

Trading Volume-Induced Spatial Autocorrelation in Real Estate Prices

S. K. Wong · C. Y. Yiu · K. W. Chau

Published online: 25 May 2012

© The Author(s) 2012. This article is published with open access at Springerlink.com

Abstract Spatial dependence is often seen as a problem in econometrics rather than in economics. This study seeks to find an economic explanation for spatially correlated real estate prices. We posit spatial dependence as a process to discover price information from neighboring property transactions. Weaker spatial dependence is expected when price information in the immediate vicinity of a subject property is abundant. In the context of apartment buildings, in addition to the more commonly known horizontal dependence, there is also spatial dependence in the vertical dimension within the same building. Based on more than 18,000 transactions of highly homogeneous apartment units in Hong Kong, we found that the trading volume of a building depresses horizontal spatial dependence, but raises vertical spatial dependence. This not only confirmed the role of trading volume in the real estate price discovery process, but also questioned the validity of constant spatial autocorrelation assumption adopted in many studies.

Keywords Spatial dependence · Trading volume · Price discovery · Real estate

Introduction

Spatial dependence or autocorrelation has been analyzed in the real estate field for more than three decades, but the focus has so far been on correcting for bias or improving efficiency in the estimation process rather than on finding the underlying cause of spatial dependence. In other words, spatial dependence is treated more as an

S. K. Wong · C. Y. Yiu · K. W. Chau (✉)

Department of Real Estate and Construction, The University of Hong Kong, Pokfulam Road, Hong Kong, Hong Kong
e-mail: hrrbckw@hku.hk

S. K. Wong
e-mail: kelvin.wong@hku.hk

C. Y. Yiu
e-mail: cyyiu@hku.hk

econometric problem than as a phenomenon subject to economic inquiry. This paper takes the latter perspective and attempts to test the time dependence motivation for spatial dependence, i.e. present decisions are influenced by past neighboring actions (LeSage and Pace 2009).

In explaining the spatial pattern of real estate prices, the fundamental question is: why are these prices spatially correlated even after controlling for quality differences through a hedonic model? There are, typically, two conjectures for the cause, namely: (1) omitted variables and (2) information spillovers or searches.¹ The former is a problem of the researcher—spatial dependence is detected simply due to specification error or data limitation. We are not interested in why the researcher erred or data was imperfect. The latter, in contrast, concerns the search behaviour of market participants to be investigated; whether spatial information affects their search and hence market pricing requires economic inquiry. To date, there are neither theoretical justifications nor empirical tests on the information search conjecture. Our study seeks to fill this research gap by examining if spatial dependence varies with the amount of information available.

The other motivation of this study is that most spatial models developed in the literature so far were applied, or applicable, to real estate on a two-dimensional plain only.² But real estate developments are undoubtedly three-dimensional, especially when high-rise buildings or elevated topographies are concerned. Although it is technically straightforward to extend a spatial matrix to measure distance in a three-dimensional space, whether or not market participants treat distance as directionless is somewhat unknown. In our study of high-rise buildings, units are identified not only by their X-Y coordinates on a plane (called “horizontal dimension”), but also by their altitudes (called “vertical dimension”). Our spatial models will be flexibly formulated such that spatial dependence between buildings in the horizontal dimension is allowed to differ from spatial dependence between floor levels in the vertical dimension.

By dividing space into different dimensions, we can subject the information search conjecture to a more critical test. In a high-rise setting, units within the same building are typically more comparable with each other than units outside of the building, but the amount of price information (e.g. in terms of the number of prior sales) within the same building is generally small compared to the volume of transactions in other buildings. This is the tradeoff that market participants face when deciding which information—vertical or horizontal—to rely more on. A testable implication is that when information in the vertical dimension is more abundant, participants should rely *relatively* more on information from the same building. That is, the percentage of spatial dependence in the vertical dimension should increase and that for the horizontal counterpart should decrease. The separation of the price information into

¹ Other causes include: spatial heterogeneity, externalities, and model uncertainty (LeSage and Pace 2009). In the context of a hedonic price model, Bell and Bockstael (2000) contended that spatial autocorrelation is either due to: (1) structural spatial dependencies across the observations of the dependent variable and (2) spatial dependence across the error term caused by omitted explanatory variables, which are themselves spatially correlated.

