J Real Estate Finan Econ (2013) 47:341–369
DOI 10.1007/s11146-011-9359-3

Spatial and Temporal Dependence in House Price Prediction

Xiaolong Liu

Published online: 14 January 2012

Abstract This paper incorporates spatial and temporal dependence among housing transactions in predicting future house prices. We employ the spatiotemporal autoregressive model and structure the spatial and temporal weighting matrices as in Pace et al. (1998). We control for the time variation of both the attribute prices and the spatial and temporal dependence parameters through performing the analysis on an annual basis. Spatial heterogeneity is accounted for using experience-based definition of submarkets by real estate professionals. Using a comprehensive housing transaction data set from the Dutch Randstad region, we show that integrating the spatial and temporal dependence within the hedonic modeling improves the prediction power as compared to traditional hedonic model that neglects these effects.

Keywords Spatial dependence · Temporal dependence · House price prediction · Hedonic model

Introduction

Predicting property values are of great interest to various parties in an economy. Individuals are interested in knowing the values of their properties before setting up their list prices. Tax authorities rely on the estimates of the properties' value as the basis for levying property taxes. Banks and mortgage providers conduct housing collateral valuation to qualify the borrowers for their mortgage applications. Real estate investors and portfolio managers devise and carry out their investment decisions based on periodic evaluations of their real estate portfolios.

Due to the heterogeneous nature of housing and the thin housing market with infrequent transactions, the house price determination process has traditionally been

X. Liu (✉)
School of Finance, Renmin University of China, No. 59, Zhongguancun Main Street, Haidian District, Beijing 100872, China
e-mail: x.liu@ruc.edu.cn

modeled within the hedonic framework, following the pioneer work by Rosen (1974). Within this framework, houses are viewed as bundles of attributes that offer utilities to their users, and the market equilibrium dictates house prices to be determined by the amount of utility-generating attributes possessed by these houses. The observable housing transaction prices are thus the value realization of heterogeneous bundles with varying amounts of different attributes which these houses contain. Within the hedonic framework, hedonic regression techniques, normally taking the log-linear functional form, are widely applied using real housing transaction data to filter out the implicit attributes' prices and make house price predictions. The hedonic regression specification takes structural as well as location attributes as explanatory variables with log transaction price as the dependent variable, and the coefficients can be interpreted as either elasticities or semi-elasticities. A notable problem with this standard ordinary least squares (OLS) specification is that errors may not be independent from one another, and such error correlation will lead to inefficient as well as erroneous inference concerning these parameter estimates. This is of particular concern in the housing market due to the spatial and temporal feature of housing transactions. Spatial dependence occurs when houses close in proximity tend to be correlated which violates the assumption that the housing transactions are independent in traditional hedonic regressions. Basu and Thibodeau (1998) argue that house prices are spatially correlated for two reasons. First, neighborhoods tend to be developed at the same time so that neighborhood properties have similar structural characteristics. Second, neighborhood residential properties share location amenities. The inclusion of location indicators in the hedonic regression specification does not necessarily remove the spatial dependence among the properties within the neighborhoods since they only capture the spatial effects that are shared by all the properties within certain geographical boundaries. Therefore, hedonic regressions that control for well-defined spatial locations may still have house price residuals exhibiting spatial dependence (Pace et al. 1998; Dubin 1998).

Another source that causes correlated error structure in traditional hedonic regression specification is that housing transactions tend to be correlated temporally. Such temporal dependence could arise due to earlier housing transactions containing information that is relevant to the pricing of the target property. Past housing transactions may proxy for the market trend of general house price development, and can also capture changes in institutional setting, such as changes in tax laws, as well as local amenities, all of which could lead to temporal correlation among housing transaction prices. Consequently, both the spatial and temporal dependence need to be accounted for if one aims to address the correlated errors in the traditional hedonic regression specification. Such exercise is not only beneficial to improving estimation efficiency and inference accuracy, but also having implication on house price prediction which is one of the most important tasks of empirical application of hedonic pricing models.

The purpose of this paper is to integrate the spatial and temporal dependence within the traditional hedonic regression specification, and assess if accounting for spatial and temporal effects in the housing market improves prediction performance relative to hedonic regression specification with such effects neglected. Given our focus on the empirical applications of hedonic price models, we limit our analysis to the spatiotemporal autoregressive (STAR) model of Pace et al. (1998) that corrects for both the spatial and temporal correlations in the error structure. In addition, the sparse structure of spatial and temporal weighting matrices in the STAR model offers computational efficiency especially when dealing with large data sets (Pace 1997; Pace and Barry 1997a and b). This computational advantage is of particular relevance to the empirical analysis of this paper since the working data set has over 400,000 housing transactions covering the Dutch Randstad region for the period from 1997 to 2007.

Most of the prior research has put emphasis on the model comparison between the default hedonic pricing model and various models that account for spatial correlated errors (Can 1990, 1992; Basu and Thibodeau 1998; Dubin et al. 1999; Gelfand et al. 2004; Militino et al. 2004; Case et al. 2004; Bourassa et al. 2007). These studies either employ small samples of housing transactions or do not explore the temporal structure in the housing market in making out-of-sample predictions since only prior housing transactions convey information relevant to the pricing of later transactions and not vice-versa. Moreover, as shown in a recent study by Nappi-Choulet and Maury (2009), the temporal heterogeneity may well exist in the property market when using property transaction data extending over a long period of time. This paper contributes to the existing literature on the following grounds. First, we employ a large non-U.S. data set, which, to our knowledge, is one of the first papers studying the prediction performance of spatiotemporal hedonic model in the Dutch housing market. Second, we control for spatial heterogeneity by including experience-based definition of submarkets in our regression analysis. To account for temporal heterogeneity of parameter estimates, we perform the analysis on an annual basis. Third, the prediction exercise recognizes the time structure in the housing market such that only earlier housing transactions are taken for predicting future house values. Our findings show that accounting for spatial and temporal dependence in traditional hedonic pricing model is not only theoretically warranted, but also contributes to better prediction performance. As a natural application of the spatiotemporal model in the context of house price index construction due to its better predictive power, we show that the house price index constructed using the OLS hedonic model consistently understates the more accurate house price development as captured by the spatiotemporal model on the basis of the single random sample that is used to represent the overall housing market.

The paper is structured as follows. "Literature Review" briefly reviews relevant literature. "Spatiotemporal Model and Estimation Procedure" illustrates the STAR model and highlights some estimation issues. "Data and Empirical Results" presents the data and discusses the estimation results. We identify the conclusions reached in "Conclusion".

Literature Review

Spatial Heterogeneity and Housing Submarkets

Spatial dependence and spatial heterogeneity are distinct phenomena in the housing market. Spatial heterogeneity refers to the low degree of substitutability of houses on the basis of both observable and unobservable characteristics, which occurs across well-defined submarkets. Houses within each of the submarkets exhibit high degree of homogeneity and are highly substitutable with similar implicit attributes' prices. In this context, spatial dependence still exists due to the substitutability of houses within each submarket. Therefore, if the housing submarkets are inappropriately defined, it would aggravate the problem of spatial correlated error terms in hedonic equations. Recognizing this fact, we give special consideration to the inclusion of well-defined submarkets into our regression analysis to control for the spatial heterogeneity effect.

Can (1990) is among the first to address both the spatial heterogeneity and spatial dependence in the modeling of spatial data. She uses Casetti's (1972) spatial expansion method where the structural parameters of the hedonic pricing model are allowed to vary over space. Spatial dependence in the error structure is treated using the so-called mixed regressive-autoregressive model which incorporates a spatially lagged dependent variable into the hedonic regression specification. Using a sample of 577 single family transactions in Columbus, Ohio MSA, she shows that accounting for both the spatial heterogeneity and spatial dependence explains the urban house price variations better than the traditional hedonic model on the basis of some in-sample statistics.

A number of studies advance in the search of well-defined housing submarkets such that spatial heterogeneity can be optimally controlled for within an OLS setup, and results are mixed. Bourassa et al. (1999) apply the statistical method that using principal component analysis to first extract a set of factors and performing cluster analysis on the basis of these factors to identify the composition of housing submarkets. Hedonic price equations are estimated using a priori classified submarkets and statistically generated submarkets. Using a survey data with 4,661 observations from Sydney and Melbourne, they find that submarkets based on the statistical procedures, with one exception, do not produce results that are much better than the a priori submarket classification.

Dunse et al. (2001) follow the same statistical procedures as Bourassa et al. (1999) to derive submarkets for the office market in two cities, Glasgow and Edinburgh. They show that different factors are important in influencing the structure of the office market in Scotland's major centers. In a follow-up study, Dunse et al. (2002) test parameter stability across different submarkets with both a priori and statistical definitions. They find little evidence of spatial heterogeneity of parameter estimates on the basis of Chow test. In addition, the test results do not differ between two submarket definitions. Put it differently, there is little difference in the submarket model based on agent's knowledge compared with a more complex structure constructed a combination of statistical procedures.

