for statistically significant variables at a 0.01 level only (see Table A1 in appendix for the complete results). For comparison purposes, the summary results of the BHM's are also shown. For the SAR models, observations in the first period are dropped to avoid bias due to non-existing temporally nearby transactions in data. Therefore, asymmetric numbers of observations between the BHM and the SAR are reported. The critical distance threshold for the temporal relations is 36 months to reflect the rent increase regulation in the market and for the spatial relations is 2 km as the specification provides the best performance for comparison between rent and sale markets. The fit to the data is reasonable and improved by capturing spatial dependence for both rent and sale models compared to the BHM. However, comparable results are observed in most variables in sign, value and significance between the two models, suggesting that the spatial effects captured are highly likely to be related to locational and neighbourhood effects and not to property-specific physical attributes. Most coefficients show the expected signs.

---

<sup>6</sup> Except for the renter-occupied variable, which is used only for sale.

**Table 3 Summary results for the BHM and the SAR**

Variables	Rent		VARIABLES	Sale	
	BHM	SAR		BHM	SAR
Size	0.010*** (0.000)	0.010*** (0.000)	Size	0.011*** (0.000)	0.010*** (0.000)
Rooms	0.023*** (0.003)	0.027*** (0.003)	Rooms	0.033*** (0.004)	0.039*** (0.004)
Floor	-0.006*** (0.001)	-0.005*** (0.001)	Floor	0.011*** (0.001)	0.012*** (0.001)
Age	-0.006*** (0.000)	-0.005*** (0.000)	Parking	0.072*** (0.007)	0.058*** (0.006)
Age sq.	0.000*** (0.000)	0.000*** (0.000)	C_heating	-0.086*** (0.007)	-0.080*** (0.006)
parking	0.042*** (0.004)	0.039*** (0.004)	U_heating	0.066*** (0.009)	0.055*** (0.008)
C_heating	-0.046*** (0.004)	-0.045*** (0.004)	SC_heating	-0.076*** (0.011)	-0.072*** (0.010)
U_heating	0.117*** (0.009)	0.103*** (0.009)	F_kitchen	0.047*** (0.007)	0.047*** (0.006)
SC_heating	-0.051*** (0.006)	-0.049*** (0.006)	P_floor	0.038*** (0.008)	0.032*** (0.007)
Furnished	0.089*** (0.020)	0.087*** (0.019)	F_basement	-0.721*** (0.124)	-0.627*** (0.125)
F_kitchen	0.057*** (0.004)	0.053*** (0.004)	Garden	0.026*** (0.007)	0.018*** (0.006)
P_floor	0.074*** (0.006)	0.062*** (0.006)	Lift	-0.022*** (0.007)	-0.018*** (0.007)
Attic	0.022*** (0.005)	0.022*** (0.004)	Terrace	0.022** (0.010)	0.030*** (0.009)
Cellar	-0.010*** (0.004)	-0.013*** (0.004)	Villa	0.165*** (0.019)	0.128*** (0.018)
Garden	0.047*** (0.005)	0.046*** (0.005)	Chimney	0.053*** (0.020)	0.056*** (0.019)
Balc_log	0.022*** (0.006)	0.019*** (0.005)	Southward	-0.113*** (0.040)	-0.107*** (0.038)
Maisonette	0.044*** (0.012)	0.038*** (0.012)	N_build	0.213*** (0.007)	0.179*** (0.007)
Villa	0.081*** (0.016)	0.051*** (0.016)	R_occupied	-0.205*** (0.007)	-0.185*** (0.006)
CBD	-0.000*** (0.000)	-0.000*** (0.000)	CBD	-0.000*** (0.000)	-0.000*** (0.000)
Stations	0.016*** (0.004)	0.011*** (0.004)	Stations	0.064*** (0.006)	0.052*** (0.006)
Rho( $\rho$ )		0.176*** (0.006)	Rho( $\rho$ )		0.187*** (0.004)
Constant	5.699*** (0.011)	4.564*** (0.037)	Constant	11.075*** (0.019)	8.850*** (0.051)
LL		3735.77	LL		-5385.15
AIC		-7365.55	AIC		10878.30
BIC		-30160.20	BIC		-29875.00
(Adj)R <sup>2</sup>	0.825	0.835	(Adj)R <sup>2</sup>	0.805	0.826
Observations	17,500	17,010	Observations	17,500	17,325

Standard errors in parentheses \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Note: The results are based on Equation (1) for the BHM and Equation (2) for the SAR with controls for all the 54 variables including location and time dummy variables respectively, but represented with only statistically significant variables at a 0.01 level for the SAR (See Appendix for the entire results).

