A Comprehensive Spatiotemporal Framework for Hedonic Pricing: Integrating the Comparable Sales Approach and Reducing Omitted Variable Bias

Document Version
Submitted manuscript

Link to publication record in Manchester Research Explorer

Citation for published version (APA):

Citing this paper
Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights
Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy
If you believe that this document breaches copyright please refer to the University of Manchester’s Takedown Procedures [http://man.ac.uk/04Y6Bo] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.
A Comprehensive Spatiotemporal Framework for Hedonic Pricing: Integrating the Comparable Sales Approach and Reducing Omitted Variable Bias

Sotirios Thanos ¹, Jean Dubé ², Diègo Legros ³,

Abstract
This paper develops a theoretical and methodological framework that integrates Hedonic Pricing (HP), grid comparable sales approach (CSA), and nearest neighbors into a general spatiotemporal specification. By explicitly providing a theoretical justification for introducing spatial (or spatiotemporal) econometrics to HP, this approach is not only relevant to house price forecasting and automated valuation models (AVM) but also to valuing environmental goods capitalized in housing and to all other fields employing house pricing models. The resulting econometric CSA and a spatiotemporal Durbin model provide higher prediction accuracy to alternatives and minimize the spatially-delineated omitted variable bias (OVB) common in HP. Spatiotemporal autoregressive (STAR) and error models (STEM) are also derived, providing specific conditions, under which their application can be justified. Our analysis reinforces the common real estate practice of selecting a small number of comparables in grid CSA and challenges AVM approaches, in which hundreds or thousands of comparables are introduced.

KEY WORDS: Comparable Sales; Hedonic Pricing; Spatial Econometrics; Automated Valuation Models (AVM); Nearest Neighbors; Spatiotemporal Models; STAR.

JEL Code: D62, C21, C40, Q51, R10, R21, R31

The copyright of this Working Paper remains with the authors. If you wish to quote or cite this paper please contact the corresponding author. A more recent version of the paper may have been published elsewhere. CC licensing: CC BY-NC-ND 4.0

¹ Corresponding author. Address: Manchester Urban Institute, Bridgeford Street Building, University of Manchester, M13 9PL, UK. Tel: +44(0)161-306-6439. Email: sotirios.thanos@manchester.ac.uk
² Centre de Recherche en Aménagement et Développement, Laval University, Canada
³ Laboratoire d’Économie et de Gestion, Université de Bourgogne-Franche Comté, France
1. Introduction

The capability of predicting house prices or values effectively, efficiently, and with precision is of paramount importance not only for local housing markets and the real estate sector but also for national economies and world economic stability (Eriksen et al., 2019). Eriksen et al. (2019) stress how the appraisal bias of residential property prices played a key role to the housing market crash and subsequent recession initiated in the US housing market (Martin, 2011). Precise house valuations/price-predictions are essential for key economic tools and activities, such as tax estimates, price indices, insurance policies, and investment portfolios. These activities and tools are indispensable to a range of economic actors from individual consumers, companies, and large corporations to the government and wider public sector (Glumac and Des Rosiers, 2018).

Not surprisingly, there has been a shift in attitudes regarding house valuation tools over the last decade, moving from grid comparable sales and relatively simpler statistical techniques to advanced spatial econometric approaches (Bidanset and Lombard, 2014; Borst and McCluskey 2007, 2008), and to the incorporation of big data and machine learning in the process (McCluskey, et al., 2012; Kok et al., 2017). This shift in perspective is certainly not limited to academic research, as illustrated by Zillow’s (2017) competition to beat their in-house automated valuation model (AVM) with prizes of 1.2 million US dollars awarded to the participants with the best price predictions. Nevertheless, Eriksen et al. (2019) identify the weights assigned to selected comparable transactions after adjusting for observable attributes to be the primary source of price appraisal bias. This issue is directly relevant to AVMs and a focus point of this paper.

Past sales is directly or indirectly one of the main sources of information, no matter the technique employed for price prediction or house valuation. Employing past house sales data, Hedonic Pricing (HP) has solid grounding in economic theory (Rosen, 1974) and is the key method for house price predictions and for valuing non-market goods, such as environmental, transportation, and other neighborhood amenities capitalized in housing. Nevertheless, HP estimates suffer from several types of biases (Kuminoff and Pope, 2014)
that have driven researchers to develop quasi-experimental approaches and boundary discontinuity designs based on HP (Greenstone and Gayer, 2009; Black, 1999; Chay and Greenstone, 2005). Conversely to HP, the comparable sales approach (CSA) has mainly been a practitioner’s tool that effectively values houses under informational and computational constraints with scant basis on economic theory. There has been a number of papers (Kang and Reichert, 1991; Pace, 1998; Pace and Gilley, 1998) that sought to merge the HP and CSA frameworks in order to improve the performance of the former and the theoretical grounding of the latter.

Pace (1998) and Pace and Gilley (1998) did derive a spatial error model (SEM) by combining HP and CSA frameworks. However, that the data are treated as strictly cross-sectional creates some problems. It has thereafter been established in a rapidly expanding literature that house sale data are spatio-temporal in nature (Can and Megbolugbe, 1997; Pace et al., 1998 and 2000; Huang et al., 2010; Thanos et al., 2012 and 2015; Dubé and Legros, 2013a, Dubé and Legros, 2014; Nase et al., 2016; Dubé et al., 2018; Hyun and Milcheva, 2018). When only the spatial aspect of the data is modelled and the temporal aspect is ignored, arrow of time violations arise (Thanos et al., 2016). In the context of CSA, houses sales that will be occurring in the future cannot possibly be used as comparables. Nevertheless, an extensive spatial econometric literature applied on housing data implicitly assumes just that. Spatial econometric methods largely stem from the disciplines of Geography and Regional Science, and as such are geared for aggregate geographical analysis (Anselin and Griffith, 1988) and ill-suited to disaggregate housing data and HP.

Pace et al. (1998) and Can and Megbolugbe (1997) were the first to start a new research stream by adjusting the spatial econometric paradigm to take into account the distinct spatiotemporal nature of housing data and the resulting HP models. However, Can and Megbolugbe (1997) only mention CSA and do not derive spatiotemporal models from the

---

4 House sales data do not have a spatial panel data format, as individual observations occur in a specific place at a particular time (Thanos et al., 2016).
5 We can cite upon request at least 50 peer reviewed papers that commit arrow of time violations through spatial econometric applications to house sale data.
combination of HP and CSA. The influential Pace et al. (1998) paper explicitly states that it extends the Pace and Gilley (1998) framework based on CSA to derive an STDM type of specification. Despite the novelties and the subsequent innovations it sparked in the field, the Pace et al. (1998) specification leaves room for a number of further improvements, which we seek to address in this paper.

Even though the theoretical justification of including spatial or spatiotemporal autoregressive terms in HP can only be based on CSA, a considerable gap of deriving all the widely used spatial econometric specifications from a theoretically sound basis remains in the literature. This paper incorporates and further develops recent advances in spatiotemporal models to provide a framework that combines CSA and HP and derives a range of spatiotemporal HP specifications. These specifications include Spatiotemporal Autoregressive (STAR), econometric CSA, Spatiotemporal Durbin (STDM), and Spatiotemporal Error (STEM) models. The comparables are explicitly accounted for and provide the framework for all the spatiotemporal terms in econometric CSA, STAR, STDM, and STEM. In the process of achieving this, we meticulously account for the grid CSA in existing real estate professional practice, providing a parsimonious exposition of the process. This helps to draw a clear distinction between two types of CSA weights: house attribute weights; and “comparability” or nearest neighbor weights. This in turn allows us to show that the econometric CSA derives from HP and complies with its theoretical imperatives under certain conditions.

Another key contribution emerging from the combination of CSA and HP rests in deriving specific conditions for realizing reductions in spatially-delineated omitted variable bias (OVB) that is common in HP (Kuminoff and Pope, 2014). Furthermore, we provide the conditions under which the STAR application is justified instead of STDM. We also use out-of-sample estimation to compare the price prediction performance among HP, nearest neighbors (NN), econometric CSA, STDM, STAR, and STEM; this comparison provides key insights to the practice of house price prediction and AVMs.
The structure of the paper is as follows. Section 2 provides the methodological framework that combines HP with CSA and derives the spatiotemporal models. Section 3 presents the data and discusses the estimation procedure. Section 4 presents the results of the statistical models and compares their prediction performance. Section 5 draws conclusions and discusses further research.