² But three-dimensional data is often involved in the geostatistics literature, e.g. kriging.

horizontal and vertical dimensions, therefore, provides a strong test of the information search conjecture.

Our method also enables us to avoid the omitted variable problem due, for example, to unmeasured location characteristics. Omission of such characteristics should only affect the *level* of spatial dependence. Our information search argument, however, focuses on *changes* of spatial dependence in the different dimensions as a result of changes in the amount of price information. Even if we cannot avoid omitting certain location variables, our conclusion about the information search conjecture should remain intact.

In summary, our study aims to explain spatial dependence in real estate prices by way of a information search conjecture. The next section selectively reviews the relevant literature on the spatial autocorrelation of real estate prices. After that, we will introduce our models based on the general proposition that the reliability of past information increases with trading volume. “[Data and Variables](#)” describes our data and variables. “[Results](#)” presents the empirical test. In our case of high-rise buildings, we will test if an increase in trading volume within a building would induce a positive effect on vertical spatial dependence and a negative effect on horizontal spatial dependence. The last section is the conclusion.

Literature Review

Spatial autocorrelation has played an increasingly important role in analyzing spatial data since the late 1970s. Earlier texts include Silk (1979), Cliff and Ord (1981), Miron (1984), Upton and Fingleton (1985), Anselin (1988), and Odland (1988). One of the earliest studies in the real estate field is that of Dubin (1988), who proposed a maximum likelihood method for adjusting for spatial autocorrelation in a hedonic pricing model. Can (1992) then introduced different ways to specify a spatial autoregressive process for housing prices. In general, a spatial hedonic pricing model takes one of two forms: the spatial error model or the spatial lag model. The former incorporates a spatial process into the error term to account for omitted variables (e.g. Dubin 1992; Dubin et al. 1999), whereas the latter uses a spatially lagged dependent variable to account for spatial spillovers (e.g. Can and Megbolugbe 1997; Kim et al. 2003).

Over the past decade, most real estate research has been devoted exclusively to devising “better” spatial estimation methods. Pace and Gilley (1997) incorporated spatial autoregression into a hedonic price model and found substantial improvement in estimation efficiency. Can and Megbolugbe (1997) reported that spatial dependence affected hedonic price estimates, particularly the accuracy of estimated housing price indices. Pace et al. (1998) documented improved out-of-sample forecasts with the use of a spatiotemporal model. A number of new estimators were proposed for spatial models, including kriged EGLS (Basu and Thibodeau 1998), generalized spatial two-stage least squares (Kelejian and Prucha 1998), and generalized moments (Kelejian and Prucha 1999). Sun et al. (2005) developed a Bayesian estimation method with a vertical spatial-temporal effect for a multi-unit high rise residential market, but they did not consider the role of information in the price discovery process. From all these studies, it is clear that spatial autocorrelation matters a lot in real estate prices; what we do not know well is why.

A possible explanation is information search. Can and Megbolugbe (1997) put forward a similar argument on the search for property comparables, but they did not conduct any test. We propose a new test that associates trading volume with spatial dependence. Research on liquidity or trading volume effects in real estate markets was done, for example, by Kluger and Miller (1990), Sirmans et al. (1995), and Stein (1995). Recently, more empirical evidence confirmed the significance of trading volume effects on the real estate pricing process (Benveniste et al. 2001; Ho 2003; Yiu et al. 2009, 2006, 2008). These studies established that liquidity or trading volume plays an important role in affecting the search cost of property information. Yet, the effects of trading volume on spatial dependence have never been investigated.

Another limitation of the existing literature is that of spatial weight. In all previous works except for that by Sun et al. (2005), the weight was defined by distance or contiguity between properties on a two-dimensional plain (see, for example, Haining 2003).³ However, high-rise developments are three-dimensional. In valuation, it is a common practice to give more weight to comparables found in the same block of a building. Similarly, topographical differences, like contour lines on a map, may also result in a three-dimensional space. Applying a two-dimensional spatial model to these three-dimensional situations may not be appropriate and could result in a distortion of crucial spatial information.

Model Development

Consider a typical multi-period hedonic price model for real estate:

$$P_{it} = X_i\beta + \tau_t + \varepsilon_{it} \quad (1)$$

where P_{it} is the log sale price of property i at time t ; X_i is a vector of property i 's characteristics (with 1 as its first element); β is the implicit real price of the characteristics; τ_t is the market-wide price level at time t ; and ε_{it} is an unobserved random element in each transaction with zero mean. Assuming a perfectly competitive market, this model essentially says that a property can be valued solely by its own characteristics and the prevailing implicit nominal prices.