### 5.1. The berlin housing market

The results indicate the same signs in some variables but suggest different preferences in housing attributes between tenants and homeowners. Both rents and sales prices of apartments tend to increase with the size of the unit, the number of rooms and the existence of parking spaces. As heating carries a significant part of the running costs, three out of five heating types show a highly significant coefficient in the dataset. Underfloor heating is seen as an improvement to the living atmosphere since no heaters are attached to walls, effectively blocking useable space. Our results indicate that apartments with these features have a higher price. While our dataset, unfortunately, does not give additional information on the sustainability features of the units, one could use underfloor heating as a proxy to measure this. At the same time, apartments that offer central heating are sold or let with a discount as indicated through the negative sign in front of the coefficients. This result is slightly surprising, as central heating usually offers the cheapest way for tenants to heat their apartments. Therefore, the negative sign is not expected. Other studies, such as Kholodilin and Mense (2012), mention the availability of energy performance certificates in their dataset, but these variables remain insignificant.

Moving on to the fitted kitchen and furnished apartments, the results are within the expectation. Two decades ago, a fitted kitchen was seen as a luxury feature of an apartment. The law expected the landlord to provide a sink and a stove in the kitchen. This has changed, and fitted kitchens are becoming more common in the housing market. A fitted kitchen is usually included in the selling price for apartments offered for sale. However, if the apartment is sold as part of a larger development, fitted kitchens can become cheaper due to the discount developers can achieve. Therefore, the high significance of the variable is not surprising and confirms the findings of Kholodilin and Mense (2012). Overall, it is fair to say that landlords are more likely to provide a fitted kitchen as this allows to increase the rent. Looking at the second variable, whether the apartment is furnished or not, the results are within our expectations. People who buy an apartment for their private use are more likely to buy an empty apartment. However, the letting market has seen an increase in temporary living, alongside AirBnB lettings in the last years. Therefore, a furnished apartment is desired by some people, and it will increase the rent. Looking at some more general amenities and features, parquet flooring, the existence of a garden, and if it is located in a villa lead to an increase in the monetary value. The presence of a cellar, on the other hand, shows a mixed result in both markets. This is surprising, as a basement offers additional space to store items. However, the rent model suggests that a cellar harms the rent.

Housing owners and tenants might have different preferences on property-specific characteristics, and the coefficients can further confirm that for various variables. The presented difference between highly

significant variables between the two datasets is interesting. For instance, what is surprising is the fact that the variable for the existence of a lift in the sales data carries a negative sign. A lift indicates that the apartment might be located on a higher floor, allowing easier access to the apartment in general. Also, a full basement or the apartment's orientation harms the price. While the second one is surprising, the first might be a feature that would play a more significant role in purchasing a single-family house.

The last two variables that differ from the rent data set are whether the apartment is newly built and currently let. A new property is traded at a premium, while an occupied apartment is sold with a discount. The reason for this is that the apartment has been blocked for some time by the new owner, meaning she cannot move in right away. Secondly, the apartment is currently disconnected from the market rent. So, if the tenant does live in the dwelling for some time, recent rent market rent developments have not been realised due to German law. Increasing the rent to the market level plus 10% can only be realised by vacating it. However, the German law is tenant-friendly and protects the tenant here, as described earlier. Therefore, a property that is currently let will be traded at a discount.

Different to the sales model, the existence of an attic, whether the apartment has a balcony or loggia and whether it is a maisonette or not have become highly significant variables. All three results confirm our expectation that all three features should increase the price of the apartment. For the accessibility to public transport, we measure the distance to the nearest subway station (either S-Bahn or U-Bahn) and the number of stations within 300, 400, 500 and 1,000 metres. While the results are insignificant with the former measurement, data are best fitted with the number of stations within 300m for both models. More stations within 300m positively influence sales prices and rents.

We further test the distance to the CBD with three different areas, Alexander Platz, Zoologischer Garten and Potsdamer Platz, which are generally considered central business districts in Berlin. While the variable is highly significant in both markets, the coefficient does not influence the price much. A reason for this finding can be found in the structure and size of the city. While the CBD is an economical and or touristic hub, many inhabitants of Berlin prefer to spend their time in their local neighbourhoods. Each district and borough have some centres offering amenities which make it obsolete to travel far. Therefore, the direct distance to the CBDs might play a minor role.

Figure 1 illustrates locational fixed effects to control for regional heterogeneity at the district level. All the coefficients are intuitively acceptable. For example, Charlottenburg-Wilmersdorf, used as a reference, is considered the most expensive district due to the wide range of Wilhelminian style buildings and villas. Within the wider district, there is a borough called Grunewald, which is famous for its old mansions, which are mainly located in the forest and close to the river, Havel. The figure further illustrates

Steglitz-Zehlendorf, Friedrichshain-Kreuzberg and Mitte come 2nd, 3rd and 4th, representing our expectations. Especially, Mitte and Friedrichshain-Kreuzberg have become popular districts in the last years. While the gentrification in some parts of the city, such as Prenzlauer Berg (borough within Pankow – 5<sup>th</sup> place), has reached its peak, these districts still offer sites for development. The other aspect of Mitte being in the top ranks is the closeness to the government district, which makes it a popular location to live in for people working for and with the government.