2. Theoretical and Methodological Framework

The setup of the methodological framework starts with the exposition of the grid CSA and proceeds with the theoretical framework. This prepares the ground for introducing CSA into HP and for deriving the basis of the spatiotemporal models. This is followed by an examination of the approaches to the selection of comparables, which informs the specification of the spatiotemporal weight matrices. The combination of the weight matrices with the CSA-HP basis of spatiotemporal models allows the complete derivation of STAR, STDM, and STEM.

2.1. Grid comparable sales approach

The real estate professional wants an estimate for the price \( y \) of house \( h \), for which only \( \kappa \) characteristics, \( x_{\kappa} \), are observed as shown in equation 1.

\[
y_h = \alpha + \sum_{k=1}^{K} \varphi_k x_{hk}
\]  

(1)

Where the parameter \( \varphi_k \) is the estimate of the impact of housing characteristic \( \kappa \) on the price of house \( h \).

From a practical point of view, the real estate professional does not usually employ any statistical models, only arbitrary/expert knowledge of \( \varphi_s \). Equation 1 cannot provide a price estimate for the real estate professional, as the constant is unknown, and she/he does not have a complete map of all characteristics and their relative weightings. However, past sale
information is available. House \(i\) in equation 2 was sold recently, and it is in close proximity and comparable to house \(h\).

\[
y_i = \alpha + \sum_{k=1}^{K} \varphi_k x_{ik}
\]  

Substituting equation 2 into 1 provides a way of expressing the sale price of house, \(h\), as a function of a comparable past sale \(i\):

\[
y_h = \alpha + \sum_{k=1}^{K} \varphi_k x_{ik} + [y_i - \alpha - \sum_{k=1}^{K} \varphi_k x_{ik}] \iff y_h = y_i + \sum_{k=1}^{K} \varphi_k (x_{hk} - x_{ik})
\]  

Equation 3 provides the real estate professional the ability to value house \(h\), even under significant informational and computational constraints. If houses \(h\) and \(i\) are similar and in close spatial proximity, all characteristics depending on locality drop out, as do the constants. The only information required for a price estimate is expert knowledge of the weights (\(\varphi_k\)) on limited number of the housing characteristics that diverge between houses \(h\) and \(i\). The weights (\(\varphi_k\)) can be additive or multiplicative (percentage changes or elasticities) depending on the specification of the functional (i.e. linear, log-linear, double-log).

Equation 3 was first introduced by Colwell et al. (1983) and employed widely in the real estate literature to show the similarity between the CSA and HP. However, equation 3 holds only for one comparable. By introducing \(c\) comparable sales \((c = 1, \ldots, C)\), it is transformed into the format shown in equation 4.

\[
y_h = (1 - c) \alpha + \sum_{c=1}^{C} y_c + \sum_{k=1}^{K} \varphi (x_{hk} - \sum_{c=1}^{C} x_{ck})
\]  

Pace and Gilley (1998) start from equation 4 and introduce thereafter two additional types of weights on top of \(\varphi\) for each \(x\) to complete the algebraic derivation, but with limited theoretical justification for the extra weights. Pace and Gilley (1998) and Pace (1998) fall into the time-travel problem of using spatial econometrics for spatiotemporal housing data, as is also typical for subsequent papers in the field (Borst and McCluskey, 2007). This issue is fundamental, since the complete map of relations between comparables and price estimates cannot be specified when interrelationship is assumed between house \(i\) and house \(h\), as house \(h\) was sold after house \(i\) and price information cannot travel backward in time.
(Thanos et al., 2016). Using the Pace and Gilley (1998) derivation, Pace et al. (1998, 2000) were one of the first papers to rectify the “time-travel” problem, but we need to discuss their approach in section 2.4 after we define the spatiotemporal weight matrices.

2.2. Theoretical framework of Integrating HP and CSA

This section provides the theoretical framework for integrating HP and CSA. Rosen’s (1974) model conditions and its assumptions about consumer and producer/seller behavior are the inevitable starting point: production and consumption of potentially continuous quantities of housing attributes; market participants with no market power; perfect information; differentiable cost and utility functions; myopic consumers; and free mobility. Under these conditions, Rosen (1974) demonstrates that the marginal price function for a housing attribute equals the buyer’s willingness to pay for a marginal change in that attribute (Kuminoff and Pope, 2014). Palmquist (1992) argues that in the special case of “localized” change, it can be valid to construct ex ante welfare measures from the HP coefficients without departing from Rosen’s (1974) framework.

As price-taking households are assumed to freely choose a house with any combination of physical attributes in a neighborhood that provides their desired bundle of spatially-delineated amenities (Kuminoff and Pope, 2014), their utility \( V \) maximization problem is:

\[
\max_{\xi, x, \eta, \zeta} V(\xi, x, \eta; \zeta), \text{ subject to } \iota = \eta + y(\xi, x; b) \tag{5}
\]

where we deliberately distinguish between \( x \) that is a vector of individual property characteristics and \( \zeta \) that is a vector of spatially-delineated amenities. Both \( x \) and \( \zeta \) may include omitted variables that are captured in the error term of the econometric model. Given its preferences \( \zeta \), income \( \iota \), and the after-tax house price \( y \), a household selects the quantity of the composite good \( \eta \) and the housing attributes/amenities to maximize utility. The price of housing, \( y(\xi, x; b) \), is expressed as a general parametric function of \( x, \zeta, \) and a parameter vector, \( b \). The first-order conditions are:

\[
\frac{\partial y(\xi, x; b)}{\partial x_k} = \frac{\partial V}{\partial x_k} / \frac{\partial V}{\partial x_k} \tag{6a}
\]
\[
\frac{\partial y(\xi, x; b)}{\partial \xi_\lambda} = \frac{\partial V/\xi_\lambda}{\partial V/\eta}
\]  

(6b)

Where \(x_k\) is the \(k\)th \((k = 1, \ldots, K)\) structural characteristic and \(\xi_\lambda\) is the \(\lambda\)th \((\lambda = 1, \ldots, \Lambda)\) spatially-delineated amenity. Equations 6a and 6b represent each household choosing the house and neighborhood that respectively provide quantities of \(x_k\) and \(\xi_\lambda\) at which their willingness to pay (WTP) for an additional unit equals their marginal implicit price.

The producer’s cost function is given by \(Z(\xi, x, \psi; j)\), where \(\psi\) is the number of houses sold and \(j\) is a parameter vector denoting the idiosyncratic costs. To follow Rosen (1974) and to simplify the exposition, each producer is assumed to be a price taker who is free to vary the number of houses sold, and specializes in producing exactly one type of house. The profit maximization problem of this producer is given in equation 7.

\[
\max_{x, \psi} \pi = \psi \cdot y(\xi, x; b) - Z(\xi, x, \psi; j)
\]  

(7)

Equation 8 specifies the first order conditions.

\[
y(\xi, x; b) = \frac{\partial Z(\xi, x, \psi; j)}{\partial \psi}, \quad \frac{\partial y(\xi, x; b)}{\partial x_k} = (1/\psi) \frac{\partial Z(\xi, x, \psi; j)}{\partial x_k}
\]  

(8)

The number of houses \(\psi\) is set by the producers so as the offer price of the marginal house equals its production cost, and \(x_k\) is set to the point that the per unit marginal cost of the housing attribute equals its implicit price. Equilibrium is achieved when the first-order conditions in equations 6 and 8 are simultaneously satisfied for all consumers and producers, implicitly defining the equilibrium hedonic price function that clears the market.

The pricing of a house is a complicated matter and a potential cognitive or computational burden to individual households or small producers. It can be challenging for individual households to gauge the market clearing price for a whole bundle of housing characteristics, or indeed the implied market price for each individual attribute. Such challenges can compromise the household’s decision-making process of evaluating their own reservation price for a house with regard to the market clearing price. Equivalent burdens are applicable to a producer with no market power, who has to adjust the number of houses \(\psi\) and the amount of housing attribute \(x_k\) according to the (potential) market clearing prices. These issues are assumed away by the perfect information, rationality, and
the other conditions in Rosen’s (1974) framework. For the purpose of this paper, we do not address information asymmetries, strategic behavior, and other market mechanism issues (Nanda and Ross, 2012; Thanos and White, 2014). Nevertheless, the whole pricing process does not take place in an informational vacuum, rather, there is a widely used source of information for buyers and producers that can provide a precise picture of the local housing market and of the market clearing prices: comparable sales.

Real estate common practice has long employed, implicitly and explicitly, a nearest neighbor (NN) approach of comparable sales (Isakson 1986) that can only have a spatiotemporal basis. Space is quite an important consideration, as comparables in close spatial proximity provide a precise reflection of how a specific combination of spatially-delineated amenities is valued in the market (Kuminoff and Pope, 2014). Time is also key, as the information from the comparables has to be up-to-date, reflecting current market conditions. The myopic consumer assumption in Rosen (1974) is relevant here and potentially not very far from actual behavior in the housing market, as households have limited capacity for long term foresting. Myopia can also concern the past in the form of limited memory in the market (Dubé et al., 2018).