A similar result was also found in Bourassa et al. (2003), in which they explore the effects of alternative definitions of submarkets on the accuracy of predictions for mass appraisal purposes. Using 8,421 residential transactions in the city of Auckland, they compare submarkets based on small geographical areas defined by the real estate appraisers with a set of submarkets that are generated via statistical procedures. They find more accurate price prediction based on the housing market segmentation used by appraisers.

These previous research shed light on the validity of using submarkets defined by real estate professionals rather than those produced otherwise through statistical procedures. Two notes are put forward for the empirical unsuccessful application of statistically derived submarkets. First, the cluster analysis that has been widely applied in deriving the submarkets empirically implicitly gives equal weights to the attributes of housing stock. It seems likely that some attributes are more important than others. For instance, location may well conform to the ideal sense of submarket more precisely than do other factors in the clustering process to derive submarkets (Bourassa et al. 2003). Second, taking factors from supply side alone is not sufficient in classifying submarkets in practice. One can think of one good and one bad neighborhood that are contiguous with similar housing attributes. The two neighborhoods are treated as one submarket if only housing attributes are considered in the statistical process. However, the result would change if, for example, household income and education are considered as well. The submarkets defined by the real estate professionals, on the contrary, correct for these issues related to statistically generated submarkets since their submarket classification is experience-based that incorporates the usual confidential demand side factors as well as observable housing attributes.

Spatial and Temporal Dependence Modeling and Applications

The spatial nature of housing data has attracted significant amount of attention in the academic literature regarding the optimal way to model the spatial data. As summarized in Dubin (1998), there are two common ways to address the spatial dependence in the error structure of traditional hedonic pricing models. The first method is to model the process itself through using a spatial lag structure, and the second approach is to model the spatial error covariance matrix explicitly.[1] However, the current literature does not endorse the use of one approach over the other, which may depend on the working data at hand and the complexity involved in the implementation process. Nevertheless, model comparison exercises generally show better performance of spatial models relative to the traditional hedonic specification (Can 1990, 1992; Pace et al. 1998; Case et al. 2004; Militino et al. 2004) except Bourassa et al. (2007).

Can (1992) advances Can (1990) by applying rigorous testing of the presence of spatial effects. Similar models as in Can (1990) were estimated using 563 single-family houses for 1980 sold in the Franklin County of the Columbus metropolitan area. Results indicate the spatial models that incorporate both spatial dependence and spatial heterogeneity are better than traditional hedonic specification that includes neighborhood effects in terms of explaining house price variations. She also shows that neighborhood quality variable seems capable of capturing the spatial heterogeneity without introducing spatially varying marginal prices of structural housing attributes.

Militino et al. (2004) compares four spatial models as well as OLS without out-of-sample prediction. They estimate these four spatial models using 293 property transactions from Pamplona, Spain, and find that most inferences seem robust with respect to the spatial technique. In addition, between the two commonly used spatial models, the spatial autoregressive model (SAR) and conditional autoregressive model (CAR), CAR offers better performance than SAR on the basis of AIC or BIC values.

[1] See Dubin (1998) for the details of these two modeling approaches.

Bourassa et al. (2007) is an exception to an overwhelming favor of the spatial models over the OLS hedonic pricing model. They use submarket defined by valuers and apply alternative methods of controlling for the spatial dependence of house prices to a housing transaction database with 4,880 observations in Auckland. Without recognizing the temporal structure in the housing market, they apply repeated random sampling of 100 times to obtain prediction results. Moreover, their prediction exercises only consider estimated structural parameters without fully exploring the spatial information in the spatial models. Their prediction outcome favors the inclusion of submarket variables in OLS specification over alternative spatial models.

There have been few studies that address both the spatial and temporal features inherent in modeling housing data with exceptions such as Pace et al. (1998 and 2000), Case et al. (2004), Gelfand et al. (2004) and Nappi-Choulet and Maury (2009). Pace et al. (1998) investigates the best way to model spatial and temporal effect by comparing two models. One model uses indicators of time and location, and the other incorporates the spatiotemporal dependence in the error specification such that the target house price is affected not only by the early transactions in the housing market but also by prior transactions of neighbors. Using a large data set with 70,822 observations from 1969 to 1991 from Fairfax County Virginia, they demonstrate improved goodness of fit of using the STAR model to account for both the spatial and temporal dependence than the indicator based hedonic model. They conclude that the house price variations are strongly influenced by the sales prices of previously sold, neighboring properties.

Case et al. (2004) apply the standard hedonic model and three spatial models to a similar sample as used in Pace et al. (1998), and out-of-sample prediction accuracy is used for comparison purposes. Their final results show that two out of three spatial models outperform OLS with residuals from 15 nearest neighbors included in out-of-sample prediction. Moreover, their prediction results indicate the importance to incorporate the nearest neighbor transactions for predicting housing values.

Nappi-Choulet and Maury (2009) adopt the STAR model as in Pace et al. (1998) in the analysis of Paris office market while incorporating temporal heterogeneity to control for the time varying structural parameters as well as spatial and temporal dependence coefficients. Applying Paris office transaction data set with 2,587 observations between 1991 and 2005, they find that spatial and temporal dependence parameters differ strongly according to the transaction date, which implicates the assessment of price changes from 1991 to 2005 for the Paris office market.

Overall, previous literature demonstrate the need to account for spatial effects to address the correlated error in an OLS setup, but the optimal way of doing so remains unclear. In addition, the intrinsic temporal structure in housing transactions deserves attention when making out-of-sample predictions such that recent observations would not be used in predicting house values that are transacted in the past. Spatial heterogeneity should be modeled on the basis of experience-based definition of submarkets rather than those derived through complex statistically procedures, while temporal heterogeneity is of concern when the data set includes housing transactions over a long period of time.

Spatiotemporal Model and Estimation Procedure

We proceed to present the STAR model by Pace et al. (1998) which models the spatial and temporal dependence in the hedonic error structure through using a weighting matrix. The traditional hedonic regression is specified as follows,

$$Y = X\beta + u \tag{1}$$

where Y is a $n\times 1$ vector with log transaction prices, X denotes a $n\times k$ matrix, which normally includes time dummies, location indicator variables, and housing structural characteristics, β is a $k\times 1$ vector of parameters corresponding to k independent variables, and u refers to a $n\times 1$ vector of error terms. However, such specification overlooks the correlated errors. To account for this, Pace et al. (1998) subsume an autoregressive error process such that

$$u = Wu + \varepsilon \tag{2}$$

where W is a $n\times n$ weighting matrix, and ε is a $n\times 1$ vector of white noise. If W consists of only a spatial weighting matrix, the model is essentially a SAR model. However, as noted in Pace et al. (1998), in a temporal context, simply multiplying the dependent and the independent variables by the spatial weighting matrix does not remove all the autocorrelation effects. This is due to the fact that neighboring property transactions relative to the subject property may have taken place long time ago and therefore do not contain too much relevant information in the pricing of that particular property. In addition, as mentioned above, earlier housing transactions may convey other relevant information to the pricing of the target property. Therefore, there is need to take into account the previous transactions which are not necessarily close in proximity relative to the target property. Within the model specification, it leads to the inclusion of a temporal weighting matrix into W besides a spatial weighting matrix. The spatial and temporal elements in W thus generalize the SAR model to be the STAR model.

Following Pace et al. (1998), we structure all observations to be ordered according to time, such that first element in Y and first row of W correspond to the oldest housing transaction with the first row in X containing its attributes. Therefore, the implicit temporal feature of the housing transactions is recognized and accounted for in the STAR model. Moreover, it will be shown later that such time ordering of W contributes to computational efficiency in estimating the model. Combining (1) and (2), we arrive at a compact form of the STAR model as

$$(I - W)Y = (I - W)X\beta + \varepsilon \tag{3}$$

A general specification of W, as in Pace et al. (1998), would be

$$W = \phi_S S + \phi_T T + \phi_{ST} ST + \phi_{TS} TS \tag{4}$$

where S and T are spatial and temporal weighting matrices while ϕ_S and ϕ_T are spatial and temporal dependence parameters. ST and TS are the interaction matrices that allow for the modeling of potentially compound spatiotemporal effects with ϕ_{ST} and

ϕ_{TS} as their coefficients.[2] As a further generalization of (3) and (4), we write the unrestricted STAR model as follows,

$$Y = \alpha + X\beta_1 + SX\beta_2 + TX\beta_3 + STX\beta_4 + TSX\beta_5 + \phi_T TY + \phi_S SY + \phi_{TS} TSY + \phi_{ST} STY + \varepsilon \tag{5}$$

where β_1, β_2, β_3, β_4, β_5 are $k\times 1$ vectors of coefficients that correspond to independent variables as well as the spatial, temporal and spatial-temporal lagged independent variables. Our empirical analysis will be based on the unrestricted form of the STAR model (5).