Interestingly, the model confirms that Marzahn-Hellersdorf is, in both terms, a district that is at the other end of the spectrum. The model shows that apartments are traded with the most significant discount for rents and sale prices. Reasons are the distance to the various city centres and many people's perception. The district has a higher number of low-income households and a matching clientele.

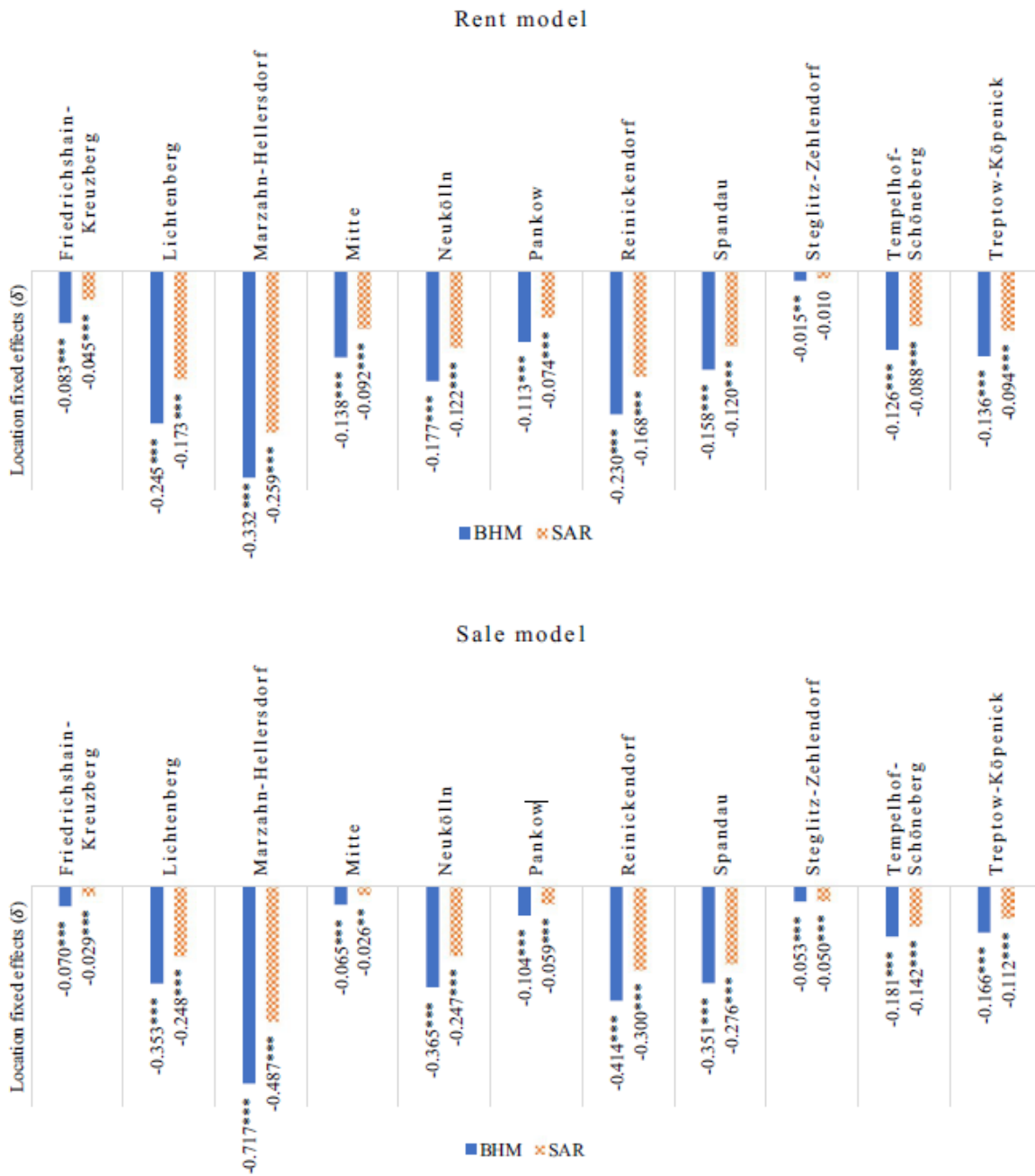
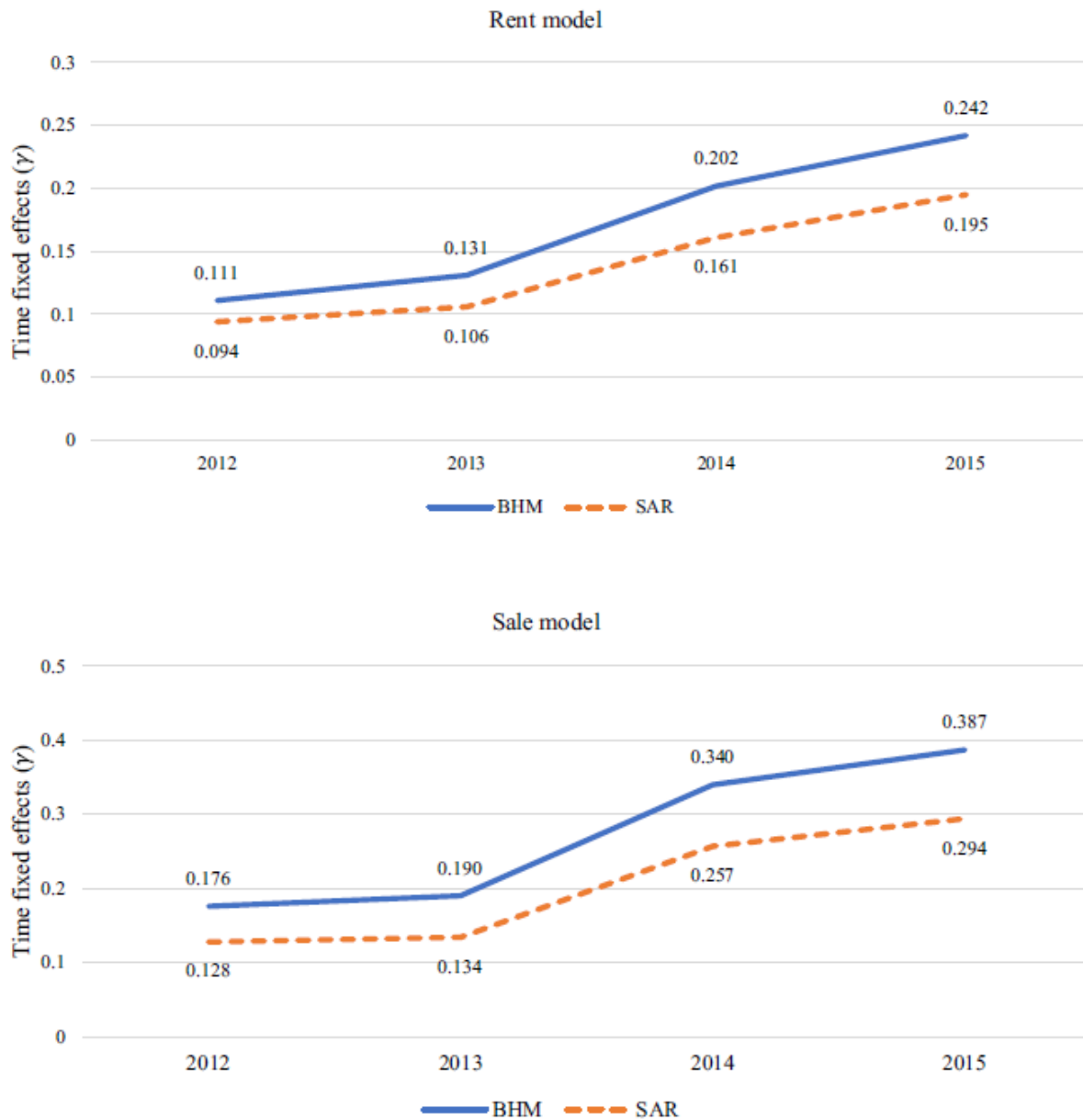