Following Kuminoff and Pope (2014) and without departing from Rosen (1974), we need to rewrite the price function to acknowledge its dependence on model primitives and to scrutinize on the role of comparables in the price formation process. The primitives of the model include: the distribution of consumer types, \( A(i, \varsigma) \sim \delta \), the distribution of producer types, \( B(j) \sim \theta \), and the spatial distribution of spatially-delineated amenities, \( \Gamma(\gamma) \sim o \). Information on comparable sales is provided by a vector of comparable prices with the corresponding matrices of individual property characteristics and spatially-delineated amenities \( y_c(\xi_c, x_c; b) \sim u_c \). The Real Estate literature and practice tends to explicitly parametrize the level of comparability with a number of complicated metrics (McCluskey and Borst, 2017), rather than assume perfect comparability, as is the case with the Grid CSA equations 3 and 4. For the purposes of this exposition we parametrize the comparability effect in \( \rho \). When \( \rho = 1 \), it denotes perfect comparability, meaning that the set of comparable sales is in very close spatial and temporal proximity to perfectly reflect
local market conditions. The comparables are also broadly similar regarding structural housing attributes\(^6\), so as to inform market participants of the specific market clearing house prices. Housing attribute distance that is discussed in sections 2.3 and 2.4 can address some limited divergence. Otherwise, \(\rho\) may capture the discount of information that is not completely proximal in space, time, and attributes to provide a perfect reflection of local market conditions and prices. The different approaches to comparability measurement in the real estate literature and their comparison to our approach are discussed in section 2.4.

Equation 9 provides the price function acknowledging primitives and comparables.

\[
y(\xi, x; b) \equiv y[\xi, x; b[\delta(\rho v_c), \theta(\rho v_c), o(\rho v_c)]]
\]

where \(\delta\), \(\theta\), and \(o\) are vectors of parameter, \(v_c\) is a vector containing information on the comparable sales, and \(\rho\) is a parameter capturing the “comparability” effect. The reduced form parameters, \(b\), describing the shape of the price function are endogenously determined by the structural parameters (\(\delta\), \(\theta\), and \(o\)) (Kuminoff and Pope, 2014). The structural parameters are in turn affected by the exogenous comparable sales’ information in \(v_c\) and the level of comparability in \(\rho\). Given the Rosen (1974) conditions that allow us to assume a common parameter vector \(b\) between comparable sales and the exogeneity of \(v_c\), we can include \(\rho\) in the price function and estimate it based on comparable sales information as shown in equation 10.

\[
y[\xi, x; b[\delta(\rho v_c), \theta(\rho v_c), o(\rho v_c)]] \Rightarrow y(v_c, \xi, x; b, \rho)
\]

This equation is consistent with CSA and produces two types of parameters: \(b\) capturing the marginal WTP housing attributes, and which can also be interpreted as the impact of each housing on the price; and \(\rho\) capturing the information comparability of the nearest neighbors. Section 2.3 discusses the conditions under which the combination of \(b\) and \(\rho\) comes up in an empirical setting and how it can be interpreted. It is stressed that there is no basis in economic theory for weighing the impact of housing characteristics to the price other than the unbiased HP parameters in \(b\). In equations 1 to 4, \(\phi\) is assumed to be a

\(^6\) This implicitly raises the issue of housing submarkets that is not addressed here, as a small one bedroom flat sale may not serve well as a comparable for a mansion, even when it is in close spatial and temporal proximity.
subjective estimate, based on expert local knowledge, of the impact of a housing characteristic to the price. A real estate professional with a perfect knowledge of the market or an excellent statistical model would be expected to recover a $\varphi$ that in theory converges to the equivalent unbiased HP parameter. As the HP theoretical framework assumes the coefficients to be stable and relevant across observations in a specific local market, we use the HP coefficients, instead of $\varphi$, for any comparables introduced to the HP equation.

2.3. Introducing CSA to HP Estimation

Operationalizing the estimation, the typical 1st stage HP equation consistent with Rosen(1974) and Palmquist (1992) is given in equation 11.

$$y_h = \alpha + \sum_{k=1}^{K} \beta_k x_{hk} + \sum_{\lambda=1}^{A} \gamma_\lambda \xi_{h\lambda} + \epsilon_h \quad (11)$$

Where $\beta_k$ is the estimate of the impact of the $k$ observable structural characteristics $x_k$, $\gamma_\lambda$ is the estimate of the impact of the $\lambda$ spatially-delineated amenity, $\xi_{h\lambda}$, on the sale price of house $h$, and $\epsilon_h$ is an independent and identically distributed (IID) error term ($\epsilon_h \sim N(0, \sigma^2)$). By definition, $\xi$ are observed and unobserved spatially-delineated amenities that are common for very close neighbors.

We introduce comparable sales into HP in equation 11 by taking the pairwise difference. This is consistent with CSA in equations 3-4, which Colwell et al (1983) hinted at, and with the theoretical framework culminating in equation 10. We introduce the equivalent of $\rho \mathbf{u}_1$ that includes information on the house price and all attributes of comparable sale 1: $\rho(y_1 - \alpha - \sum_{k=1}^{K} \beta_k x_{1k} - \sum_{\lambda=1}^{A} \gamma_\lambda \xi_{1\lambda})$. This is repeated for $C$ comparable sales ($c = 1, \ldots, C$) to produce the system of the pairwise difference equations shown in equation 12a.

$$y_h = a - \rho a + \rho y_1 + \sum_{k=1}^{K} \beta_k (x_{hk} - \rho x_{1k}) + \sum_{\lambda=1}^{A} \gamma_\lambda (\xi_{h\lambda} - \rho \xi_{1\lambda}) + \epsilon_h$$

$$y_h = a - \rho a + \rho y_2 + \sum_{k=1}^{K} \beta_k (x_{hk} - \rho x_{2k}) + \sum_{\lambda=1}^{A} \gamma_\lambda (\xi_{h\lambda} - \rho \xi_{2\lambda}) + \epsilon_h \quad (12a)$$

$$y_h = a - \rho a + \rho y_3 + \sum_{k=1}^{K} \beta_k (x_{hk} - \rho x_{3k}) + \sum_{\lambda=1}^{A} \gamma_\lambda (\xi_{h\lambda} - \rho \xi_{3\lambda}) + \epsilon_h$$

$$
\vdots
$$

$$y_h = a - \rho a + \rho y_C + \sum_{k=1}^{K} \beta_k (x_{hk} - \rho x_{Ck}) + \sum_{\lambda=1}^{A} \gamma_\lambda (\xi_{h\lambda} - \rho \xi_{C\lambda}) + \epsilon_h$$
\[
y_h = (1 - \rho)\alpha + \rho C^{-1} \sum_{c=1}^{C} y_c + \sum_{k=1}^{K} \beta_k (x_{hk} - \rho C^{-1} \sum_{c=1}^{C} x_{ck}) + \sum_{\lambda=1}^{A} y_{\lambda}(\xi_{h\lambda} - \rho C^{-1} \sum_{c=1}^{C} \xi_{c\lambda}) + \epsilon_h
\]

(12b)

Adding vertically the comparables in Equation 12a produces Equation 12b, showing the price of a house \(h\) to be given by the average of \(C\) comparable prices adjusted by the weighted pairwise difference of their attributes to the house \(h\) attributes. Equation 12b cannot be estimated in its current form and needs to be transformed into a spatiotemporal error model (STEM) equivalent to that of Pace and Gilley (1998) without the arrow of time violation, as shown in equation 13.

\[
y_h = \alpha + \sum_{k=1}^{K} \beta_k x_{hk} + \sum_{\lambda=1}^{A} y_{\lambda} \xi_{h\lambda} + \rho C^{-1} \sum_{c=1}^{C} (y_c - \alpha - \sum_{k=1}^{K} \beta_k x_{ck} - \sum_{\lambda=1}^{A} y_{\lambda} \xi_{c\lambda}) + \epsilon_h \equiv y_h = \alpha + \sum_{k=1}^{K} \beta_k x_{hk} + \sum_{\lambda=1}^{A} y_{\lambda} \xi_{h\lambda} + \rho C^{-1} \sum_{c=1}^{C} (y_c) + \epsilon_h
\]

(13)

The spatially-delineated amenities do not drop out, as \(\sum_{\lambda=1}^{A} y_{\lambda} \xi_{h\lambda}\) are included in the spatial error term in the brackets; thus, this specification does not provide the OVB reduction benefits shown below. As \(\rho v_c\) is essentially the spatial (and temporal) error term \(\rho C^{-1} \sum_{c=1}^{C} (\epsilon_c)\), an additional error term for estimating the coefficients of comparable \(c\) is not justified, as its price and its characteristics are known\(^7\) and the HP coefficients are common across the market.