The spatial weighting matrix S is generally structured on the basis of cardinal distances among housing transactions; see, for example, Can (1990 and 1992) and Militino et al. (2004). Another specification of the spatial weighting matrix is based on ordinal distances which is equivalent to the use of a fixed number of neighbors in space as in Pace et al. (1998), which reduces the problems posed by uneven housing densities. We follow Pace et al. (1998) in the specification of the spatial weighting matrix, and we limit the existence of spatial dependence within 30 (m_S=30) neighbors which are identified on the basis of Euclidean distance d_{ij} between every pair of observations j and i for every prior observation j relative to the ith observation ($j<i$). Consequently, we are able to sort these calculated distances and find our required number of neighbors relative to every housing transaction. We form neighbor matrices S_i ($i \in [1, m_S]$), where S_1 includes the closest previously sold neighbor, S_i stores the ith closest previously sold neighbor relative to every observation. The overall spatial weighting matrix S is computed as follows,

$$S = \frac{\sum_{i=1}^{m_s} \rho^i S_i}{\sum_{i=1}^{m_s} \rho^i} \tag{6}$$

where ρ is the spatial decay parameter capturing the impact of neighboring transactions based on proximity and falls between 0 and 1. By construction, S matrix is row standardized and lower triangular with zeros on its diagonal.

The temporal weighting matrix is structured in a similar fashion. We take the temporal dependence to be within 150 (m_t=150) previously sold houses relative to the current transaction. The ith row in the temporal matrix T contains m_t prior observations which are equally weighted with weight $\frac{1}{m_t}$.[3] Therefore, the T matrix is structured to be lower triangular and contains zeros on its diagonal with the ij th element as

$$T_{ij} = \frac{1}{m_t} \tag{7}$$

only if $i > j$ and $i - j \le m_t$, and $T_{ij} = 0$ otherwise.

[2] As noted in Pace et al. (1998), the interaction terms essentially account for the order of filtering both the dependent and the independent variables, especially when we do not have prior knowledge on which, space or time, to filter first.

[3] Pace et al. (1998) discuss the choice of the equal weight attached to each of the prior transaction to the current transaction which is based on an acceptable performance from preliminary fitting.

Since we are taking ordinal distances in both space and time in specifying the S and T matrices, both S and T are sparse with densities $\frac{m_s}{n}$ and $\frac{m_t}{n}$ respectively. The low densities in S and T matrices avoid the time consuming calculation of the determinant in maximizing the log-likelihood function and contribute to the computational efficiency (Pace 1997; Pace and Barry 1997a and b). Taking the spatial and temporal weighting matrix W as specified in (4) and substituting it into (3), we arrive at our working model which can be estimated using maximum likelihood. The concentrated log-likelihood function in parameters ϕ, as shown in Pace et al. (1998), is

$$L(\phi) = \ln|I - \phi_s S - \phi_t T - \phi_{st} ST - \phi_{ts} TS| - \left(\frac{n}{2}\right) \ln[SSE(\phi)] \tag{8}$$

where SSE is the sum of squared errors. Given the structure of S and T matrices, the log-determinant term drops out in the calculation of the log-likelihood, which greatly speeds up the estimation of the STAR model, especially when using large data set.

The implementation of the STAR model requires the specification of the optimal spatial decay parameter ρ. In this study, the optimal ρ is set through maximizing the likelihood function, and we implement this with a rough grid search process. Specifically, we take the optimal ρ to be falling within 0.2 and 0.85, and preliminary fitting is carried out for different ρ values with an interval 0.05 within the pre-specified range, for instance, 0.2, 0.25, 0.3, 0.35, …, 0.85. The optimal ρ is chosen such that the log-likelihood of the STAR model is maximized. This process of choosing the optimal ρ does not necessarily produce the true optimal spatial decay parameter that corresponds to the global maximum of the log-likelihood, but it is adopted as a trade-off between computational feasibility and bias arising from a priori subjective choice.

We pay special attention to several issues that have been widely discussed in the previous literature related to the application of the hedonic pricing model. First, when using housing transaction data that extend over a long period of time, we should account for temporal heterogeneity of the parameters in the model. It is reasonable to assume that parameters such as marginal housing attribute prices are also time-varying. In the short-run, the supply of reproducible housing attributes is inelastic, which underlies the fluctuation of housing attribute prices as the demand for certain attributes changes. However, in the long run, the housing attributes prices will be driven by the supply of these reproducible attributes with non-constant production costs. For irreproducible housing attributes such as a lake view and highway access, the prices of these attributes will be determined by demand both in the short run and in the long run. For example, as people become better well-off over time, houses with a lake view appreciate in prices since the lake view is irreproducible and there is no supply to accommodate the increasing demand of such housing attribute over time. Knight et al. (1995) relax the temporal parameter stability assumption, and apply seemingly unrelated regressions (SUR) to account for the time variation of implicit housing attribute prices in the context of house price index construction. Munneke and Slade (2001) focus on the commercial real estate price construction, and control for the temporal heterogeneity through running hedonic regressions for each of the sub-periods. Besides the time varying structural parameters, both the temporal and spatial dependence terms are also subject to temporal heterogeneity as established in Nappi-Choulet and Maury (2009). In this study, we perform both estimation and prediction on an annual basis, which is similar to Munneke and Slade (2001), to take into account temporal heterogeneity of parameter estimates.

The second issue related to the empirical application of hedonic pricing model is the existence of spatial heterogeneity such that parameters of interest in a hedonic model may differ according to the submarkets to where the property belongs. Can (1990) models the spatial heterogeneity via interactions between neighborhood quality variable and housing structural attributes. Can (1992) finds that hedonic model without such interactions perform equally well as the hedonic model estimated in Can (1990). In this paper, we model the spatial heterogeneity through incorporating submarkets defined by real estate professionals. Using the spatial drift terms to capture the spatial heterogeneity has been applied in Bourassa et al. (2007) and Nappi-Choulet and Maury (2009). The validity of adopting the experience-based submarket definition over driving the submarket statistically has been briefly discussed above and supported by the existing literature, such as Palm (1978), Michaels and Smith (1990), Dunse et al. (2001 and 2002) and Bourassa et al. (2003 and 2007).

Our main prediction exercise is performed on an annual basis as follows. For every year, we sort the housing transactions in our data from the earliest to the latest, and we use the first 80% of all observations as the in-sample for estimation, and the other 20% as the out-of-sample for prediction. In doing so, we recognize the intrinsic temporal structure in housing transactions, since only previous transactions are relevant for predicting the current house value and not vice-versa. As a robustness check to see if our prediction results are driven by a particular out-of-sample specification, we specify two other out-of-samples that are composed of 80–90% and 90–100% of all temporally ordered observations respectively. As a further robustness check, we retain 90% of all observations as the in-sample, and observations within 90–100% range as the out-of-sample to check if our results are sensitive to different specifications of in-sample. It is to note that, in predicting future house values, we not only utilize the estimated structural parameters as in Bourassa et al. (2007) but also explore information contained in the spatial and temporal neighbors relative to these transactions. Put it differently, for each observation in our out-of-sample, we first identify its spatial and temporal neighbors from a pool of temporally ordered housing transactions within the in-sample, and this information will be combined with parameter estimates from the in-sample in predicting the value of each house in the out-of-sample. Our main prediction exercise also guarantees that identical information set is taken for predicting house value using both the OLS model and the STAR model, so that model comparison on the basis of predictive power is performed on a level ground.

To check the robustness of our prediction results as to an alternative definition of submarket, we also examine the prediction performance of an OLS model that includes submarkets using pre-specified geographical areas represented by postcodes which are more refined as compared to the submarket defined by real estate brokers according to our data. This is a relevant issue since using more refined submarkets to an alternative definition of submarket may alleviate the spatial dependence among housing transactions, which makes the explicit modeling of hedonic errors to account for the spatial dependence be less of a concern.

Data and Empirical Results

Randstad Housing Transaction Data

Our working data set comprises housing transactions from the Dutch Randstad region. The Dutch Randstad region is a conurbation in the Netherlands, which consists of parts of four Dutch provinces, North Holland, South Holland, Utrecht and Flevoland. The country's largest cities, Amsterdam, Rotterdam, The Hague and Utrecht are all in Randstad, as well as the world's largest port in Rotterdam, one of the European busiest airports at Schiphol, and the major railway terminal in Utrecht. It has a population of seven million which is around 46% of the total Dutch population on 26% of the country's land area, making it one of the most densely populated areas in Europe. The region hosts a wide range of economic activities with 45% of total employment and accounts for a significant share of the Dutch economy with nearly half of the total Dutch GDP generated. Due to its affluence of employment opportunities and limited land area, the Randstad region faces a strong housing demand with 15% at any growth rate and, if rapid growth should occur, even 30% of the current housing supply (van der Burg and Vink 2008). Such high competition on the demand side of the housing market makes the Randstad region the most vibrant housing market in The Netherlands.

The housing transactions in our data set span a period from 1997 to 2007. These transaction data are collected by the Dutch Association of Real Estate Brokers and Real Estate Experts (NVM) which represent approximately 70% of total housing transactions in the Dutch housing market in recent years. The NVM transaction database has also been used by Francke and Vos (2004) and Theebe (2004), where they control for the spatial dependence in addressing different issues in the Dutch housing market. Francke and Vos (2004) model spatial dependence among housing clusters or neighborhoods by spatial error structure with submarket dummy variables in the study of house price index construction. Theebe (2004) adopts the SAR model to control for spatial dependence in the analysis of the impact of traffic noise on house prices.