Fig. 1 Location fixed effects for the BHM and the SAR. Note: Charlottenburg-Wilmersdorf has been used as the location reference. \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$

Figure 2 depicts trends in the sales price and rent levels with controls for all other variables. The data suggest that the Berlin housing market experienced a gradual increase in sales price and rent during the study period. Although gaps are observed in coefficients for the location and time fixed effects between sale and rent, their comparable patterns imply that market participants might have similar preferences for neighbourhood factors and respond to macroeconomic conditions similarly to set their asking prices for sale and rent.



**Fig. 2** Time fixed effects for the BHM and the SAR. Note: The year 2011 has been used as the reference year. All the coefficients are statistically significant at the 0.01 level

## 5.2. Spatial dependence

The SAR for rent yields a significant spatial coefficient ( $\rho$ ) of 0.176. The statistically and economically significant coefficient suggests that tenants and landlords rely on price information of recently transacted nearby properties to agree upon their lease transaction. In other words, the sales comparison approach is also applied in the rental market. 0.176 of the coefficient means that, for example, if there is a 1.00 € increase in rents of properties located in a 2km radius and transacted in the past 36 months on average, it would cause a rise of 0.176 € in a given asking rent if the effect is purely pecuniary (Thanos et al., 2016). Put another way, that amount would be driven by the prices of nearby properties, which the market participants consider as benchmarks for their asking prices. We can also observe a significant spatial coefficient in the SAR for sale. The value is slightly higher than rent, but the difference is marginal<sup>7</sup>. Therefore, we can conclude with the comparable spatial dependence in the rent and sale market that the sales comparison approach is typical for market participants in rent and sales markets. However, they have different preferences in housing characteristics.