When comparables are defined as having common spatially-delineated amenities \((\xi_{h\lambda} = \xi_{c\lambda})\), the term \(\sum_{\lambda=1}^{A} y_{\lambda} (\xi_{h\lambda} - \rho C^{-1} \sum_{c=1}^{C} \xi_{c\lambda})\) in equation 12b is simplified to \(\sum_{\lambda=1}^{A} y_{\lambda} (1 - \rho) \xi_{h\lambda}\), and we get equation 14:

\[
y_h = (1 - \rho)\alpha + \rho C^{-1} \sum_{c=1}^{C} y_c + \sum_{k=1}^{K} \beta_k (x_{hk} - \rho C^{-1} \sum_{c=1}^{C} x_{ck}) + \sum_{\lambda=1}^{A} y_{\lambda} (1 - \rho) \xi_{h\lambda} + \epsilon_h
\]

(14)

When \(\rho\) is 1, then this process essentially describes the basic grid CSA. The \(K\) differences between each variable \(x\) and the average of \(C\) comparables is shown in equation 15.

\[
y_h = C^{-1} \sum_{c=1}^{C} y_c + \sum_{k=1}^{K} \beta_k (x_{hk} - C^{-1} \sum_{c=1}^{C} x_{ck}) + \epsilon_h
\]

(15)

Equation 15 can only be estimated in this form when \(\rho\) is 1, in which case the \(\beta\)'s are not affected by spatially-delineated OVB, as the (observed and unobserved) \(\xi\) terms completely

\(^7\) The sale price \(y_c\) of comparable \(c\) and its characteristics are observed. These are not estimates \((\hat{y}_c)\) to require an extra error term, as \(y_c = \hat{y}_c + \hat{\epsilon}_c\). The temporal ordering and arrow of time constraints to spatial matrices discussed in section 2.3 ensure that there are no such endogeneities.
drop out. It is stressed that unobserved $x$ terms can potentially still cause OVB, hence this specification reduces but may not completely eliminate OVB.

If the comparable effect is different to one ($\rho \neq 1$), equation 12b needs to be transformed by breaking up the term in the brackets into the Durbin model structure shown in equation 16.

$$y_h = \omega + \rho C^{-1} \sum_{c=1}^{C} y_c + \sum_{k=1}^{K} \beta_k x_{hk} - \sum_{k=1}^{K} \mu_k C^{-1} \sum_{c=1}^{C} x_{ck} + \sum_{\lambda=1}^{A} \gamma_{\lambda} \xi_{h\lambda} - \sum_{\lambda=1}^{A} \nu_{\lambda} C^{-1} \sum_{c=1}^{C} \xi_{c\lambda} + \epsilon_h$$  \hfill (16)

We substitute $\rho \beta_k$ and $\rho \gamma_{\lambda}$ with $\mu_k$ and $\nu_{\lambda}$ respectively, allowing $\mu_k$ and $\nu_{\lambda}$ to vary independently of $\rho$, $\beta_k$, and $\gamma_{\lambda}$. The “spatial constant” term $(1 - \rho)\alpha$ is denoted as $\omega$. If $\rho$, $\beta_k$, and $\gamma_{\lambda}$ do not vary independently, STDM becomes the STEM shown in equation 13, which is something to be tested empirically.

It is reiterated that comparables must be in very close spatial proximity to house $h$ in order to have common spatially-delineated amenities ($\xi_{h\lambda} = \xi_{C\lambda}$). If that is the case, then we can follow the simplified equation 14 and get the STDM in equation 17.

$$y_h = \omega + \rho C^{-1} \sum_{c=1}^{C} y_c + \sum_{k=1}^{K} \beta_k x_{hk} - \sum_{k=1}^{K} \mu_k C^{-1} \sum_{c=1}^{C} x_{ck} + \sum_{\lambda=1}^{A} \gamma_{\lambda} (1 - \rho) \xi_{h\lambda} + \epsilon_h$$  \hfill (17)

Where $\gamma_{\lambda} = (1 - \rho)$ that is not the HP coefficient anymore but the coefficient minus the spatial effect. If the spatial effect dominates ($\rho \gg 1 \vee \rho \ll 1$), it can have a significant impact on the magnitude and sign of $\gamma_{\lambda}$. Nevertheless, with a bit of algebraic manipulation we can estimate the HP coefficient $\gamma_{\lambda} = \gamma_{\lambda} (1 - \rho)^{-1}$. To limit spatially-delineated OVB, $\rho$ needs to be close to one ($\rho \sim 1$), so as the spatial effects ($rs$) of the unobserved $\xi$ terms disappear completely and stop affecting the $\beta$s.

It is stressed that the STDM $\mu$s (and vs, if equation 16 is specified instead of 17) should be interpreted strictly in the context of attribute difference between the comparables and the house $h$. This is an adjustment process that makes the price of comparable house $c$ comparable to house $h$. We contend that there is scant theoretical justification in HP for spatially weighted independent variables that are not part of this comparable adjustment process.
As the number of comparables increases, so does the probability for spatially-delineated amenities to diverge between comparables:

\[
E(\sum_{c=1}^{C} \sum_{\lambda=1}^{\Lambda} (\xi_{h\lambda} - \xi_{C\lambda})) \sim 0, \quad (c \to 1, \xi_{Low} \leq c \leq \xi_{High})
\]  

(18)

Where \( \xi \) is the range of comparables within which the common spatially-delineated amenities \( (\xi_{h\lambda} = \xi_{C\lambda}) \) assumption holds. The whole process needs at least one comparable to work, otherwise the equation 12b is reduced to the typical HP. The STDM specifications in equations 16 and 17 provide an opportunity to obtain the \( \xi \). If \( \xi_{h\lambda} \neq \xi_{C\lambda} \), then we need to estimate \( \nu_\lambda \) in equation 17 that drops out \( (\nu_\lambda=0) \) when \( \xi_{h\lambda} = \xi_{C\lambda} \). Therefore, we can test for the range of NNs, in which the extra constraint \( \nu_\lambda \) in equation 17 is superfluous when compared to equation 16. Essentially, we are testing \( r_\lambda = \gamma_\lambda + \nu_\lambda \) across equations.

The maximum number of comparables \( (\xi_{High}) \), which can be introduced before the common spatially-delineated amenities assumption does not hold anymore, should be market specific and can potentially depend on market “heat” (Thanos and White 2014) and built environment characteristics, such as dwelling density (Dunse et al., 2013). This issue is exemplified in the common real estate practice of selecting as a low number of comparables in grid CSA as 5 (Ratcliffe 1972), which may not be the case for AVMs, where one can introduce hundreds or thousands of comparables. Such high number of comparables of course would invalidate the common spatially-delineated amenities assumption. Ratcliff (1972) also argues that using below 5 comparables tends to be unstable and thus is avoided by real estate professionals. Our methodological framework distinguishes and systematizes the different elements in the CSA process and, by so doing, allows us to test the effects of different numbers of comparables.

Making explicit the assumptions required for STAR application in HP using housing data, STAR can be derived from equation 17 with an added requirement for the \( x \)s process to be superfluous or detrimental to the statistical modelling. It is relatively easy to test superfluity \( (\mu_k = 0) \), which we do, but unlikely to hold, given the amount of variation typically captured in this term \( (\sum_{k=1}^{K} \mu_k C^{-1} \sum_{c=1}^{C} x_{ck}) \). Examining the statistical detriment is a bit more complicated to analyze. In the case of a very high number of compatibles \( (c \to \)
the comparables’ attribute \( x_k \) will asymptotically equal its mean, as shown in equation 19.

\[
\sum_{k=1}^{K} \mu_k C^{-1} \sum_{c=1}^{C} x_{ck} \sim \sum_{k=1}^{K} \mu_k \bar{x}_k , \quad (c \to n - 1)
\] (19)

The mean (\( \bar{x}_k \)) in equation 19 does not really reflect the comparable adjustment process that is integral to explaining the inclusion of the comparables’ attributes in the HP model. Simultaneously accounting for the effects of independent variables, of independent variables’ means, and of the comparable price mean raises an issue of overestimation and (Smith, 2009).