After removing transactions with missing or unreliable attributes' values, we have access to 437,734 housing transactions, which is one of the largest real estate working data set in the existing literature. Moreover, our data set is very comprehensive in its coverage of housing structural attributes and related information. Specifically, as shown in Table 1, for each transaction, we have information on the transaction price, transaction date, exact location, housing quality, housing attributes, and the submarket to where it belongs. The submarkets are classified on the basis of any region where at least 80% of the moving taking place within itself. This definition of the submarket takes into account the substitutability or homogeneity of houses with respect to their attributes within each submarket as well as consumer tastes and preferences which are normally clustered as well and influenced by unobservable demand side factors. As indicated in Table 1, housing transactions in each of the four broker regions exceed 10% of our total observations. Unsurprisingly, these four submarkets, being broker region 34, 42, 46, and 49, include the four largest cities in the Randstad as well as the whole country, which are Amsterdam, Utrecht, The Hague, and Rotterdam respectively. The house type variable provides detailed information regarding the type of transacted owner occupied dwellings. Houses are classified into 5 different categories from detached house to houses sharing one roof or row house, while

Table 1 Descriptive statistics

Panel A—Continuous variables		
Variable	Mean	Standard Deviation
Size in Square Meters	108.73	42.99
Number of Rooms	4.05	1.40
Log Transaction Price	12.09	0.51
Panel B—Yearly observations		
Year	Proportion in the Sample	Observations
1997	5.4%	23838
1998	7.3%	31744
1999	7.7%	33569
2000	8.3%	36123
2001	9.3%	40609
2002	9.5%	41645
2003	9.6%	41840
2004	10.0%	43770
2005	10.9%	47584
2006	11.2%	48832
2007	11.0%	48180
Panel C—Binary variable frequency		
Variable	Proportion in the Sample	Observations
Parking Possibility	18.5%	80930
Lift	13.9%	60665
Attic	14.9%	65064
Living Room with Sunlight	9.9%	43378
Building year		
1500 to 1905	7.5%	32696
1906 to 1930	16.9%	74136
1931 to 1944	12.2%	53595
1945 to 1959	7.5%	32685
1960 to 1970	12.3%	53804
1971 to 1980	11.7%	51081
1981 to 1990	13.3%	58381
After 1991	18.6%	81356
House/Apartment type		
Detached House	2.4%	10492
Row House	30.6%	133960
Town House	1.1%	4854
Corner House	10.5%	45889
Semi-detached House	4.3%	19031
Ground-level Apartment	7.4%	32229

Table 1 (continued)

Panel C—Binary variable frequency		
Variable	Proportion in the Sample	Observations
Non-ground-level Apartment	15.3%	67099
Maisonnette Apartment	4.3%	18795
Front-door-in-hall Apartment	14.6%	64000
Gallery Apartment	9.0%	39535
Ground-front-door-in-hall Apartment	0.4%	1850
House quality		
Inside Good	89.9%	393600
Outside Good	93.9%	411080
NVM Submarkets classification		
Broker Region 31	1.8%	8034
Broker Region 32	3.0%	13185
Broker Region 33	4.7%	20767
Broker Region 34	16.1%	70289
Broker Region 35	1.0%	4392
Broker Region 36	3.2%	13968
Broker Region 37	5.2%	22778
Broker Region 38	2.5%	11084
Broker Region 39	3.7%	16198
Broker Region 41	2.5%	10870
Broker Region 42	10.4%	45649
Broker Region 44	1.9%	8142
Broker Region 45	3.0%	13314
Broker Region 46	18.8%	82201
Broker Region 47	2.1%	9004
Broker Region 48	1.8%	7832
Broker Region 49	13.0%	56798
Broker Region 50	0.5%	2170
Broker Region 51	1.7%	7514
Broker Region 52	3.1%	13545

apartments are divided into 6 types, ranging from apartments on the ground floor to apartments on the ground floor with front door in the hallway.

We have applied aggregation over the domain of some categorical variables in the creation of dummies due to limited observations within some detailed classifications. For instance, there are nine different descriptions of housing maintenance status ranging from "bad" to "excellent". Since there are very few observations with bad or excellent inside and/or outside maintenance, we create only "good" or "bad" housing quality indicator variables through aggregating from "excellent" to "reasonable" and from "below reasonable" to "bad" respectively. After the transformation, there are around 90% of total observations having good interior and exterior. The same reasoning applies to the transformation of other variables as well, such as "parking possibility" and "living room with sunlight".

Estimation Results

In our empirical analysis, we estimate the unrestricted STAR model as in (5), and X matrix is structured as $X = \begin{pmatrix} D_S & \widetilde{X} \end{pmatrix}$, where D_S is a $n \times I$ matrix storing submarket dummies, and $\widetilde{X}$ is a $n \times J$ matrix of housing structural attributes. We specify the OLS hedonic regression and estimate

$$\ln(\text{Price}) = \alpha + \sum_{i=1}^{I} D_{S,i}\gamma_i + \sum_{j=1}^{J} \widetilde{X}_j\beta_j + \varepsilon \tag{9}$$

where we explain the variation of $n \times 1$ log transaction price using J housing structural attributes represented by $n \times 1$ vectors of $\widetilde{X}_j$ while controlling for spatial heterogeneity via the inclusion of I submarket dummy variables $D_{S,i}$ with broker region 31 as the base case. For building year dummy, we use houses built between 1500 and 1905 as the base, and, for dwelling type dummy, detached house serves as the base. In addition, to control for the possibility of nonlinear relationship between the log price and size and number of rooms, we add size squared and number of rooms squared into the regression.

The main goal of this paper is to examine the prediction performance of the STAR model as compared to the traditional hedonic pricing model. Before our prediction exercise, we first perform analysis based on both models using the full annual sample from 1997 to 2007 to examine their descriptive power. For the sake of space, we present only one example result for the year 1997 in Table 2.[4] Moreover, since we include the submarket dummies as control variables for spatial heterogeneity, we do not report them in the table. The reported standard errors are heteroskedastic robust standard errors following White (1980). We first turn to the OLS results. For the OLS model as a whole, the explanatory power is satisfactory with R squared over 77%. For 1997, one square meter increase in the size of the dwelling will cause the transaction price to rise by roughly 1.3% and the effect is statistically significant. With the coefficient of the size squared term is zero, it seems that the log linear relationship between the log transaction price and size is adequate for this transaction sample. Having a lift in the building will positively affect the transaction by approximately 1.3%, which is reasonable since people are willing to pay for convenience in their living surroundings. Both the number of rooms and its squared term do not have a significant impact on the housing transaction prices. The dwelling age, on the other hand, is an important factor in explaining house price variations. With reference to houses built prior to 1905, all houses built in other years except after 1991 are transacted with a lower price. This is confusing if we take the building year as a benchmark for housing depreciation, and we would expect old houses trading with a significant discount. Our findings seem to suggest that the omitted variables which are highly correlated with dwelling age contribute to the positive age effect. For instance, houses built long time ago may have been artistically designed and

[4] Results for other years are available upon request.

Table 2 Example result for year 1997

Year 1997	OLS model		STAR model	
Structural variables	Coeff.	S.E.	Coeff.	S.E.
Size	0.013[a]	0.000	0.009[a]	0.000
Size squared	0.000	0.000	0.000	0.000
Lift	0.013[b]	0.006	0.040[a]	0.004
Number of rooms	0.013	0.007	0.072[a]	0.005
Number of rooms squared	0.000	0.001	−0.005[a]	0.001
Age1906 to 1930	−0.104[a]	0.008	−0.035[a]	0.006
Age1931 to 1944	−0.117[a]	0.008	−0.009	0.006
Age1945 to 1959	−0.127[a]	0.009	−0.033[a]	0.007
Age1960 to 1970	−0.171[a]	0.008	−0.039[a]	0.006
Age1971 to 1980	−0.123[a]	0.008	0.019[a]	0.006
Age1981 to 1990	−0.029[a]	0.008	0.104[a]	0.006
Age After 1991	0.070[a]	0.008	0.193[a]	0.006
Semi-detached house	−0.332[a]	0.013	−0.337[a]	0.011
Town house	−0.232[a]	0.016	−0.226[a]	0.013
Corner house	−0.294[a]	0.013	−0.296[a]	0.011
Row house	−0.169[a]	0.013	−0.177[a]	0.011
Ground-level apartment	−0.367[a]	0.015	−0.398[a]	0.013
Non-ground-level apartment	−0.432[a]	0.015	−0.474[a]	0.013
Maisonnette apartment	−0.387[a]	0.015	−0.451[a]	0.012
Front-door-in-hall apartment	−0.373[a]	0.014	−0.458[a]	0.012
Gallery apartment	−0.392[a]	0.014	−0.469[a]	0.012
Ground-front-door-in-hall apartment	−0.359[a]	0.034	−0.340[a]	0.027
Attic	−0.018[a]	0.004	0.002	0.003
Living room with sunlight	0.017[a]	0.004	0.012[a]	0.003
Inside good	0.099[a]	0.006	0.104[a]	0.004
Outside good	0.107[a]	0.007	0.076[a]	0.005
Parking possibility	0.094[a]	0.004	0.085[a]	0.003
$ø_T$			−0.162[a]	0.038
$ø_S$			0.875[a]	0.005
$ø_{ST}$			−0.110[a]	0.039
$ø_{TS}$			0.469[a]	0.049
R squared	0.777		0.877	
Loglik	−82420		−75425	
k	47		235	
F statistic			101	
ρ			0.75	

[a], [b] denote statistical significance at 1% and 5% level respectively

decorated that are appreciated nowadays. Dwelling age can also capture locational effects since older homes are near the historical center. In general, relative to the base case, houses built between 1960 and 1970 are the cheapest in the market, while modern constructions are traded with significant premium.