Participants in the real estate market, however, have incomplete information about property characteristics and/or prevailing implicit prices. Our model focuses *solely* on the latter imperfection—since trades are infrequent and decentralized, no one has perfect knowledge of the current sale price of other properties or current market trends. What traders can do is to look back at the prices of recently sold neighboring properties. Past price information is useful when the real estate market is less than efficient (Case and Shiller 1989), transaction prices are a noisy signal of the true price (Quan and Quigley 1989), or the search for buyers and sellers is time-consuming (Wheaton 1990).

This backward-looking behaviour can be incorporated into the hedonic price model as a spatial autoregressive process motivated from a time-dependence

³ An exception is Goetzmann and Spiegel (1997), who used socio-economic distance rather than physical distance.

perspective (LeSage and Pace 2009, p.25):

$$P_{it} = \rho \sum_{j=1}^n W_{ij} P_{j,t-k} + X_i \beta + \tau_t + \varepsilon_{it} \quad (2)$$

where $P_{j,t-k}$ is the sale price of a neighboring property j at time $t-k$ and W_{ij} is a spatial weight governing the proximity (e.g. inverse distance) between properties i and j . Given that $\sum_{j=1}^n W_{ij} = 1$, the spatial lag term $\sum_{j=1}^n W_{ij} P_{j,t-k}$ denotes a weighted average of space-time lagged price information. The spatial dependence parameter ρ therefore measures the degree of reliance on such past information in the price formation process. If past information is useful, ρ should be non-zero.

There is, however, an alternative explanation for a non-zero ρ : we—the researchers, not the traders—may have omitted some characteristics from Eq. (2) and the spatial lag term simply acts as an instrument for the omitted variable. Even if ρ turns out to be significantly different from zero, we are not able to tell whether this is driven by our model inadequacy or traders' backward-looking behaviour.

To solve this problem, we identified a unique implication for the backward-looking price formation process. According to Quan and Quigley (1989), transaction prices are bargaining outcomes that produce noisy signals of the true market price. Assuming individual bargaining strengths are random with constant variance, traders would be less (more) capable of inferring the true price from these noisy signals when the sample size, or amount of past transactions, is small (large). We therefore expected that the spatial dependence parameter should not be fixed, but should vary with lagged market trading volume L_{t-k} . The higher the past trading volume, the more that traders would rely on past information. Giving spatial dependence a testable economic interpretation is the first innovation of this study.

We incorporated L_{t-k} into Eq. (2) and rewrote the equation in a stacked form without the subscripts i and j :

$$P_t = (\rho + \gamma R) W P_{t-k} + X \beta + \tau_t + \varepsilon_t \quad (3)$$

P_t is an $n \times 1$ vector of sale prices, with the subscript t retained to emphasize the time lag structure. W is an $n \times n$ spatial weight matrix. It is triangular because traders are only allowed to look back, not forward. R is an $n \times n$ scaling matrix with lagged trading volume L_{t-k} on its diagonal:

$$R_{ij} = \begin{cases} L_{t-k} & \text{when } i = j \\ 0 & \text{when } i \neq j \end{cases}$$

γ is a scalar parameter for the trading volume-spatial lag interaction term. It should have a positive sign if the reliance on past information depends on trading volume. This prediction comes straight from a backward-looking price formation process, not the omitted variable argument.

Our second innovation was to generalize Eq. (3) to cases with a more complex, but rather common, spatial setting. Consider an apartment unit in a high-rise building. Its neighbors include units above and below it within the same building, as well as units in other adjacent buildings. There is certainly no technical problem measuring distance between units in three-dimensional spaces, but whether or not spatial dependence is the same across different dimensions is an economic question.