In contrast to the comparable results for variables describing property-specific characteristics between the BHM and the SAR, differences are quite noticeable in the location and time fixed effects in Figures 1 and 2, respectively. The apparent spatial heterogeneity across the 12 districts and the gradual yearly increasing trends in the Berlin housing market appear in the BHM and the SAR model. However, this pattern is found at a remarkably reduced magnitude in the SAR model in comparison. This provides additional evidence that the spatial coefficient captures a part of the location and time variation of the transaction price determination process. Based on this, we can infer that the reduced amount might account for the spatial dynamics of the nearby property prices (i.e., the practice of the sales comparison approach). The comparable values of spatial coefficient across rents and sales depict similar reduced patterns in the location and time fixed effects. This translates into the assumption that the more challenging rental market conditions have an insignificant impact on market participants' dependence tendency in price determination.

The critical distance threshold for the SAR is set for 36 months in this study to reflect the rental market regulation. As this is a market criterion, other time lengths might have been more suitable. This leads to difficulties in the spatial-hedonic modelling process, as there is no consensus on the selection in empirical research. The same accounts for the physical distance parameter. It rests with the investigator choosing which kilometre and month are considered nearby property transactions. As the goodness-of-fit criteria are usually applied in empirical studies, we test the SAR with various spatial and temporal

---

<sup>7</sup> Asymptotic t-tests reject the null hypothesis that the two spatial effects are not equal to each other at the 0.01 level.

thresholds, respectively, from 1km to 6km with a 1km interval and from 12 months to 60 months with a 12 months interval. Summary results are presented in Table 4 with values of spatial dependence ( $\rho$ ) and estimators generally used for the model selection.

**Table 4 Comparison of the SAR specifications**

	Rent					Sale				
Threshold	1km-12m	1km-24m	1km-36m	1km-48m	1km-60m	1km-12m	1km-24m	1km-36m	1km-48m	1km-60m
Rho( $\rho$ )	0.047*** (0.003)	0.086*** (0.004)	0.093*** (0.004)	0.095*** (0.004)	0.095*** (0.004)	0.063*** (0.003)	0.086*** (0.003)	0.092*** (0.003)	0.095*** (0.003)	0.094*** (0.003)
LL	3342.19	3487.52	3518.16	3528.18	3526.11	-6128.53	-5970.61	-5922.57	-5902.15	-5902.89
AIC	-6578.38	-6869.05	-6930.31	-6950.36	-6946.23	12365.07	12049.21	11953.15	11912.30	11913.78
BIC	-29373.0	-29663.7	-29724.9	-29745.0	-29740.8	-28388.3	-28704.1	-28800.2	-28841.0	-28839.5
(Adj)R <sup>2</sup>	0.827	0.830	0.830	0.830	0.830	0.811	0.814	0.815	0.816	0.816
Threshold	2km-12m	2km-24m	2km-36m	2km-48m	2km-60m	2km-12m	2km-24m	2km-36m	2km-48m	2km-60m
Rho( $\rho$ )	0.116*** (0.005)	0.168*** (0.006)	0.176*** (0.006)	0.173*** (0.005)	0.172*** (0.005)	0.160*** (0.004)	0.186*** (0.004)	0.187*** (0.004)	0.188*** (0.004)	0.187*** (0.004)
LL	3496.45	3691.12	3735.77	3732.28	3729.22	-5617.16	-5409.25	-5385.15	-5376.37	-5378.70
AIC	-6886.90	-7276.23	-7365.55	-7358.56	-7352.44	11342.31	10926.50	10878.30	10860.75	10865.40
BIC	-29681.5	-30070.8	-30160.2	-30153.2	-30147.0	-29411.0	-29826.8	-29875.0	-29892.6	-29887.9
(Adj)R <sup>2</sup>	0.830	0.834	0.835	0.835	0.834	0.822	0.826	0.826	0.827	0.827
Threshold	3km-12m	3km-24m	3km-36m	3km-48m	3km-60m	3km-12m	3km-24m	3km-36m	3km-48m	3km-60m
Rho( $\rho$ )	0.222*** (0.007)	0.249*** (0.007)	0.250*** (0.007)	Not calculable		0.224*** (0.005)	Not calculable			
LL	3722.83	3884.73	3918.41	Not calculable		-5863.96	Not calculable			
AIC	-7339.66	-7663.46	-7730.81	Not calculable		10728.30	Not calculable			
BIC	-30134.3	-30458.1	-30525.4	Not calculable		-30025.0	Not calculable			
(Adj)R <sup>2</sup>	0.834	0.837	0.838	Not calculable		0.828	Not calculable			
Threshold	4km-12m	4km-24m	4km-36m	4km-48m	4km-60m	4km-12m	4km-24m	4km-36m	4km-48m	4km-60m
Rho( $\rho$ )	0.308*** (0.009)	Not calculable				0.322*** (0.005)	Not calculable			
LL	3858.16	Not calculable				3858.16	Not calculable			
AIC	-7610.32	Not calculable				9810.15	Not calculable			
BIC	-30404.9	Not calculable				-30943.2	Not calculable			
(Adj)R <sup>2</sup>	0.837	Not calculable				0.837	Not calculable			
Threshold	5km-12m	5km-24m	5km-36m	5km-48m	5km-60m	5km-12m	5km-24m	5km-36m	5km-48m	5km-60m
Rho( $\rho$ )	0.297*** (0.009)	Not calculable				Not calculable				
LL	3778.77	Not calculable				Not calculable				
AIC	-7451.55	Not calculable				Not calculable				
BIC	-30246.2	Not calculable				Not calculable				
(Adj)R <sup>2</sup>	0.835	Not calculable				Not calculable				
Threshold	6km-12m	6km-24m	6km-36m	6km-48m	6km-60m	6km-12m	6km-24m	6km-36m	6km-48m	6km-60m
Rho( $\rho$ )	0.286*** (0.009)	Not calculable				Not calculable				
LL	3715.19	Not calculable				Not calculable				
AIC	-7324.39	Not calculable				Not calculable				
BIC	-30119.0	Not calculable				Not calculable				
(Adj)R <sup>2</sup>	0.834	Not calculable				Not calculable				