For STAR to be applied robustly by satisfying both equations 18 and 19, the available data need to include comparables that: (a) are in close geographical proximity to have common spatially-delineated amenities; (b) are numerous enough to cause overestimation or multicollinearity when spatiotemporal “lagged” \( x \)s are included in the estimation. Only under these conditions the STAR model in equation 20 can be robustly employed instead of STDM.

\[
y_h = \omega + \rho C^{-1} \sum_{c=1}^{C} y_c + \sum_{k=1}^{K} \beta_k x_{hk} + \sum_{\lambda=1}^{A} r_\lambda \xi_{h\lambda} + \varepsilon_h
\] (20)

Condition (b) essentially introduces a constraint to STDM estimation that becomes problematic in high number of comparables. When condition (a) is not satisfied one needs to estimate both \( \nu_\lambda \) and \( \gamma_\lambda \) in equation 16, essentially turning STAR into a partial Durbin model form. It is stressed that single coefficients in STAR for spatially-delineated amenities are biased if condition (a) is not satisfied. Even if condition (a) holds, when the spatial effect dominates (\( \rho \gg 1 \lor \rho \ll 1 \)), it can have a significant effect on the magnitude and sign of \( r_\lambda \) and cause significant divergence from the HP coefficient \( \gamma_\lambda \). STAR is not appropriate for capturing spatially-delineated amenities \( \gamma_\lambda \). At best, \( r_\lambda \) may be close to the true HP coefficient; at worst a completely biased \( r_\lambda \) is estimated, being dominated by \( \nu_\lambda \) that has a negative sign, as shown in equation 16. These issues are tested in the empirical model section.
2.4. Introducing weight matrices to CSA, STDM, STAR, and STEM

Let matrix $C$ map the NN or comparables for each observation. Let spatiotemporal distances from each $h$ to all houses $j \neq h$ be ranked from closest to furthest in space and time. This naturally includes only relations for past to future and results in a lower triangular matrix. For each $c = 1, \ldots, n - 1$ the set $G_c(h) = \{j(1), \ldots, j(c)\}$ contains the $C$ closest comparables. The nearest neighbor weight matrix $C$, the elements of which are shown in Equation 21, is row-standardized to obtain the averaged comparables in equations 12 - 17, and 20.

$$c_{hj} = \begin{cases} 1 & \forall j \in G_c(h) \\ 0 & \forall j \notin G_c(h) \end{cases} \quad (21)$$

Thanos et al. (2016) and Dubé et al. (2018) suggested that information from sales is discounted accordingly to spatio-temporal distance. Therefore, instead of $C$, a spatiotemporal matrix $W$ also specified in equation 22 is used to account for the spatiotemporal distance of each NN.

$$w_{hj} = \begin{cases} f(s_{hj}) \times m(\tau_{hj}) & \forall j \in G_c(h) \\ 0 & \forall j \notin G_c(h) \end{cases} \quad (22)$$

Where $s_{hj}$ is a function of spatial distance multiplied by a function of temporal distance $\tau_{hj}$. In matrix notation this is a Hadamard product of spatial and temporal distance matrices (Thanos et al., 2016).

Differencing within our framework takes place only in a very short time span between comparables, as opposed to differencing across long periods in first difference or difference-in-difference models. Sales with the maximum available proximity in time and space are typically sought as information sources for the equilibrium between housing supply and demand that can change across time and space (Thanos et al., 2016). Spatial proximity typically denotes that houses are in the same local market. However, this may not be always true, as the local market can be split by “barriers” and be defined in terms of house similarity regarding type, size, and other key attributes that signify catering to similar types of consumers. Most of real estate comparable valuation literature, as reviewed by
McCluskey and Borst (2017), employs a range of metrics, such as the Minkowski, Manhattan, and Mahalanobis distance, to measure the degree of attribute similarity between different houses. These house attribute distance metrics can potentially be useful empirically in defining the most appropriate comparables.

Krause and Kummerow (2009) propose adding a geographic distance penalty to a house attribute distance metric (Mahalanobis distance). We contend that the opposite is dictated by theory: adding “comparability” constraints (e.g. houses of the same type, similar in size or a selection of other key attributes) or potentially a Mahalanobis distance penalty to spatiotemporal NN. Besner (2002) did attempt to combine spatial distance and house attribute distance in the same weight matrix. The constraint of this approach, beyond Besner (2002) not accounting for temporal distance explicitly, is the difficulty of combining in a meaningful way distances measured in different units. Temporal distance is conceptualized as a time decay function when combined to spatial distance (Thanos et al., 2012; Dubé and Legros, 2013b; Thanos et al., 2016). Combining spatial, temporal, and attribute distance may not be feasible or at least it requires the development of a new methodological framework and extensive empirical testing. Nevertheless, adding “similarity” constraints to spatiotemporal NN does not create the same problems (Yousfi et al., 2019); a discussion which is outside the scope of this paper.

Having shown the underlying equations in section 2.3 and specified the comparables by employing row-standardized matrices \( C \) and \( W \), we can derive the spatiotemporal models. Matrix notation simplifies the exposition, hence the HP model in equation 11 is expressed as equation 23. The HP model in equation 23 is expected to suffer from OVB of unobservable spatially-delineated amenities.

\[
y = \mathbf{1}\alpha + \mathbf{X}\beta + \Xi\gamma + \epsilon
\]  
\[\text{(23)}\]

Where the sale price vector \( y \) is a function of observable structural attributes in matrix \( X \) multiplied by a vector of parameters \( \beta \), of observable spatial attributes in matrix \( \Xi \) multiplied by a vector of parameters \( \gamma \), the identity vector \( \mathbf{1} \) multiplying a constant \( \alpha \), and the error term vector \( \epsilon \).
To derive the econometric CSA model, we replace the comparables’ averages in equation 15 with the mathematically equivalent matrix $C$ and arrive at equation 24.

\[ y = Cy + (X - CX)\beta + \epsilon \]  
\[ y = Wy + (X - WX)\beta + \epsilon \]

(24a)  
(24b)

The key advantage of equation 24 is that these spatiotemporal models are not expected to suffer from the OVB of equations 12b-13, as the observable and unobservable spatially-delineated amenities $\zeta$ drop out and are no longer relevant to the estimation. As each price is only compared with very similar houses and prices, the following typical HP problems are foregone or significantly ameliorated: a) potential OVB owing to unobserved spatially-delineated amenities; b) certain nonlineairties due to large distance between observations over space and time.

To derive the STDM, we replace in equation 17 the comparables’ averages with the mathematically equivalent matrix $C$ (or $W$). The parametrization of comparables’ effect in $\rho$ and $\mu$ (vector of coefficients) is especially important in flexible specifications ($W$) when the spatiotemporal distance between observations is explicitly included.

\[ y = \omega + \rho Cy + X\beta - CX\mu + \Sigma r + \epsilon \]  
\[ y = \omega + \rho Wy + X\beta - WX\mu + \Sigma r + \epsilon \]

(25a)  
(25b)

Where $r$ is a vector of parameters.

House attribute distance metrics, or comparability weights, can inform the comparable selection process and thus $\rho$, but not the weight/importance of each housing characteristic described by $\beta$ or $\gamma$. This distinction is clear in equation 25, where $\mu$ and $r$ indicate the exact cases where the coefficients combining the “comparability” and housing characteristic weights are estimated.

We also estimate the spatiotemporal adaptation (STEM) of Pace’s and Gilley’s (1998) SEM in equation 26 by introducing $C$ or $W$ in equation 13, where the error term $(y_c - \alpha - \sum_{k=1}^{K} \beta_k x_{ck} - \sum_{\lambda=1}^{A} \lambda y_{\lambda c})$ is denoted by the error vector.

\[ y = \omega + X\beta + \Sigma \gamma + \rho \epsilon + \epsilon \]  
\[ y = \omega + X\beta + \Sigma \gamma + \rho W \epsilon + \epsilon \]

(26a)  
(26b)
As mentioned already, Pace et al. (1998) not only rectified the SEM “time-travel”, but their “spatiotemporal models essentially generalize conventional hedonic regressions and the adjustment grid method (Pace and Gilley 1998)” (Pace et al., 1998, p: 16). However, beyond not explicitly deriving STEM or STAR, there is room for further developments in their approach that our paper builds upon. The process of transforming the SEM into a spatial (and temporal) Durbin model in Pace et al. (1998, 2000) ignores the issues of common spatially-delineated amenities and their final model provides no advantages in reducing OVB as illustrated in section 2.3. Pace et al. (1998, 2000) also use a matrix multiplication of separate spatial, $S$, and temporal, $T$, matrices to obtain the spatiotemporal matrix $W = \rho_T T + \rho_S S + \rho_{ST} ST + \rho_{TS} TS$. $ST$ and $TS$ are different from each other and are explained as compound spatio-temporal effects. Following Smith and Wu (2009) and Thanos et al. (2016), we argue that the Hadamard product in equation 22 provides a clearer specification of the spatiotemporal NN or comparables and their effect on the house price. A key advantage of the Hadamard product is that it specifies a spatiotemporally unique influence from each comparable and avoids having to parametrize comparable influences separate in time ($T$), space ($S$), “space-time” ($ST$), and “time-space” ($TS$). The Hadamard product specification also forgoes the problem noted by Smith and Wu (2009) of interpreting the economic significance of $\rho_{ST}$ and $\rho_{TS}$.