Suppose there are two comparables equidistant from a subject unit: one located on a higher floor level, and the other located in a different building. Which one should be relied more upon? If comparables in the vertical and horizontal dimensions are treated differently, a more flexible functional form for Eqs. (2) and (3) is needed.⁴

To allow spatial dependence to vary across dimensions, we defined two $n \times n$ spatial weight matrixes: 1) W^V for proximity (e.g. inverse vertical distance) between units within the same building and 2) W^H for proximity (e.g. inverse horizontal distance) between buildings. For W^V , proximity was set to zero whenever a pair of transactions did not occur within the same building. Conversely, for W^H , proximity was zero whenever a pair of transactions occurred within the same building. Substituting these matrixes to Eq. (2), we got:

$$P_t = \rho^H W^H P_{t-k} + \rho^V W^V P_{t-k} + X\beta + \tau_t + \varepsilon_t \quad (4)$$

Equation (4) explicitly separated two dimensions of past price information: a weighted average price within the same building, $W^V P_{t-k}$, and a weighted average price outside of the building, $W^H P_{t-k}$. Their impacts on the price formation process are governed, respectively, by vertical spatial dependence within the same building, ρ^V , and horizontal spatial dependence outside of the building, ρ^H . To price an apartment unit, transactions within the same building are generally considered “closer,” or more comparable, than those in other buildings. But at the same time, transactions within the same building are much more limited. Whether ρ^V or ρ^H should be larger becomes an empirical question.

Although the relative magnitude of ρ^V and ρ^H cannot be ascertained *a priori*, we can deduce their direction of change when trading volume varies across buildings and time. Different from the case in Eq. (3), trading volume is expected to affect the price formation process in a more subtle way. Compare two apartment buildings with different trading volumes. The more transacted building should allow traders to rely *relatively* more on the price information from units within it than those outside it. The same applies to a single building with varying trading volumes over two periods: during a more heavily-transacted period, traders can rely *relatively* more on the price information from units within it than outside it. A higher trading volume should, therefore, strengthen vertical spatial dependence within a building and weaken horizontal spatial dependence with other buildings. We can modify Eq. (4) to take into account such effects:

$$P_t = (\rho^H + \gamma^H R) W^H P_{t-k} + (\rho^V + \gamma^V R) W^V P_{t-k} + X\beta + \tau_t + \varepsilon_t \quad (5)$$

where R is an $n \times n$ scaling matrix with L_{t-k} on its diagonal and L_{t-k} is the lagged trading volume of the building in which the subject unit is located. The new parameters γ^H and γ^V are expected to be negative and positive, respectively.

⁴ We do not consider interactions between the vertical and horizontal processes because they represent two distinct sources of price information, i.e. those within the same building and between different buildings.

Data and Variables

To test the above models, we used the transaction data for a huge housing development in Hong Kong—Taikoo Shing. The development consists of 61 30-storey high homogeneous apartment buildings (Fig. 1). This entailed the spatial modeling of the vertical and horizontal dimensions, as in Eqs.(4) and (5). We chose Taikoo Shing mainly because of its high transaction frequency: on average, there were 250 transactions per quarter from 1992 to 2009. As shown in Fig. 2, its transaction frequency varied widely over time, enabling us to identify the effect of trading volume on spatial dependence.

Another reason why we chose Taikoo Shing was that information on its property characteristics was rather complete—many real estate agents specialize in trading units at Taikoo Shing and can provide prospective buyers with all the essential property information. What buyers and sellers do not know well is the current price information, as transaction details are usually not available to the public until one month after a deal is made. Traders, therefore, could only look back to extract price information from past sales. This motivated us to develop our models solely from the price formation perspective.

In our models, we set current price (P_t) to depend on lagged price information available 1 month prior to a sale (P_{t-1}). Lagged price information spans a t -month period—transactions that occurred within the past 3 months to a sale are deemed relevant for determining current prices. Such a short reference period is not unreasonable, given the relatively high efficiency of Hong Kong's real estate market. We performed robustness checks on shorter and longer reference periods (from a one-month period to a six-month period) and the results were more or less the same.