Standard errors in parentheses \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Notes: the results are based on Equation(2) with controls for all the 54 variables including location and time dummy variables`

All the specifications yield a significant value of  $\rho$  at a 0.01 level, implying that spatial dependence exists in the sale and the rent data. As we found in Table 3, values of the spatial effect are generally slightly higher for sales than for rents within the same threshold. Still, the gap is not remarkable, indicating again that market participants' dependence tendency for setting their asking price is standard practice in the housing market regardless of the transaction type. Also, the rent increase regulation impacts the rent level per se but seems to have no direct impact on the dependence tendency; we can observe no clear differences in the spatial coefficient between specifications with 36 months threshold and 24 months threshold for rent. The difference is somewhat vast compared to the specification with 12 months threshold. However, we can find similar patterns in the sales data as well. Therefore, it might be related to spatial modelling issues, not the rent increase regulation itself. One reason could be that not all landlords either use or even have the ability to increase the rents on a regular period. Maybe, because the rents might already exceed the CRCL or because the landlords do not follow an ongoing rent increase strategy.

Regarding econometric concerns in spatial modelling, we can find in Table 4 that spatial dependence tends to be more assertive with a higher value of spatial and temporal thresholds, respectively. Both lead to an overall improvement of model fit. However, beyond the threshold of 36 months, values of  $\rho$  do not increase but tend to be stable or even decrease slightly. This implies that the spatial effect's estimation does not automatically increase by setting a more comprehensive range of spatial and temporal areas as 'being nearby' in our spatial modelling. This finding aligns with Dubé and Legros (2014) concerning the spatial boundary. Their empirical results from the Paris housing market show spatial dependence in house prices increases with a spatial distance cut-off value within 3km but drops within 5km. Our results, furthermore, suggest a similar pattern in temporal effects in spatial dependence under appropriate control for unidirectional temporal dimensions in housing transactions.

Another issue in spatial modelling is the construction of the spatial weight matrix, which requires a tremendous amount of calculation. Consequently, only a limited size of data could be applied. 17,500 observations are the maximum size we could use in preliminary tests from over 50,000 observations in the raw data. However, as shown in Table 4, we could not get results with the same sample size over a particular spatial and temporal threshold as computations are not possible with a modest storage size. As mentioned above, a model selection in the spatial application is usually based on the goodness-of-fit criteria. That means that we could not confidently conclude which specification is best-fitted to our data, although the main aim of this study is not to find the best spatial specification. Such issues in econometric concerns need to be discussed further for the empirical application of spatial modelling.

## 6. Conclusion

Housing markets are intransparent and illiquid. Such nature makes it difficult for market participants with fewer transaction experiences to value houses with non-identical bundles of attributes. Practically, the most straightforward way to value housing for a transaction, in this sense, is to gather information from a real estate agency about recent transaction prices of comparable properties in the same local market. It is, in fact, the practice we can commonly observe in housing markets that prices are anchored around a reference price at a specific location. The practice would be fundamentally found in most housing market participants for a transaction, although each might have different housing and locational preferences. It would also be found in a market with restricted settings, such as a price cap.