The general form of a STAR model is given in equation 27, as derived from equation 20.

\[
\begin{align*}
y &= t\omega + \rho Cy + \Xi r + X\beta + \varepsilon \\
y &= t\omega + \rho Wy + \Xi r + X\beta + \varepsilon
\end{align*}
\] (27a) (27b)

As mentioned in the previous section, if comparables are in close geographical proximity to have common spatially-delineated amenities, $\Xi W v + \Xi y$ become $\Xi r$.

3. Data and Empirical Strategy

To estimate the spatio-temporal extensions of the HP model and test their performance, single family home transaction data from Lucas County, Ohio, 1993-1998, are employed.
The database is available in the MatLab software library (LeSage 1999) and has been used by Lesage and Pace (2004) and Dubé and Legros (2013b). It consists of 25,357 transactions and contains information on the nominal sale price in US dollars. An advantage of publicly shared data lies on the easily verifiable estimation procedure by any researcher.

An out-of-sample comparison for testing the performance of the different models and specifications requires splitting the data. Data between 1993 and 1997 are used for estimation, while transactions in 1998 are used to make predictions. A total of 20,979 transactions are used for the in-sample estimation, and 4,378 transactions are used for the out-of-sample prediction exercise. Instead of using all the past in-sample observation for prediction, we also employ one-step-ahead forecasting, in which for an observation \(i\) in time \(t+1\) we select comparables in time \(t\) provided by spatiotemporal nearest neighbors and predict sale price \((\hat{y}_{i,t+1})\). We do so by using coefficients estimated on data from previous periods, as dictated by the temporal distance constraint mentioned above. This process is repeated for all observations within one time period \((t+1)\) and then we proceed to all observations in subsequent time periods \((t+2...t+p)\). This process is closer to comparable sales approach than using the coefficient from the whole in-sample for prediction.

Table 1 provides the descriptive statistics for the in-sample used for estimation. Given that OVB plagues HP no matter the amount of housing characteristics included, we simplify the exposition by employing seven key observed structural variables as well as one spatial variable for robustness checks. Both age and its logarithmic transformation are included as independent variables to account for non-linear effect, reflecting a convex relationship that is common in the literature and fits this data best\(^8\). The temporal distribution of the transactions evolves over time, with more transactions for each year, as shown in Figure 1. It is noted that the time fixed effects drop out in STDM and econometric CSA.

\(^8\) Evidence is provided upon request
Nothing in the theory suggest that the OLS estimator in the absence of OVB is not BLUE, as spatiotemporal effects in lower triangular systems are exogenous (Pace et al., 1998). Given the double-log specification, all such coefficients can be directly interpreted as elasticities, as $C$ and $W$ are lower-triangular and the typical “spatial spillovers” of more aggregate models are not applicable here (Dubé and Legros, 2018). Matrix $W$ is specified as a Hadamard product of spatial and temporal distance functions, as shown in equation 22. For simplicity, we use linear distance functions and constrain the temporal distance to be between 10 to 180 days before the sale. This is consistent with Dubé et al. (2018), who show that price information beyond 6-8 months in the past does not improve model estimation.

Table 1: Descriptive Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Notation</th>
<th>Description</th>
<th>Mean</th>
<th>Std.Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>sale price</td>
<td>$y$</td>
<td>Sale price</td>
<td>79018</td>
<td>59655</td>
<td>2000</td>
<td>875000</td>
</tr>
<tr>
<td>ln_price</td>
<td>$y$</td>
<td>Ln of Sale price</td>
<td>11.020</td>
<td>0.763</td>
<td>7.601</td>
<td>13.682</td>
</tr>
<tr>
<td>age</td>
<td>$x_1$</td>
<td>Age of the house</td>
<td>50.367</td>
<td>27.931</td>
<td>0.000</td>
<td>161.000</td>
</tr>
<tr>
<td>ln_age</td>
<td>$x_2$</td>
<td>Ln of age</td>
<td>3.672</td>
<td>0.940</td>
<td>0.000</td>
<td>5.088</td>
</tr>
<tr>
<td>ln_lotsize</td>
<td>$x_3$</td>
<td>Ln of lotsize in sqft</td>
<td>8.976</td>
<td>0.790</td>
<td>6.554</td>
<td>12.969</td>
</tr>
<tr>
<td>ln_area</td>
<td>$x_4$</td>
<td>Ln of living area in sqft</td>
<td>7.216</td>
<td>0.368</td>
<td>4.787</td>
<td>8.938</td>
</tr>
<tr>
<td>ln_bathroom</td>
<td>$x_5$</td>
<td>Ln of bathroom number</td>
<td>0.788</td>
<td>0.185</td>
<td>0.000</td>
<td>2.079</td>
</tr>
<tr>
<td>Stories2+</td>
<td>$x_6$</td>
<td>Over 2 stories dummy</td>
<td>0.317</td>
<td>0.465</td>
<td>0.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Garage</td>
<td>$x_7$</td>
<td>Garage dummy variable</td>
<td>0.862</td>
<td>0.344</td>
<td>0.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Dist_Center</td>
<td>$\xi_1$</td>
<td>Distance to Toledo city center</td>
<td>8802</td>
<td>5102</td>
<td>384</td>
<td>38581</td>
</tr>
</tbody>
</table>

N = 20,979

---

9 Nevertheless, to pre-empt any potential issues (see footnote 1), we also attempted to employ the IV/GMM estimator of Kelejian and Prucha (1998), as maximum likelihood estimation is not feasible for models with lower triangular matrices that are not of full rank. However, IV/GMM produced unfeasibly high values for $\rho$, which is a known issue as Elhorst (2014, pp: 17) puts it “one disadvantage of the IV/GMM estimator is the possibility of ending up with a coefficient estimate outside its parameter space”. This is possibly exacerbated by the lower triangular nature of the matrices. Nonetheless, the OLS and IV/GMM coefficient magnitudes, except $\rho$, exhibit close consistency and only insignificant differences that reinforce our belief that the OLS estimator can be unbiased and efficient for these spatiotemporal models.
Selecting a limited number of comparables is the common practice in real estate, for instance Ratcliffe (1972) argues that less than 5 comparables is unsafe and more than 10 comparables is probably unnecessary. LeSage and Pace (2014) show that for purely spatial matrices, marginal increases to the number of NN do not provide any significant improvements, as correlation is very high. Our methodological framework suggests that certain specifications may be quite sensitive to increasing NN. Hence, we follow an iterative process to cover 1 - 1000 nearest spatio-temporal neighbors, which involves running 8 spatiotemporal specifications at 68 regular NN intervals\(^\text{10}\). Our iterative approach allows to test the following hypotheses that are suggested by our methodological framework:

1. The coefficient of garage is expected to pick up unaccounted spatially-delineated amenities as it could reflect unobserved transportation attributes causing OVB in HP. As a result, we expect garage HP coefficient to have significantly higher

---

\(^{10}\) 1 NN interval for 1 - 20NNs, 5 NN intervals for 25 - 250NNs, and two further estimations at 500 and 1000 NNs.
magnitude than that of the spatiotemporal models in general. We also expect econometric CSA to have the lowest garage coefficients for low NN levels, since unaccounted spatially-delineated amenities drop out. At higher NNs, we expect the spatiotemporal models to start converging to the HP coefficient.

2. The methodological framework suggests that the HP constant term should also be found in STEM. The “spatial constant” terms in STDM and STAR should be significantly different to the HP constant.

3. We test whether STDM should be reduced to STEM ($\mu + \rho \beta = 0$)

4. Through the iterative estimation we test if and when STDM, STAR, and STEM start exhibiting overestimation issues where $\rho > 1$.
   a. If at a certain NN level STDM $\rho > 1$ and STAR $\rho < 1$, the second applicability condition of STAR in equation 19 comes into effect.

5. The spatiotemporal coefficient $\rho$ should not be statistically different to one ($\rho = 1$) for the econometric CSA. We test whether this holds and for which range of NNs.

6. We perform a robustness check by introducing the spatially-delineated amenity of distance to the city center. If our framework is correct, we expect this variable not to be statistically significant ($\gamma_1 = 0$) for lower NNs in econometric CSA. Conversely, we expect this variable to be statistically significant in HP and STEM.