As expected, dwelling types also significantly influence the transaction prices. The detached houses, which serve as the base case, are the most expensive on the market. This makes sense since detached house normally includes a garden and spacious parking, which become a luxury in highly concentrated urban areas with mostly apartment buildings. Moreover, the building costs are high for detached houses as well which naturally lead to higher transaction prices. Comparing with other house types, normal apartments, which are not on the ground floor, are the cheapest. It is puzzling to find a negative sign of the attic dummy, since having an attic provides convenience for storage purposes which should have added value. However, if such utility is not valued, for instance, due to poor design with multiple storage possibilities within the dwelling, or such storage function needs to be shared with others in the case of apartment buildings, we would find a negative effect of having an attic on house prices, which seems likely to be the case here. Unsurprisingly, having a sun-through living room increases the house value, and the same holds for good maintenance for both the interior and exterior of the dwelling. Parking possibility is valued at 9.4% of the house value, which is quite significant. Nonetheless, it seems reasonable if we take into account the high population density within the Randstad region.

Next, we turn to the estimation results of the STAR model. We do not report coefficients of *TX*, *SX*, *STX*, and *TSX*, which are estimated through our specification of the autoregressive process of the hedonic error terms, but do not have a straightforward interpretation as the OLS coefficients (Nappi-Choulet and Maury 2009). A first look at the STAR model results reveal much better in-sample fit with R squared over 87% as compared to less than 78% for the traditional hedonic model. Since we control for spatial heterogeneity in both models using dummies of submarkets defined by real estate brokers, it seems to suggest the existence of spatial and temporal dependence within housing transactions, which contribute to explaining house price variations. Most of the signs of coefficients are consistent with our findings using the OLS model. However, there are few notable exceptions. First, the effect of the number of rooms becomes significant, and its squared term is able to capture the nonlinear relationships between the number of rooms and log transaction price. Second, the order of the impact of building year on the transaction price is changed. In the OLS model, where error terms are taken as independent, we find houses with building year prior to 1905 are the second most expensive on the market next to the most recent constructions. This is no longer the case once we incorporate the spatial and temporal dependence in the housing transactions. The STAR model results show that houses built after 1971 are more expensive as compared with older houses. Overall, neglecting the error correlation in the OLS model leads to larger estimates of the effect of housing construction period on the transaction prices. Third, having an attic is no longer significantly affecting the house prices as compared to the puzzling finding using the OLS model.

We now focus on the spatial and temporal dependence parameter estimates as well as their interactions. All of these parameter estimates are statistically significant, with the spatial dependence parameter ϕ_S exhibiting greater magnitude than the temporal dependence parameter ϕ_T. The statistical significance of the interaction terms *TS* and *ST* suggests their significant impact on the house price determination process. For our

1997 transaction data, the large magnitude of ϕ_{TS}, as compared to ϕ_{ST}, implies the need to filter space first and subsequently for time and not vice-versa, contrary to the findings in Pace et al. (1998 and 2000).

A priori, if there is no spatial and temporal dependence in housing transactions such that $\phi_S = \phi_T = \phi_{ST} = \phi_{TS} = 0$, the STAR model would be reduced to the OLS model. We perform a F-test on the joint significance of the spatial and temporal dependence parameter estimates in the STAR model, and the large F statistic in Table 2 clearly rejects the null hypothesis that these dependence terms are jointly zero. This confirms the existence of both the spatial and temporal dependence in the Randstad housing market.

Our STAR model results implicate the necessity to take into account the previous neighboring transactions in determining the value of the current house in the Randstad region. Our finding of large spatial dependence parameter, which is 0.875, is consistent with those found in other studies being 0.811 in Pace et al. (1998), 0.886 in Pace et al. (2000), and above 0.9 in Militino et al. (2004). On the contrary, Nappi-Choulet and Maury (2009) finds the spatial dependence coefficient to be 0.5 in Paris office market, which seems to indicate residential housing market exhibits stronger spatial dependence than that in the office market. As compared to the spatial dependence parameter, the temporal parameter is of relatively smaller magnitude. This is not surprising, since part of the temporal effect has been absorbed through the specification of the spatial weighting matrix with implicit temporal structure where only previously sold neighbors are considered. In other words, the explicit temporal ordering in the spatial weighting matrix dampens the temporal effect which is supposed to be captured through the specification of the temporal weighting matrix.

Table 3 presents the summary of estimation results using full annual sample for other years from 1998 to 2007. For all the other years, we find large spatial dependence coefficients, all of which are larger than that in 1997. This consistent finding reinforces our earlier conclusion that it is necessary to take into account the previous neighboring transactions in pricing the current house in the Randstad housing market. The temporal dependence parameter estimates, on the other hand, exhibit much greater variations over time as compared to the spatial dependence parameter estimates, ranging from −0.209 in 2004 to 0.289 in 1999, which seems to suggest that the temporal heterogeneity is more relevant to the temporal dependence parameter than to the spatial dependence parameter. We find the optimal spatial decay parameter ρ is not constant over time, but concentrated between 0.75 and 0.8, which are in line with Pace et al. (1998 and 2000). With respect to the standard error of the regression, the STAR model outperforms the OLS model in terms of in-sample fit during our entire sampling period in a consistent manner. The STAR model, on average, reduces the regression standard error by 27% relative to the alternative OLS model.

In addition to examining the in-sample fits using annual full sample as shown in Tables 2 and 3, we also undertake one-step-ahead forecast for both the OLS model and the unrestricted STAR model (5). It is to note that the one-step-ahead forecast is essentially a robustness check of the in-sample fits using a subset of the full sample.[5] Specifically, for every year, we use the first 30%, 60% and 90% of all temporally

[5] See, for example, Plackett (1950).

Table 3 Summary of estimation results from 1998 to 2007

	1998		1999		2000		2001		2002		2003		2004		2005		2006		2007	
	OLS	STAR	OLS	STAR	OLS	STAR	OLS	STAR	OLS	STAR	OLS	STAR	OLS	STAR	OLS	STAR	OLS	STAR	OLS	STAR
$ø_T$		0.088		0.289		0.007		−0.071		−0.069		−0.084		−0.209		−0.045		−0.158		−0.163
$ø_S$		0.900		0.901		0.890		0.883		0.878		0.890		0.896		0.892		0.933		0.933
ρ		0.75		0.75		0.75		0.75		0.75		0.75		0.75		0.80		0.80		0.75
S.E.	0.230	0.164	0.252	0.179	0.245	0.178	0.227	0.163	0.215	0.159	0.206	0.155	0.215	0.159	0.225	0.172	0.231	0.168	0.234	0.163

This table presents the results of in-sample estimates using full annual data. We do not report the coefficients of housing attributes here but only the parameters of interest ordered observations respectively as the in-sample to start one-step-ahead forecast for the next 500 observations. The one-step-ahead forecast results confirm the better in-sample fits of the STAR model relative to the OLS model.[6]

Prediction Results

Our previous analysis demonstrates satisfactory performance of the STAR model relative to the OLS model in terms of in-sample fits. However, what matters more in practice is not how well a model is capable of describing the past, but rather how the model is going to improve our prediction accuracy into the future. Therefore, our main goal is to examine if the STAR model adds to the prediction power relative to the OLS model in predicting future house values. The prediction exercise is performed annually using the first 80% of temporally sorted yearly housing transactions as the in-sample, and the other 20% as the out-of-sample.

Table 4 summarizes the prediction results for each individual year from 1997 to 2007 obtained using the OLS model, and the unrestricted STAR model respectively. The prediction performance is evaluated using the mean absolute error, the median absolute error, the mean squared error (MSE) and the root mean squared error (RMSE) criterion. All of these criteria penalize prediction error in both directions in the prediction error distribution.

On the basis of mean absolute prediction error, the STAR model consistently outperforms the OLS model. On average, the STAR model reduces the OLS mean absolute prediction error by 27.6%. Relative to the OLS model, the best prediction outcome of the STAR model is found to be in 1999 with 32.4% decline in the OLS mean absolute prediction error. For the STAR model, 1998 corresponds to the year with the smallest reduction, which is 22.3%, in the OLS mean absolute prediction error. These findings indicate that the incorporation of both the spatial and temporal dependence in the traditional hedonic price model pays off in obtaining more accurate prediction of future house prices.