In this light, Berlin qualifies as an exciting testing ground for the methodology applied. As the German housing market offers various safeguards for tenants, people tend to rent rather than buy properties. The current market situation gives landlords, developers, or property vendors more power when estimating prices. Limited supply and an increasing demand lead to an increase in prices. This demand is fuelled by an influx of national and international people, which, different to original inhabitants, might be able to pay more for their housing and increase prices subsequently. Another generator of demand is the low-interest rate of the European Central Bank at the time. This allows more people to buy properties, and this, therefore, adds another layer of pressure on the market, as more properties will be transformed from rent to sales. Overall, this situation leads to the fact that both potential tenants and potential property buyers have just a small chance to negotiate the price for their housing. They are, therefore, likely not to bother with the time-consuming exercise of evaluating rents or prices.

In this study, we examine the argument by comparing the rental and sales market with a measurement of spatial dependence on prices. We incorporate a spatially lagged house price as an additional variable into a hedonic price model. We used a rich dataset of asking rents and sales prices for Berlin apartments. The market is exciting as it offers an ambivalent nature to the lease and the sales side. The difference is that the rental market is heavily regulated regarding rent increases and much less controlled for setting the sales price of properties. The empirical results strongly support the arguments. We find that the spatial dependency is statistically and economically significant for rents. That implies that landlords tend to set their housing leases at prices closely related to nearby property rents controlling for property-specific attributes. Besides, the effect is quite comparable in magnitude with that in sales prices, providing evidence for the common practice of the sale comparison approach over rental and sales markets. The restricted market setting for rents does not exert influence on the practice of landlords directly as we find no noticeable differences in the estimation over various temporal specifications, including 36 months. The

results are robust over different spatial specifications as well. Meanwhile, we can confirm asymmetric preferences on property-bearing attributes over rental and sales markets with distinct estimations for other hedonic variables.

We employ spatial econometrics as an empirical strategy. The spatial application has gained popularity recently in housing studies. Still, main interests tend to be too focused on econometric concerns, so statistically, outperformance over the non-spatial model with improved information criteria such as AIC and BIC are the primary motivation to use it in a hedonic framework. However, the main finding in this study shows empirical difficulties in finding the best-fitted specification due to extremely time-consuming computation for spatial relations between each pair of properties. Even we could not analyse specifications which are expected to produce statistically more significant results due to the non-calculable sample size. Given noticeable changes in values of the spatial coefficients according to the specification, as shown in our results, it would be a crucial issue for precise spatial estimation. Therefore, it should be dealt with in further research.

Another limitation of this study is spatial dependence as a compound effect. We consider the estimated spatial coefficient as a proxy for the practice of the sales comparison approach, but the value might partly capture the effects of sharing unobserved neighbourhood characteristics between property prices. The results indicate that locational heterogeneity over 12 districts keeps the same patterns after capturing significant spatial dependence in the SAR, implying that the spatial coefficient might be explained by other effects such as market participants' dependence tendency rather than unobservable neighbourhood effects. The results are robust over various specifications, and the comparison between rental and sales markets shows consistent patterns over various specifications. However, the results would be more confident with more direct measurement, precisely separating the dependence tendency from other possible neighbourhood and locational effects.

The relevance of our dataset is another limitation. While it is unquestionable that the data is outdated and does not reflect the current market developments, it can be used as evidence of a working market. We like to point out that the Berlin housing market has seen unprecedented development in the last eight years, making the market less affordable and leading to increased social tensions. The dataset used precedes these developments and allows an uninfluenced analysis of features in a hedonic sense. Tensions caused by a supply shortage and increased demand would blur the picture.