7. Through iterative estimation we can arrive to $\xi$ in equation 18 for econometric CSA, STDM, and STAR. $\xi$ is the NN range under which the common spatially-delineated amenities assumption holds. For CSA this means $\gamma_1 = 0$, which needs to hold for all unobserved $\gamma$s to be a sufficient condition for minimizing spatially-delineated OVB. For STAR and STDM this means $\nu_1 = 0$. An extra condition of $r_1 = 0$ holding for all unobserved $r$s is required for limiting spatially-delineated OVB.
   a. In conjunction with 4a, we test whether in our data both STAR applicability conditions come into effect, rendering it preferable to STDM: $\nu_1 = 0$ and $\rho < 1$ for STAR at the same NNs when STDM $\rho \geq 1$.
   b. We also test across the range of NN if STAR is $r_1$ dominated by $\nu_1$. $\nu_1$ is expected to have a negative sign as shown in equation 16.
4. Econometric Modeling Results and Predictions

This section first discusses the results of the econometric models applied on data with informational constraints as is the common practice for real estate professional, testing the hypotheses illustrated in the empirical strategy section. This is followed by a subsection that compares forecasting accuracy and volatility between statistical and non-statistical methods of price prediction. There is also a subsection that presents robustness checks by introducing a spatially-delineated amenity to further examine our hypotheses.

4.1. Econometric model results

It is stressed that the purpose of this paper is not to achieve the best fitting models, but to compare techniques under informational and computational constraints. Hence in the first stage, we intentionally employ only basic housing attributes ($x_s$) without any spatial information ($γ_s$), in order to emulate real estate practice with informational constraints. As our empirical strategy requires to run and test a multitude of models, we present the results of each spatiotemporal specification only at the NN level that achieves the best fit to the data (lowest AIC/BIC). Table 2 provides the results for each econometric model and the NN level of each specification. It is generally noted that spatiotemporal models achieve their best fit at low levels of NNs. This is underlines that abundant information on the attribute similarity between different houses, a common in many AVMs, may not be very helpful in explaining house prices without the inclusion of time and space proximity.

All $β$s in Table 2 are statistically significant; their signs are consistent with economic theory expectations. The coefficient magnitudes are relatively consistent across specifications, with a few key divergences that require further discussion. The relationship of house age to price is convex in all models, which is reasonable and consistent with the literature. The parsimonious HP model with only a few structural variables achieved a decent goodness of fit, explaining 72% of the variation in the data. The spatiotemporal models provide significant improvement in goodness of fit to over 80%, which implies, not unexpectedly, unexplained variation with spatiotemporal patterns in HP.
STDM W at 15 NNs provides the best fit across all models and levels of NNs, with CSA W at 15 NNs a close second best. CSA coefficients are quite consistent to the STDM ones, despite the former being a parsimonious no-constant model. We reject the null hypothesis (hypothesis 3 in section 3) that STDM should be reduced to STEM ($H_0: \mu + \rho \beta = 0$), employing a Wald test across the full range of NNs (NN15: $\chi^2[7]=561.70$, Prob $> \chi^2 =0.000$)\(^\text{11}\). STEM and HP have almost identical constants\(^\text{12}\) and quite different to STDM and STAR constants that reflect the spatial component. This is predicted by our methodological framework, detailed in hypothesis 2.

It is noted that only the STEM specifications in Table 2 exhibit potential overestimation reflected in $\rho > 1$. CSA $\rho$ is expected and seems to be equal to 1. This issue merits further investigation through our iterative estimation process, the results of which are shown in Figure 2. At low NNS, all models have $\rho < 1$, as illustrated in Figure 2, except CSA $\rho$ that stays near unity for the full range of NNs. The overestimation problem, $\rho > 1$, in STDM and especially STEM specifications is quite sensitive to the increase of NNs. Conversely, STAR specifications are less sensitive to overestimation issues, with the $\rho$ of STAR W converging to 1 and the STAR C $\rho$ being below 1 at 1000 NNs.

Table 3 provides precise statistical testing on the NN range for which $\rho$ is lower, equal, or higher to one for every specification, addressing hypothesis 4. STEM and STDM should only be applied on at very low NNs. The use of matrix $W$ instead of $C$ allows all specifications except STAR to be applicable for a greater range of NNs without overestimation issues. The $\rho$ in CSA specifications is equal to one for a large range of NNs, consistent with hypothesis 5, but outside this range these models are biased as illustrated in equations 14-16.

\(^{11}\) The full range of tests across the full range of NNs can be provided on request.
\(^{12}\) Asymptotic $t$-tests show that the difference between HP and STEM constants is not statistically different to 0 at $p < 0.001$. 

25
<table>
<thead>
<tr>
<th>Variables/Coefs</th>
<th>HP Eq23</th>
<th>CSA C NN8, Eq24</th>
<th>CSA W NN15, Eq24</th>
<th>STDM C NN6, Eq25</th>
<th>STDM W NN15, Eq25</th>
<th>STEM C NN15, Eq26</th>
<th>STEM W NN35, Eq26</th>
<th>STAR C NN6, Eq27</th>
<th>STAR W NN8, Eq27</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_1$: age</td>
<td>-0.0195 ***</td>
<td>-0.0108 ***</td>
<td>-0.0109 ***</td>
<td>-0.0109 ***</td>
<td>-0.0108 ***</td>
<td>-0.0151 ***</td>
<td>-0.0151 ***</td>
<td>-0.0114 ***</td>
<td>-0.0109 ***</td>
</tr>
<tr>
<td>$\beta_2$: ln_age</td>
<td>0.2952 ***</td>
<td>0.1738 ***</td>
<td>0.1817 ***</td>
<td>0.1761 ***</td>
<td>0.1806 ***</td>
<td>0.1892 ***</td>
<td>0.1841 ***</td>
<td>0.2216 ***</td>
<td>0.2158 ***</td>
</tr>
<tr>
<td>$\beta_3$: ln_lot_size</td>
<td>0.1815 ***</td>
<td>0.1047 ***</td>
<td>0.1078 ***</td>
<td>0.1037 ***</td>
<td>0.1057 ***</td>
<td>0.1942 ***</td>
<td>0.1932 ***</td>
<td>0.0490 ***</td>
<td>0.0447 ***</td>
</tr>
<tr>
<td>$\beta_4$: ln_area</td>
<td>0.6406 ***</td>
<td>0.6625 ***</td>
<td>0.6493 ***</td>
<td>0.6524 ***</td>
<td>0.6454 ***</td>
<td>0.6602 ***</td>
<td>0.6607 ***</td>
<td>0.5746 ***</td>
<td>0.5570 ***</td>
</tr>
<tr>
<td>$\beta_5$: Garage</td>
<td>0.3616 ***</td>
<td>0.2155 ***</td>
<td>0.2078 ***</td>
<td>0.2337 ***</td>
<td>0.2253 ***</td>
<td>0.2514 ***</td>
<td>0.2440 ***</td>
<td>0.2725 ***</td>
<td>0.2632 ***</td>
</tr>
<tr>
<td>$\beta_6$: ln_bathrm</td>
<td>0.2258 ***</td>
<td>0.1958 ***</td>
<td>0.1880 ***</td>
<td>0.1980 ***</td>
<td>0.1898 ***</td>
<td>0.2351 ***</td>
<td>0.2234 ***</td>
<td>0.1433 ***</td>
<td>0.1308 ***</td>
</tr>
<tr>
<td>$\beta_7$: Stories2+</td>
<td>0.1856 ***</td>
<td>0.0954 ***</td>
<td>0.0920 ***</td>
<td>0.1000 ***</td>
<td>0.0956 ***</td>
<td>0.1281 ***</td>
<td>0.1270 ***</td>
<td>0.1072 ***</td>
<td>0.1017 ***</td>
</tr>
<tr>
<td>$\mu_1$: age</td>
<td>0.0062 ***</td>
<td>0.0081 ***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_2$: ln_age</td>
<td>-0.0803 ***</td>
<td>-0.1212 ***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_3$: ln_lot_size</td>
<td>-0.0746 ***</td>
<td>-0.1058 ***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_4$: ln_area</td>
<td>-0.5726 ***</td>
<td>-0.6153 ***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_5$: Garage</td>
<td>-0.0199</td>
<td>-0.0156</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_6$: ln_bathrm</td>
<td>-0.1087 **</td>
<td>-0.1387 **</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_7$: Stories2+</td>
<td>-0.0144</td>
<td>-0.0384 **</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>constant</td>
<td>3.9778 ***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time fixed effects</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.9994 ***</td>
<td>0.9994 ***</td>
<td>0.8387 ***</td>
<td>0.9476 ***</td>
<td>1.1919 ***</td>
<td>1.2800 ***</td>
<td>0.5669 ***</td>
<td>0.5938 ***</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.718</td>
<td>N/A</td>
<td>N/A</td>
<td>0.825</td>
<td>0.832</td>
<td>0.803</td>
<td>0.810</td>
<td>0.807</td>
<td>0.814</td>
</tr>
<tr>
<td>AIC/BIC</td>
<td>20951/21086</td>
<td>11496/11560</td>
<td>10488/10551</td>
<td>11139/11266</td>
<td>10209/10336</td>
<td>13595/13738</td>
<td>12848/12991</td>
<td>13112/13255</td>
<td>12353/12495</td>
</tr>
</tbody>
</table>