As for geographical boundaries, the high homogeneity and transaction frequency of Taikoo Shing allowed us to confine comparables to past sales within the development. That meant traders would not have to look outside the development for comparables. Of course, no one would consider every sale within Taikoo Shing as equally relevant, so comparables have to be weighted by proximity. For the horizontal spatial weight matrix (W^H), we measured proximity by the inverse distance between

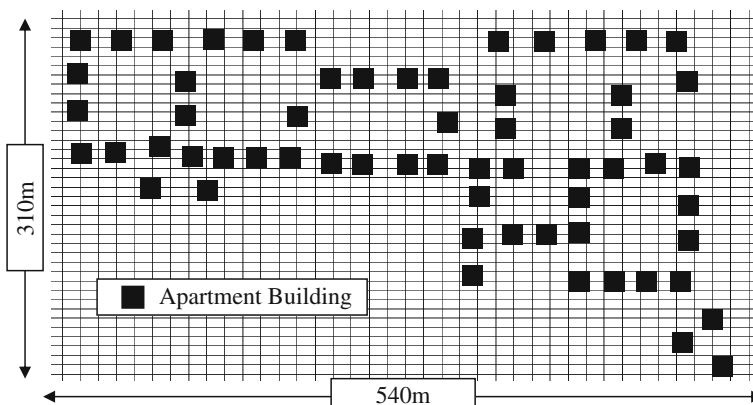


Fig. 1 Spatial distribution of the buildings in Taikoo Shing Note: the location of each building is based on the coordinates provided by the Lands Department, Hong Kong Government

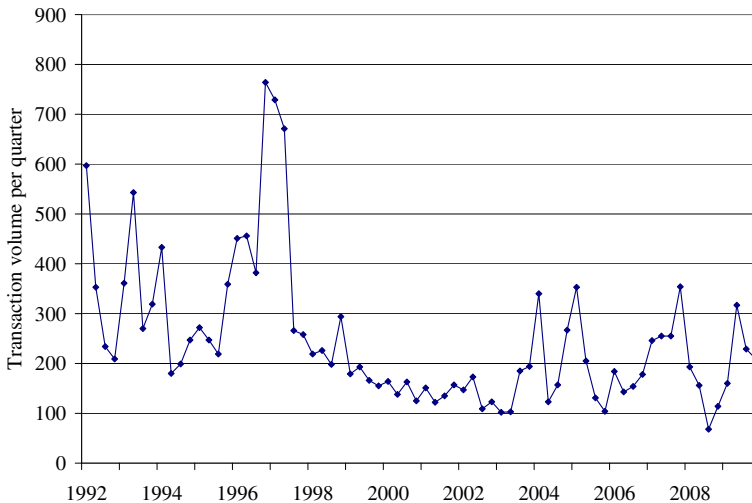


Fig. 2 Trading volume per quarter in Taikoo Shing

buildings. As shown in Fig. 1, the buildings are regularly spaced and the typical distance between two adjacent buildings is about 30 m. For the vertical spatial weight matrix (W^V), proximity is measured by the inverse distance between floors, and the typical headroom of a floor is 2.5 m. For instance, if one unit is on 5/F and the other on 8/F, then the floor distance is $(8-5)*2.5$ or 7.5 m. Both the horizontal and vertical spatial weight matrixes are row-stochastic.

Trading volume (L_{t-k}), as discussed in the previous section, was measured at the building level. Consistent with the treatment for lagged prices, the relevant timeframe for trading volume is a three-month period. In addition to this absolute measure, we will also use relative trading volume (\tilde{L}_{t-k})—the trading volume of a building divided by the trading volumes of other buildings—as a robustness check. We did not have *a priori* knowledge of whether absolute or relative trading volume shifts traders' price formation processes.

We had a comprehensive list of property characteristics for each unit, including its size ($AREA$), floor level (FLR),⁵ building age (AGE), views (full sea view $FVIEW$, partial sea view $PVIEW$, and other views), and surrounding environment (captured by a fixed effect for each of the eight phases). Since many hedonic studies have shown that the effects of $AREA$, FLR , and AGE on prices are non-linear, we added their squared terms to allow for a more flexible functional form. These characteristics, together with their squared terms, if applicable, were entered into our models as X . A series of monthly time dummies were used to capture the time effects (τ_t).⁶

Before we present the regression results of our spatial models, there is a final note on the estimation method. MLE is commonly used for spatial model estimation due to the bi-directional nature of spatial dependence—the price of one house affects and is

⁵ Note that FLR refers to the vertical location of an apartment unit, e.g. if $FLR=5$, then the unit is five storeys above the ground floor. It is different from the inverse floor distance weight in W^V , which is calculated from the vertical distance between a pair of units.