## References

- Brasington, D. (1999). Which measures of school quality does the housing market value? *Journal of Real Estate Research*, 18(3), 395-413.
- Brasington, D., & Hite, D. (2005). Demand for environmental quality: a spatial hedonic analysis. *Regional Science and Urban Economics*, 35(1), 57-82.
- Can, A. (1990). The measurement of neighborhood dynamics in urban house prices. *Economic Geography*, 66(3), 254-272.
- Can, A. (1992). Specification and estimation of hedonic housing price models. *Regional Science and Urban Economics*, 22(3), 453-474.
- Can, A., & Megbolugbe, I. (1997). Spatial dependence and house price index construction. *Journal of Real Estate Finance and Economics*, 14(1-2), 203-222.
- Dubé, J., & Legros, D. (2014). Spatial econometrics and the hedonic pricing model: what about the temporal dimension? *Journal of Property Research*, 31(4), 333-359.
- Dubin, R. A., & Sung, C.-H. (1990). Specification of hedonic regressions: non-nested tests on measures of neighborhood quality. *Journal of Urban Economics*, 27(1), 97-110.
- Gelfand, A. E., Ecker, M. D., Knight, J. R., & Sirmans, C. (2004). The dynamics of location in home price. *Journal of Real Estate Finance and Economics*, 29(2), 149-166.
- Gibbons, S., & Machin, S. (2005). Valuing rail access using transport innovations. *Journal of Urban Economics*, 57(1), 148-169.
- Glaeser, E. L., & Luttmer, E. F. (2003). The misallocation of housing under rent control. *American Economic Review*, 93(4), 1027-1046.
- Hahn, Anja M. and Kholodilin, Konstantin A. and Walzl, Sofie R. and Fongoni, Marco, Forward to the Past: Short-Term Effects of the Rent Freeze in Berlin (February 2022). DIW Berlin Discussion Paper No. 1999, Available at SSRN: <https://ssrn.com/abstract=4047633> or <http://dx.doi.org/10.2139/ssrn.4047633>
- Hanink, D. M., Cromley, R. G., & Ebenstein, A. Y. (2012). Spatial variation in the determinants of house prices and apartment rents in China. *The Journal of Real Estate Finance and Economics*, 45(2), 347-363.
- Hill, R. J., & Syed, I. A. (2016). Hedonic price–rent ratios, user cost, and departures from equilibrium in the housing market. *Regional Science and Urban Economics*, 56, 60-72.
- Hyun, D., & Milcheva, S. (2018). Spatial dependence in apartment transaction prices during boom and bust. *Regional Science and Urban Economics*, 68, 36-45.
- John, B. (1996). Mass transportation, apartment rent and property values. *Journal of Real Estate Research*, 12(1), 1-8.
- Kholodilin, Konstantin A. and Mense, Andreas, Internet-Based Hedonic Indices of Rents and Prices for Flats: Example of Berlin (February 1, 2012). DIW Berlin Discussion Paper No. 1191, Available at SSRN: <https://ssrn.com/abstract=2013076> or <http://dx.doi.org/10.2139/ssrn.2013076>
- Kholodilin, K.A., Mense, A. and Michelsen, C., 2017. The market value of energy efficiency in buildings and the mode of tenure. *Urban Studies*, 54(14), pp.3218-3238.
- Kim, C. W., Phipps, T. T., & Anselin, L. (2003). Measuring the benefits of air quality improvement: a spatial hedonic approach. *Journal of Environmental Economics and Management*, 45(1), 24-39.
- Malpezzi, S. (2003). Hedonic pricing models: a selective and applied review. In *Housing Economics and Public* (pp. 67-89). Oxford: Blackwell Science.
- McMillen, D. P., & McDonald, J. (2004). Reaction of house prices to a new rapid transit line: Chicago's midway line, 1983–1999. *Real Estate Economics*, 32(3), 463-486.
- Militino, A., Ugarte, M., & Garcia-Reinaldos, L. (2004). Alternative models for describing spatial dependence among dwelling selling prices. *Journal of Real Estate Finance and Economics*, 29(2), 193-209.
- Pace, R. K., Barry, R., Clapp, J. M., & Rodriguez, M. (1998). Spatiotemporal autoregressive models of

- neighborhood effects. *Journal of Real Estate Finance and Economics*, 17(1), 15-33.
- Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of Political Economy*, 82(1), 34-55.
- Sirmans, S., Macpherson, D., & Zietz, E. (2005). The composition of hedonic pricing models. *Journal of Real Estate Literature*, 13(1), 1-44.
- Sirmans, S., Sirmans, C., & Benjamin, J. (1989). Determining apartment rent: the value of amenities, services and external factors. *Journal of Real Estate Research*, 4(2), 33-43.
- Small, K. A., & Steimetz, S. S. (2012). Spatial hedonics and the willingness to pay for residential amenities. *Journal of Regional Science*, 52(4), 635-647.
- Smith, T. E., & Wu, P. (2009). A spatio-temporal model of housing prices based on individual sales transactions over time. *Journal of Geographical Systems*, 11(4), 333-355.
- Sun, H., Tu, Y., & Yu, S. M. (2005). A Spatio-Temporal Autoregressive Model for Multi-Unit Residential Market Analysis. *Journal of Real Estate Finance and Economics*, 31(2), 155-187.
- Thanos, S., Dubé, J., & Legros, D. (2016). Putting time into space: the temporal coherence of spatial applications in the housing market. *Regional Science and Urban Economics*, 58, 78-88.
- Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, 46(2), 234-240.
- Wilhelmsson, M. (2002). Spatial models in real estate economics. *Housing, Theory and Society*, 19(2), 92-101.
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. London: MIT press.

## Appendix

[Insert Table A here]