We explore further into other prediction performance measures to check the consistency of our findings. For all the years in our sampling period, the STAR model is capable of reducing the standard deviation of the prediction error which is measured by RMSE substantially as compared to the OLS model, ranging from a drop of 21.5% in 1998 to 27% in 2000. On average, the RMSE for the OLS model is 0.23, and is 0.17 for the STAR model. The OLS model produces much larger variation of the prediction error which is between 0.204 and 0.274 than that generated using the STAR model which falls within the range of 0.158 and 0.2. The mean reduction of the prediction error variation using the STAR relative to the OLS model is around 24.2%. Prediction performance based on the median absolute error also favors the STAR model over the OLS model. The average median prediction error for the OLS model is 0.134 as compared to that of the STAR model, which is 0.094.

Overall, our main prediction exercise illustrates that integrating both the spatial and temporal dependence among housing transactions in the empirical analysis is not only a warranted theoretical concept but also contributes to a better prediction outcome of future house values. It is to note that our findings run counter to Bourassa

[6] For the sake of space, results are not included in the paper and are available upon request.

Table 4 Prediction results

Prediction error	1997			1998			1999			2000			2001			2002	
	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR
Mean |error|	0.1732	0.1256	0.1539	0.1824	0.1418	0.1489	0.2024	0.1369	0.1537	0.1854	0.1303	0.1453	0.1662	0.1201	0.1350	0.1544	0.1157
Median |error|	0.0164	0.0226	0.1134	0.0570	0.0792	0.1141	0.0780	0.0280	0.0542	0.0415	0.0307	0.0395	0.0201	0.0019	0.0998	0.0019	0.0143
MSE	0.0533	0.0297	0.0449	0.0565	0.0348	0.0404	0.0749	0.0401	0.0558	0.0598	0.0317	0.0405	0.0487	0.0273	0.0355	0.0418	0.0249
RMSE	0.2310	0.1722	0.2119	0.2376	0.1866	0.2011	0.2737	0.2002	0.2363	0.2446	0.1781	0.2013	0.2207	0.1653	0.1884	0.2046	0.1576
N(Postcode regions)			246			319			321			367			359		

Prediction error	2002	2003			2004			2005			2006			2007		
	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)
Mean |error|	0.1274	0.1555	0.1169	0.1276	0.1613	0.1176	0.1268	0.1635	0.1193	0.1266	0.1721	0.1231	0.1299	0.1759	0.1204	0.1310
Median |error|	0.0969	0.0084	0.0133	0.0965	0.0092	0.0110	0.0956	0.0122	0.0164	0.0929	0.0157	0.0296	0.0985	0.0100	0.0157	0.0961
MSE	0.0301	0.0436	0.0267	0.0309	0.0484	0.0293	0.0329	0.0506	0.0310	0.0349	0.0545	0.0307	0.0347	0.0576	0.0319	0.0375
RMSE	0.1735	0.2088	0.1635	0.1757	0.2199	0.1712	0.1813	0.2250	0.1761	0.1868	0.2334	0.1753	0.1863	0.2400	0.1786	0.1936
N(Postcode regions)	368			381			379			400			409			409

This table presents the prediction results obtained using the first 80% of temporally ordered yearly housing transactions as the in-sample and 80–100% of temporally ordered yearly housing transactions as the out-of-sample

et al. (2007) that favors the OLS model over the spatial model on the basis of their prediction outcome. This may be attributed to the amount of information used in undertaking house price prediction. In their prediction exercise, they utilize only the in-sample parameter estimates and leave the spatial information unexplored in the residuals. On the contrary, when making out-of-sample predictions, we not only use the in-sample parameter estimates, but also take into account the pricing information contained in spatial and temporal weighting matrices. In essence, the extra information we use in our prediction exercise are the pricing information on prior spatial neighbors in the residuals that have been neglected in Bourassa et al. (2007).

We undertake two robustness checks of our main prediction results. In the first robustness check, we include an alternative specification of submarkets defined by postcode areas in the OLS model and repeat our main prediction exercise to examine if our prediction results are sensitive to a different and more refined submarket definition. The results are incorporated in Table 4. It is interesting to note that the prediction performance of the OLS model including submarkets defined by postcode areas is better than that of the OLS model using experience-based submarket definitions, which is consistent across the entire sampling period. It suggests that, at least for the Dutch Randstad housing market, using submarket defined by real estate professionals performs worse in terms of predicting future house values than the alternative submarket definition of postcode areas.

However, the STAR model consistently outperforms the OLS model including postcode submarket dummies for all the years from 1997 to 2007. In general, our first robustness check implies that including more refined postcode submarket dummies within the OLS model reduces the effect of spatial dependence among housing transactions within the Randstad region in predicting future house values to a certain extent. Accounting for spatial dependence among housing transactions is still a valid exercise for a better prediction outcome.

In the second robustness check, we check if our prediction results are driven by different in-sample and out-of-sample specifications. First, we alter the out-of-sample used for prediction. We split the original 80–100% out-of-sample into two out-of-samples containing 80–90% and 90–100% of total temporally ordered yearly observations respectively. The results are shown in Tables 5 and 6. The outperformance of the STAR model in predicting future house values relative to both OLS models with different submarket proxies is confirmed using both out-of-samples on the basis of all three evaluation criterion. Second, we focus on the in-sample specification. We extend our in-sample from using the first 80% of all yearly observations with temporal ordering to using the first 90% of all observations. Prediction performance is evaluated using 90–100% of all observations as the out-of-sample. Table 7 presents the results. Again, the STAR model demonstrates better performance in house value prediction relative to both OLS models. In sum, our second robustness check results reinforce our earlier findings regarding the better prediction ability of the STAR model which are not driven by the specifications of either out-of-sample or in-sample.

A Parsimonious Model and House Price Index Construction

As a natural application that exploits the better predictive power of the STAR model, we extend our analysis into the house price index construction. For practical

Table 5 Robustness check of the prediction results using different out-of-sample 1

Prediction error	1997			1998			1999			2000			2001			2002	
	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR
Mean |error|	0.1736	0.1274	0.1541	0.1828	0.1412	0.1470	0.2037	0.1375	0.1506	0.1855	0.1316	0.1452	0.1667	0.1188	0.1351	0.1550	0.1168
Median |error|	0.0218	0.0246	0.1157	0.0557	0.0820	0.1138	0.0750	0.0313	0.0000	0.0429	0.0295	0.0368	0.0229	0.0012	0.1017	0.0009	0.0153
MSE	0.0529	0.0294	0.0437	0.0571	0.0331	0.0401	0.0769	0.0417	0.0543	0.0610	0.0329	0.0407	0.0482	0.0259	0.0349	0.0420	0.0257
RMSE	0.2300	0.1714	0.2091	0.2390	0.1818	0.2002	0.2773	0.2043	0.2329	0.2469	0.1815	0.2016	0.2195	0.1609	0.1867	0.2048	0.1602
N(Postcode regions)			246			319			321			367			359		

Prediction error	2002	2003			2004			2005			2006			2007		
	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)
Mean |error|	0.1280	0.1536	0.1173	0.1274	0.1605	0.1182	0.1266	0.1616	0.1191	0.1269	0.1704	0.1212	0.1287	0.1761	0.1203	0.1309
Median |error|	0.0973	0.0053	0.0131	0.0955	0.0141	0.0097	0.0948	0.0113	0.0168	0.0941	0.0168	0.0301	0.0987	0.0083	0.0136	0.0949
MSE	0.0303	0.0431	0.0274	0.0315	0.0470	0.0284	0.0318	0.0488	0.0306	0.0345	0.0534	0.0299	0.0340	0.0583	0.0315	0.0378
RMSE	0.1741	0.2077	0.1655	0.1775	0.2169	0.1684	0.1782	0.2209	0.1749	0.1858	0.2311	0.1728	0.1843	0.2416	0.1774	0.1943
N(Postcode regions)	368			381			379			400			409			409

This table presents the prediction results obtained using the first 80% of temporally ordered yearly housing transactions as the in-sample and 80–90% of temporally ordered yearly housing transactions as the out-of-sample

Table 6 Robustness check of the prediction results using different out-of-sample 2

Prediction error	1997			1998			1999			2000			2001			2002	
	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR
Mean |error|	0.1728	0.1237	0.1536	0.1821	0.1425	0.1507	0.2011	0.1362	0.1568	0.1852	0.1290	0.1455	0.1657	0.1213	0.1349	0.1539	0.1146
Median |error|	0.0116	0.0196	0.1104	0.0595	0.0765	0.1145	0.0816	0.0257	0.0542	0.0402	0.0316	0.0421	0.0166	0.0038	0.0986	0.0054	0.0137
MSE	0.0538	0.0300	0.0461	0.0558	0.0366	0.0408	0.0730	0.0384	0.0574	0.0587	0.0304	0.0406	0.0492	0.0287	0.0361	0.0417	0.0240
RMSE	0.2319	0.1731	0.2146	0.2362	0.1912	0.2019	0.2702	0.1960	0.2396	0.2423	0.1745	0.2014	0.2218	0.1695	0.1900	0.2043	0.1550
N(Postcode regions)			246			319			321			367			359		