⁶ The Hong Kong housing market is highly volatile, so monthly time dummies are considered more appropriate than quarterly or yearly dummies for capturing short-term fluctuations.

affected by the price of another house (Anselin, 1999). However, with a time dimension in our study, current prices were restricted to depend on past prices, but not vice versa. In other words, the spatial lag terms in our equations were *not* endogenous, so the OLS estimator was consistent and asymptotically efficient under the usual i.i.d. assumption. Our estimation sample consisted of 18,457 transactions from 1992 to 2009. Descriptive statistics of the variables are shown in Table 1.

Results

Table 2 presents the estimated coefficient of each variable in Eqs. (1), (4), and (5), with its corresponding *p*-value in brackets. The fixed effects for phases and time are, however, not reported to simplify presentation. The bottom of the table shows the R-squared and adjusted R-squared values of each model.

Equation (1) was the traditional hedonic model. All coefficients except that for AGE^2 were significant at the 1 % level with the expected signs. The adjusted R-squared value was 90.39 %, which was reasonable given the high homogeneity of the apartment units in the development. Including the spatial terms did not seem to improve the R-squared value much. As indicated in the Introduction, the purpose of this paper was to explain how price information affects spatial dependence rather than finding a method to fit housing prices better. So, our main interest lay in the marginal effects, as we shall discuss below. The small improvement on R-squared value could indicate that omitted neighborhood characteristics were minimal in our homogenous sample, so the explanatory power of our spatial lags was not as strong as that of other spatial hedonic studies.

Equation (4) was the hedonic model with two spatial autoregressive processes: one for the horizontal dimension (between buildings) and the other for the vertical dimension (between floors). We found that both spatial effects were significant at the 1 % level: the horizontal spatial dependence ($W^H P_{t-k}$) was 0.3021 and the vertical spatial dependence ($W^V P_{t-k}$) was 0.0242. This confirmed that traders do look backwards for price information. Indeed, the size of the coefficients suggested that

Table 1 Descriptive statistics of our variables

	Unit	Mean	Std. Dev.	Minimum	Maximum
P_t	ln(HK\$M)	1.352	0.402	0.049	2.674
$W^H P_{t-k}$	ln(HK\$M)	1.319	0.274	0.582	1.943
$W^V P_{t-k}$	ln(HK\$M)	1.278	0.452	0.000	2.583
L_{t-k}	no. of sales	5.228	3.776	0.000	27.00
\tilde{L}_{t-k}	ratio	0.017	0.009	0.000	0.077
AGE	year	17.58	6.36	4.25	33.00
FLR	storey	14.43	7.43	0.00	29.00
$AREA$	ft ²	802.2	165.3	541.0	1,752
$PVIEW$	dummy	0.049	0.215	0.000	1.000
$FVIEW$	dummy	0.100	0.301	0.000	1.000

Table 2 Regression results

	Eq. (1)	Eq. (4)	Eq. (5)	Eq. (5)
$W^H P_{t-k}$	–	0.3021 (0.00)	0.3057 (0.00)	0.3319 (0.00)
$W^V P_{t-k}$	–	0.0242 (0.00)	0.0209 (0.00)	0.0161 (0.00)
$RW^H P_{t-k}$	–	–	–0.0033 (0.00)	–
$RW^V P_{t-k}$	–	–	0.0033 (0.00)	–
$\tilde{R}^H P_{t-k}$	–	–	–	–2.4029 (0.00)
$\tilde{R}^V P_{t-k}$	–	–	–	2.3365 (0.00)
Intercept	–0.6393 (0.00)	–0.8520 (0.00)	–0.8443 (0.00)	–0.8375 (0.00)
<i>AGE</i>	–0.0013 (0.00)	–0.0011 (0.00)	–0.0011 (0.00)	–0.0011 (0.00)
<i>AGE</i> ²	–2.27E-07 (0.22)	–1.16E-07 (0.53)	–2.01E-07 (0.28)	–7.52E-08 (0.68)
<i>FLR</i>	0.0130 (0.00)	0.0129 (0.00)	0.0129 (0.00)	0.0128 (0.00)
<i>FLR</i> ²	–0.0003 (0.00)	–0.0003 (0.00)	–0.0003 (0.00)	–0.0003 (0.00)
<i>AREA</i>	0.0023 (0.00)	0.0022 (0.00)	0.0022 (0.00)	0.0022 (0.00)
<i>AREA</i> ²	–5.69E-07 (0.00)	–5.41E-07 (0.00)	–5.29E-07 (0.00)	–5.07E-07 (0.00)
<i>PVIEW</i>	0.0252 (0.00)	0.0240 (0.00)	0.0239 (0.00)	0.0238 (0.00)
<i>FVIEW</i>	0.0896 (0.00)	0.0926 (0.00)	0.0925 (0.00)	0.0926 (0.00)
R ²	0.9051	0.9058	0.9059	0.9061
Adj. R ²	0.9039	0.9046	0.9046	0.9048

p-value in brackets

traders rely less on past sales in the buildings they trade than on those in other buildings. At first sight, this result looked counter-intuitive because units within the same building should be more similar or comparable to each other than to those in other buildings. But this interpretation ignored the role of trading volume in reducing information uncertainty. Consider prior sales from the same building: while their information *per sale* is richer, the *quantity* of sales within a building is smaller. If sales were equally distributed among Taikoo Shing's 61 buildings, the expected

trading volume of a building would be 60 times smaller than that of other buildings. A better interpretation of the relative magnitude of the two spatial dependence parameters would be that the trading volume effect outweighs the similarity (i.e., information per sale) effect.

Equation (5) added trading volume to the spatial processes in Eq. (4). Two types of trading volume—absolute (L_{t-k} and R) and relative (\tilde{L}_{t-k} and \tilde{R})—were used and their results are reported separately in the last two columns of Table 2. In both cases, we found that trading volume played a significant role in the price formation process. For absolute trading volume, its joint effects with the horizontal spatial lag and with the vertical spatial lag were -0.0033 and 0.0033 , respectively. Both effects were significant at the 1 % level. This confirmed our belief that a high trading volume for a building induces stronger spatial dependence between floors, but weaker spatial dependence between buildings. When a building's trading volume is high, traders place relatively greater importance to comparables found on upper or lower floors, but relatively lower importance to comparables from other buildings. Similar results hold for relative trading volume: its joint effects with the horizontal spatial lag and vertical spatial lag were -2.4 and 2.3 , respectively. The conclusion remains the same no matter which definition of trading volume is used.

Conclusion

This study made three contributions. First, it provided an economic explanation for spatial dependence in real estate prices based on the information search framework. This explanation implied that traders rely more on spatially lagged price information when lagged trading volume is high. Second, we extended the traditional spatial model to a three-dimensional setting by allowing spatial dependence to vary across dimensions. Our model can be easily applied to study spatial dependence between units in high-rise buildings. Third, we tested the information search explanation with a large number of transactions in Hong Kong, and our data fully supports the explanation. The results clearly demonstrated how traders look backwards for price information in the vertical and horizontal dimensions. Our model can be used for many practical purposes, such as mortgage valuation, investment analysis, development appraisal, taxation (e.g. stamp duty and rates), and homeowner insurance.

This study also has implications for future research. Rather than incorporating a long list of explanatory variables into a hedonic price model, spatial econometrics keeps the model simple by augmenting it with a spatial dependence process. However, the choice of a spatial lag or spatial error model is often arbitrary and is, at best, based on trial and error. Our study showed that information searches are a cause for spatial dependence, and provides future studies with a theoretical justification for the model choice, but does not necessarily imply that the omitted-variable conjecture is refuted. Moreover, we showed that spatial dependence is not fixed, but varies with trading volume. This means a more flexible functional form for the spatial lag is needed when trading volume varies over time or across locations.

Acknowledgements This project is financially supported by General Research Fund (Project Reference: HKU 7108/08E). We are grateful for the comments from our discussant Kelley Pace at the AREUEA