Prediction error	2002	2003			2004			2005			2006			2007		
	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)
Mean |error|	0.1268	0.1575	0.1166	0.1277	0.1620	0.1169	0.1269	0.1654	0.1195	0.1263	0.1738	0.1250	0.1310	0.1757	0.1205	0.1310
Median |error|	0.0967	0.0125	0.0136	0.0979	0.0031	0.0121	0.0966	0.0127	0.0158	0.0918	0.0151	0.0290	0.0982	0.0121	0.0175	0.0972
MSE	0.0299	0.0440	0.0261	0.0302	0.0497	0.0303	0.0340	0.0524	0.0315	0.0352	0.0556	0.0316	0.0354	0.0568	0.0323	0.0372
RMSE	0.1729	0.2098	0.1615	0.1739	0.2230	0.1740	0.1843	0.2289	0.1774	0.1877	0.2358	0.1778	0.1883	0.2384	0.1799	0.1929
N(Postcode regions)	368			381			379			400			409			409

This table presents the prediction results obtained using the first 80% of temporally ordered yearly housing transactions as the in-sample and 90–100% of temporally ordered yearly housing transactions as the out-of-sample

Table 7 Robustness check of the prediction results using different in-sample

Prediction error	1997			1998			1999			2000			2001			2002	
	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR
Mean |error|	0.1713	0.1222	0.1493	0.1803	0.1280	0.1477	0.1981	0.1382	0.1160	0.1845	0.1273	0.1444	0.1653	0.1227	0.1329	0.1535	0.1131
Median |error|	0.0118	0.0157	0.1076	0.0524	0.0308	0.1135	0.0725	0.0399	0.0083	0.0346	0.0246	0.0378	0.0146	0.0269	0.0968	0.0040	0.0059
MSE	0.0525	0.0297	0.0438	0.0549	0.0307	0.0394	0.0712	0.0392	0.0459	0.0584	0.0307	0.0401	0.0490	0.0289	0.0349	0.0416	0.0237
RMSE	0.2291	0.1723	0.2093	0.2344	0.1753	0.1986	0.2668	0.1981	0.2141	0.2417	0.1752	0.2003	0.2215	0.1701	0.1867	0.2039	0.1538
N(Postcode regions)			266			336			341			367			374		

Prediction error	2002	2003			2004			2005			2006			2007		
	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)	OLS	STAR	OLS (Postcode)
Mean |error|	0.1259	0.1573	0.1156	0.1262	0.1620	0.1160	0.1242	0.1652	0.1178	0.1252	0.1737	0.1216	0.1299	0.1756	0.1227	0.1302
Median |error|	0.0960	0.0116	0.0067	0.0971	0.0023	0.0130	0.0940	0.0120	0.0116	0.0910	0.0134	0.0167	0.0981	0.0100	0.0315	0.0965
MSE	0.0296	0.0439	0.0252	0.0296	0.0497	0.0300	0.0323	0.0523	0.0309	0.0348	0.0555	0.0306	0.0350	0.0568	0.0329	0.0368
RMSE	0.1719	0.2096	0.1588	0.1720	0.2229	0.1733	0.1797	0.2287	0.1759	0.1865	0.2355	0.1750	0.1872	0.2383	0.1814	0.1918
N(Postcode regions)	379			395			401			412			421			421

This table presents the prediction results obtained using the first 90% of temporally ordered yearly housing transactions as the in-sample and 90–100% of temporally ordered yearly housing transactions as the out-of-sample

purposes, we need to search for a parsimonious model that balances the performance and the reduction of dimensionality.

In light of the discrepancy in terms of the magnitude and consistency between the spatial and temporal dependence terms as shown in Table 3, we examine the performance of a simpler model that only includes the spatial dependence term as compared to the full blown STAR model on the basis of both in-sample fits and out-of-sample predictions. The parsimonious model consists of 94 variables relative to 235 variables in the STAR model. As shown in Table 8, the parsimonious model performs satisfactorily though marginally worse in terms of the in-sample fits than that of the STAR model across all the individual years in our sampling period as expected. However, the performance of the parsimonious model in terms of out-of-sample prediction power is better than that of the STAR model for most of the years in our sampling period except 2000 and 2004. Overall, the reduction of dimensionality and the increase of the degrees of freedom of the parsimonious model loses its in-sample fits, but gains in out-of-sample predictions.

We construct the market house price index following a two-step procedure as follows. First, we randomly sample 1,000 properties from transactions within 1997. Second, we predict each of the 1,000 properties in our random sample on the basis of the parsimonious model (10) at the end of each year in our sampling period,

$$\widehat{Y}_t = \widehat{\alpha} + X\widehat{\beta}_1 + S_t X\widehat{\beta}_2 + \widehat{\phi}_S S_t Y_t \qquad (10)$$

where $\hat{Y}_t$ is the predicted log house price at the end of each year, and S_t is the spatial matrix storing the updated spatial neighbors of the predicted house with the reference point at the end of each year. The house price index is constructed through chaining the average predicted price changes of the properties in the random sample.

Figure 1 demonstrates the market price indices built using both the parsimonious model and the OLS model. As expected, the OLS market index exhibits less volatility relative to the market index constructed using the parsimonious model. This is largely attributed to the fact that, in addition to the updating of the estimated structural parameters in predicting property value in the random sample as the case for the OLS model, the parsimonious model also updates the spatial weighting matrix storing the spatial neighbors on a yearly basis. Since the spatial neighbors of a particular property differs from one year to another, the standard deviation of the rate of price change in the random sample is larger for the parsimonious model than that for the OLS model.

In terms of index level, there is persistent divergence between the two indices and the OLS model reports smaller index numbers relative to parsimonious model across the years. As demonstrated above, the better predictive power of the parsimonious model implicates consistent understatement of the housing market price development that is derived using an OLS model for the period from 1997 to 2007. Through dividing our random sample on the basis of location, we construct local house price indices for four largest broker regions within Randstad respectively.[7] These indices are shown in Fig. 2. Among the four broker regions, broker region 34, where Amsterdam is included, has experienced much more volatile house price development over the period from 1997

[7] Broker region 34 includes Amsterdam. Broker region 42 includes Utrecht. Broker 46 includes The Hague. Broker region 49 includes Rotterdam.

Table 8 Performance of the parsimonious model

	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
In-sample R squared											
Parsimonious model	0.8752	0.8742	0.8734	0.8740	0.8828	0.8809	0.8808	0.8803	0.8619	0.8790	0.8917
STAR model	0.8766	0.8763	0.8765	0.8780	0.8836	0.8814	0.8814	0.8808	0.8624	0.8797	0.8923
Out-of-sample RMSE											
Parsimonious model	0.1717	0.1690	0.1987	0.2077	0.1651	0.1573	0.1627	0.1716	0.1755	0.1730	0.1781
STAR model	0.1722	0.1866	0.2002	0.1783	0.1653	0.1576	0.1635	0.1712	0.1761	0.1753	0.1786

This table presents the results of the performance of parsimonious model relative to the STAR based on both the in-sample fits and out-of-sample prediction performance. The out-of-sample RMSE is obtained using the first 80% of temporally ordered yearly housing transactions as the in-sample and 80–100% of temporally ordered yearly housing transactions as the out-of-sample

to 2007, as compared to other three broker regions. Moreover, it is also the region that has enjoyed the largest price increase within the 10-year period.

Examining house price indices produced using the OLS model and the parsimonious model on a local level, we observe discrepancy between these two relative indices across all four broker regions under consideration. In particular, relative to house price indices constructed using the parsimonious model, there is persistent understatement in terms of index level for indices produced using the OLS model. In terms of index volatility, house price indices built using the parsimonious model exhibit more volatility than the OLS indices for all four broker regions. Overall, these local level results are consistent with our earlier market level findings.

It is to note that one should not overemphasize and generalize the house price indices produced using only one random draw to obtain a sample of properties to represent the entire Randstad housing market as shown in this paper. If one is interested in the reliability of the house price indices produced above, which is not a focal point of this paper, multiple random sampling of housing transactions will make it possible to establish confidence bounds around these index numbers.