Annual Conference at Denver and participants at the Asia Pacific Real Estate Research Symposium at Hong Kong.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Anselin, L. (1988). *Spatial econometrics: Methods and models*. Dordrecht: Kluwer Academic Publishers.
- Anselin, L. (1999). *Spatial Econometrics* (Working Paper). Dallas, TX: Bruton Center, University of Texas at Dallas. http://www.csiss.org/learning_resources/content/papers/baltchap.pdf. Accessed 26 March 2010.
- Basu, S., & Thibodeau, T. G. (1998). Analysis of autocorrelation in housing prices. *Journal of Real Estate Finance and Economics*, 17(1), 61–85.
- Bell, K., & Bockstael, N. (2000). Applying the generalized moments estimation approach to spatial problems involving micro-level data. *Review of Economics and Statistics*, 82(1), 72–82.
- Benveniste, L., Capozza, D. R., & Seguin, P. J. (2001). The value of liquidity. *Real Estate Economics*, 29(4), 633–660.
- Can, A. (1992). Specification and estimation in hedonic housing price models. *Regional Science and Urban Economics*, 22, 453–474.
- Can, A., & Megbolugbe, I. (1997). Spatial dependence and housing price index construction. *Journal of Real Estate Finance and Economics*, 14, 203–222.
- Case, K. E., & Shiller, R. J. (1989). The efficiency of the market for single-family homes. *American Economic Review*, 79(1), 125–137.
- Cliff, A. D., & Ord, J. K. (1981). *Spatial process: Models and applications*. London: Pion.
- Dubin, R. A. (1988). Estimation of regression coefficients in the presence of spatially autocorrelated error terms. *Review of Economics and Statistics*, 70(3), 466–474.
- Dubin, R. A. (1992). Spatial autocorrelation and neighborhood quality. *Regional Science and Urban Economics*, 22(3), 433–452.
- Dubin, R. A., Pace, R. K., & Thibodeau, T. G. (1999). Spatial autocorrelation techniques for real estate data. *Journal of Real Estate Literature*, 7(1), 79–95.
- Goetzmann, W., & Spiegel, M. (1997). A Spatial Model of Housing Returns and Neighborhood Substitutability. *Journal of Real Estate Finance and Economics*, 14, 11–31.
- Haining, R. P. (2003). *Spatial data analysis: Theory and practice*. New York: Cambridge University Press.
- Ho, M. H. C. (2003). Liquidity and price differentials: evidence in the Hong Kong residential re-sale market. *Housing Studies*, 18(5), 745–763.
- Kelejjan, H., & Prucha, I. (1998). A generalized spatial two stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *Journal of Real Estate Finance and Economics*, 17, 99–121.
- Kelejjan, H., & Prucha, I. (1999). A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review*, 40, 509–533.
- 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, 24–39.
- Kluger, B. D., & Miller, N. G. (1990). Measuring residential real estate liquidity. *Real Estate Economics*, 18(2), 145–159.
- LeSage, J. P., & Pace, R. K. (2009). *Introduction to spatial econometrics*. Boca Raton: CRC Press.
- Miron, J. (1984). Spatial autocorrelation in regression analysis: A beginners guide. In G. L. Gaile & C. J. F. Willmott (Eds.), *Spatial statistics and models* (pp. 201–222). Dordrecht: D. Riedel.
- Odland, J. (1988). *Spatial autocorrelation*. California: Sage.
- Pace, R. K., & Gilley, O. W. (1997). Using the spatial configuration of the data to improve estimation. *Journal of Real Estate Finance and Economics*, 14, 333–340.
- 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.
- Quan, D. C., & Quigley, J. M. (1989). Inferring an investment return series for real estate from observations from sales. *AREUEA Journal*, 17(2), 218–230.
- Silk, J. (1979). *Statistical concepts in geography*. Boston: George Allen and Unwin.

- Sirmans, C. F., Turnbull, G. K., & Dombrow, J. (1995). Quick house sales: seller mistake or luck? *Journal of Housing Economics*, 4(3), 230–243.
- Stein, J. C. (1995). Prices and trading volume in the housing market: a model with down-payment effects. *Quarterly Journal of Economics*, 110(2), 379–406.
- 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.
- Upton, G., & Fingleton, B. (1985). *Spatial data analysis by example (Volumes 1 & 2)*. Chichester: Wiley.
- Wheaton, W. C. (1990). Vacancy, search and prices in a housing market matching model. *Journal of Political Economy*, 98(6), 1270–1292.
- Yiu, C. Y., Man, K. F., & Wong, S. K. (2009). Price dispersion and trading volume in housing markets. *Journal of Property Research*, 25(3), 203–219.
- Yiu, C. Y., Tam, C. S., & Lee, P. Y. (2006). Volume-related heteroskedasticity and liquidity premium in hedonic pricing model. *Journal of Property Research*, 23(1), 39–51.
- Yiu, C. Y., Wong, S. K., & Chau, K. W. (2008). Transaction volume and price dispersion in the presale and spot real estate markets. *Journal of Real Estate Economics and Finance*, 38, 241–253.