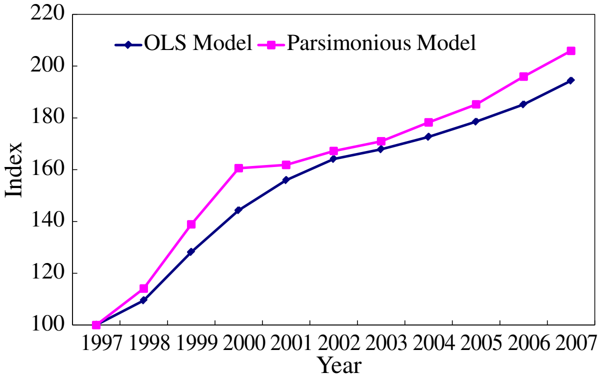

Fig. 1 Randstad housing market index

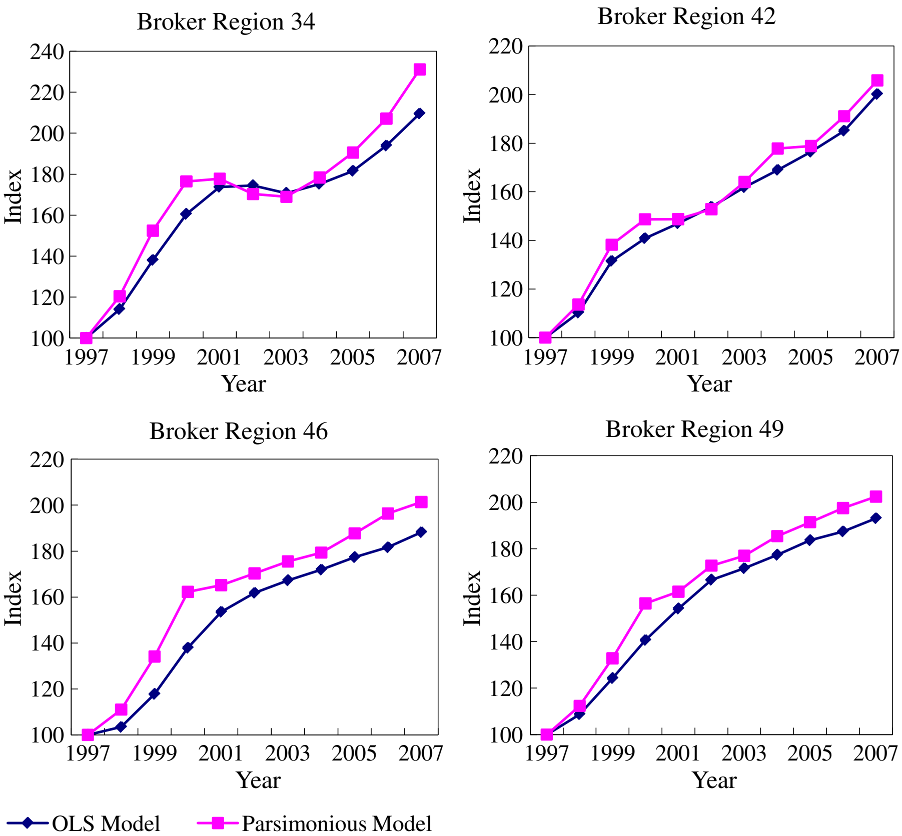

Fig. 2 Local house price indices (OLS model versus Parsimonious model)

Conclusion

This paper studies the impact of accounting for error correlations within the setup of traditional hedonic pricing model on predicting future house prices using a large non-U.S. housing transaction data set. We follow the STAR model in Pace et al. (1998) which subsumes the errors in the hedonic regression following an autoregressive process. Both the spatial and temporal dependence in the error structure are taken into account via explicit incorporation of spatial and temporal weighting matrices and their interactions. We control for spatial heterogeneity using experience-based definition of submarkets, and for temporal heterogeneity through performing analysis on an annual basis. Moreover, in our prediction exercise, we recognize the temporal nature in the housing market, and integrate the information contained in both the spatial and temporal neighbors in predicting the future house values. As a natural extension, we exploit our findings in the context of house price index construction.

Our results indicate the existence of both the spatial and temporal dependence in the Randstad housing market. Moreover, the spatial dependence is of much larger magnitude as compared to the temporal dependence, which is consistent over all the individual year during our sampling period. This finding is comparable to a number of previous spatial studies using housing transaction data from other countries, such as the U.S. and Spain. Our prediction results show better prediction capability of the STAR model relative to the OLS model, which are robust to alternative specifications of in-sample as well as out-of-sample. Therefore, it is necessary to correct for both the spatial and temporal dependence among housing transactions in the empirical applications of the hedonic pricing model for predictive purposes. Examining the consequence of exploiting the better predictive power of the STAR model in constructing house price indices on the basis of single random sample of housing transactions used to represent the overall Randstad housing market, we show that the OLS index consistently understates the more accurate house price development as captured using the parsimonious model.

Our results have implications in practice if the aim is to produce better forecast of house values, and accounting for the effect of both the spatial and temporal dependence among housing transactions can substantially reduce the prediction error. In addition to constructing house price indices, this is relevant to, for example, banks providing mortgage which need to access the associated risk via loan-to-value ratio, and rating agencies to improve the rating accuracy of structured products, such as mortgage backed securities.

Acknowledgements The author is grateful to Peter Englund, Marc Francke and Martin Hoesli for their constructive and valuable comments and suggestions. He would also like to thank the Dutch Association of Real Estate Brokers and Real Estate Experts (NVM) for providing the data for this research. In addition, he thanks the editor and the anonymous referee of this journal for their comments. Any remaining errors are the author's responsibility.

References

Basu, A., & Thibodeau, T. (1998). Analysis of spatial autocorrelation in house prices. *Journal of Real Estate Finance and Economics, 17*(1), 61–85.

Bourassa, S., Hamelink, F., Hoesli, M., & MacGregor, B. (1999). Defining housing submarkets. *Journal of Housing Economics, 8*(2), 160–183.

Bourassa, S., Hoesli, M., & Peng, V. (2003). Do housing submarkets really matter? *Journal of Housing Economics, 12*(1), 12–28.

Bourassa, S., Cantoni, E., & Hoesli, M. (2007). Spatial dependence, housing submarkets, and house price prediction. *Journal of Real Estate Finance and Economics, 35*(2), 143–160.

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.

Case, B., Clapp, J., Dubin, R., & Rodriguez, M. (2004). Modeling spatial and temporal house price patterns: a comparison of four models. *Journal of Real Estate Finance and Economics, 29*(2), 167–191.

Casetti, E. (1972). Generating models by the expansion method: applications to geographical research. *Geographical Analysis, 4*(1), 81–91.

Dubin, R. (1998). Spatial autocorrelation: a primer. *Journal of Housing Economics, 7*(4), 304–327.

Dubin, R., Pace, K., & Thibodeau, T. (1999). Spatial autoregression techniques for real estate data. *Journal of Real Estate Literature, 7*(1), 79–96.

Dunse, N., Leishman, C., & Watkins, C. (2001). Classifying office submarkets. *Journal of Property Investment and Finance, 19*(3), 236–250.

Dunse, N., Leishman, C., & Watkins, C. (2002). Testing for the existence of office submarkets: a comparison of evidence from two cities. *Urban Studies, 39*(3), 483–506.

Francke, M., & Vos, G. (2004). The hierarchical trend model for property valuation and local price indices. *Journal of Real Estate Finance and Economics, 28*(2/3), 179–208.

Gelfand, A., Ecker, M., Knight, J., & Sirmans, C. (2004). The dynamics of location in home price. *Journal of Real Estate Finance and Economics, 29*(2), 149–166.

Knight, J., Dombrow, J., & Sirmans, C. (1995). A varying parameters approach to constructing house price indexes. *Real Estate Economics, 23*(2), 187–205.

Michaels, G., & Smith, K. (1990). Market segmentation and valuing amenities with hedonic models: the case of hazardous waste sites. *Journal of Urban Economics, 28*(2), 223–242.

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.

Munneke, H., & Slade, B. (2001). A metropolitan transaction-based commercial price index: a time-varying parameter approach. *Real Estate Economics, 29*(1), 55–84.

Nappi-Choulet, I., & Maury, T. (2009). A spatiotemporal autoregressive price index for the Paris office property market. *Real Estate Economics, 37*(2), 305–340.

Pace, K. (1997). Performing large spatial regressions and autorergressions. *Economics Letters, 54*(3), 283–291.

Pace, K., & Barry, R. (1997a). Quick computation of spatial autoregressive estimators. *Geographical Analysis, 29*(3), 232–247.

Pace, K., & Barry, R. (1997b). Sparse spatial autoregressions. *Statistics and probability letters, 33*(3), 291–297.

Pace, K., Barry, R., Clapp, J., & Rodriguez, M. (1998). Spatiotemporal autoregressive models of neighborhood effects. *Journal of Real Estate Finance and Economics, 17*(1), 15–33.

Pace, K., Barry, R., Gilley, O., & Sirmans, C. (2000). A method for spatial-temporal forecasting with an application to real estate prices. *International Journal of Forecasting, 16*(2), 229–246.

Palm, R. (1978). Spatial segmentation of the urban housing market. *Economic Geography, 54*(3), 210–221.

Plackett, R. (1950). Some theorems in least squares. *Biometrika, 37*(1/2), 149–157.

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

Theebe, M. (2004). Planes, trains, and automobiles: the impact of traffic noise on house prices. *Journal of Real Estate Finance and Economics, 28*(2/3), 209–234.

van der Burg, A., & Vink, B. (2008). Randstad Holland 2040. 44th *ISOCARP* Congress.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica, 48*(4), 817–838.