***p < 0.001; **p < 0.01; *p < 0.05
Consistent with hypothesis 1 in the empirical strategy, the HP garage coefficient magnitude in Table 2 might be unreasonably high, as a garage seems to be increasing house prices by 36% on average. This drops significantly across the spatiotemporal specifications, which is seen as tentative evidence of spatially-delineated OVB in HP coefficients. This OVB is minimized in CSA C and CSA W have the lowest coefficient magnitudes. STDM garage coefficients are also very close to CSA implying that this specification may also benefit from minimized OVB. This merits further investigation in section 4.3.

Figure 2: The spatiotemporal coefficient $\rho$ for increasing NNs

![Graph showing the spatiotemporal coefficient $\rho$ for increasing NNs.](image)

Note: Each dot illustrates a different model at 5NN intervals from 5 - 250 NNs

The issue of spatially-delineated OVB in the garage coefficient is further investigated through iterative estimation, the results of which are displayed in Figure 3. All models employing a $C$ matrix converge to the HP coefficient at very high NNs, while models with the spatiotemporal matrix $W$ tend to keep their coefficients at relatively lower magnitudes. Most importantly, only STDM and CSA specifications have at very low NNs coefficients below 0.24, which increase very quickly with rising NNs. This is consistent with our
methodological framework that shows unobserved spatially-delineated amenities to be common and drop out only between few comparables in close proximity.

Table 3: The spatiotemporal coefficient $\rho$ for 1 - 1000 NNs at $p < 0.001$ significance level

<table>
<thead>
<tr>
<th>Model</th>
<th>NN range $\rightarrow \rho &lt; 1$</th>
<th>NN range $\rightarrow \rho = 1$</th>
<th>NN range $\rightarrow \rho &gt; 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSA C</td>
<td>Eq 24a</td>
<td>3 – 4</td>
<td>1 – 2, 5 – 85</td>
</tr>
<tr>
<td>CSA W</td>
<td>Eq 24b</td>
<td></td>
<td>1 – 500</td>
</tr>
<tr>
<td>STDM C</td>
<td>Eq 25a</td>
<td>1 – 11</td>
<td>12 – 16</td>
</tr>
<tr>
<td>STDM W</td>
<td>Eq 25b</td>
<td>1 – 16</td>
<td>17 – 25</td>
</tr>
<tr>
<td>STEM C</td>
<td>Eq 26a</td>
<td>1 – 7</td>
<td>8 – 9</td>
</tr>
<tr>
<td>STEM W</td>
<td>Eq 26b</td>
<td>1 – 10</td>
<td>11 – 14</td>
</tr>
<tr>
<td>STAR C</td>
<td>Eq 27a</td>
<td>1 – 1000</td>
<td></td>
</tr>
<tr>
<td>STAR W</td>
<td>Eq 27b</td>
<td>1 – 500</td>
<td>1000</td>
</tr>
</tbody>
</table>

Figure 3: The Garage coefficient for increasing NNs

Note: Each dot illustrates a different model at 5NN intervals from 5 - 250 NNs

To further analyze the issue, we also examine through iterative estimation a key structural characteristic most unlikely to pick up unaccounted spatially-delineated amenities. As
Figure 4 shows, the living area coefficient is consistent, around 0.64 to 0.66, across all models, with the exception of STAR. This implies some relative stability in estimation.

The STAR inconsistency, also reflected in a number of coefficients in Table 2, may imply bias issues in low NNs. This could stem from the lack of comparable adjustment process in this specification, as hypothesized in the methodological framework. This issue ameliorates as comparables increase in STAR C, but it is still present in STAR W, even at 1000NN. This suggests that the comparable adjustment process is essential for recovering coefficients, especially when comparable numbers are low.

Figure 4: Elasticity of living area for increasing NNs

Note: Each dot illustrates a different model at 5NN intervals from 5 - 250 NNs

4.2. Price Predictions

Typically, the real estate professional operates under informational and computational constraints. In this spirit, we also employ the simplest non-statistical mean NN price
calculation \((y = Cy \& y = Wy)\) without the attribute adjustment for comparison purposes. To compare the prediction performance of statistical and non-statistical methods, we display the mean square error (MSE) in Table 4, and the coefficient of dispersion\(^ {13} \) (COD) in Table 5. Only MSE is displayed here, as all other error metrics\(^ {14} \) present the same picture in terms of prediction accuracy comparison between the different specifications. We tested all models in the full range of NNs and found increasing number of NNs to deteriorate instead of increase prediction accuracy. Therefore, we employ 15 NNs for the spatiotemporal models in this exercise, as this is the level at which most models exhibit their best fit to the data.

Table 4: MSE between observed price and predictions

<table>
<thead>
<tr>
<th>Model (weight matrix, equation)</th>
<th>MSE of in-sample price predictions</th>
<th>One-step-ahead forecast MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>HP (eq 23)</td>
<td>0.22841</td>
<td>0.16371</td>
</tr>
<tr>
<td>NN ((y = Cy))</td>
<td>0.18840</td>
<td>0.17311</td>
</tr>
<tr>
<td>NN ((y = Wy))</td>
<td>0.16365</td>
<td>0.16474</td>
</tr>
<tr>
<td>CSA ((C, eq 24a))</td>
<td>0.09311</td>
<td>0.10715</td>
</tr>
<tr>
<td>CSA ((W, eq 24b))</td>
<td>0.08675</td>
<td>0.10674</td>
</tr>
<tr>
<td>STDM ((C, eq 25a))</td>
<td>0.09072</td>
<td>0.09688</td>
</tr>
<tr>
<td>STDM ((W, eq 25b))</td>
<td>0.08501</td>
<td>0.09596</td>
</tr>
<tr>
<td>STEM ((C, eq 26a))</td>
<td>0.27545</td>
<td>0.18928</td>
</tr>
<tr>
<td>STEM ((W, eq 26b))</td>
<td>0.26903</td>
<td>0.16363</td>
</tr>
<tr>
<td>STAR ((C, eq 27a))</td>
<td>0.11502</td>
<td>0.13533</td>
</tr>
<tr>
<td>STAR ((W, eq 27b))</td>
<td>0.10626</td>
<td>0.13318</td>
</tr>
</tbody>
</table>

Econometric CSA and STDM specifications provide superior performance to alternatives. HP and NN prediction precision seems to be more or less on par in one-step-ahead MSE, with NN performing much better at in-sample predictions. STAR models perform better than non-spatiotemporal models, but their precision is reduced in one-step-ahead process. STEM performance is poor, as it is on par with or worse than much simpler specification such as HP and NN.

\(^ {13} \)The coefficient of dispersion: COD = g/m, where g is the mean absolute deviation from the median and m is the median.

\(^ {14} \)MAPE, RMSE, and MAE metrics can be provided upon request.
COD measures dispersion or spread or variation of the predictions. Higher COD is in principle not positive (McCluskey and Borst, 2017); it may capture additional volatility in some predictors. However, the increased volatility may also mean that a model is more responsive to small (spatiotemporal) market shocks. Responsiveness to local shocks that increases prediction accuracy may also increase variability of the forecast. This is reflected in our results, as the simple NN models have the lowest COD for both prediction procedures, but also among the lowest prediction accuracy. Nevertheless, STDM with the highest prediction accuracy does not have highest COD. It is not unexpected for econometric CSA to show the highest COD in the one-step-ahead process, as these models would be more responsive to local shocks.

Table 5: COD between observed price and predictions

<table>
<thead>
<tr>
<th>Model (weight matrix, equation)</th>
<th>COD of in-sample price prediction</th>
<th>One-step-ahead forecast COD</th>
</tr>
</thead>
<tbody>
<tr>
<td>HP (eq 23)</td>
<td>5.431%</td>
<td>4.790%</td>
</tr>
<tr>
<td>NN (y = C y)</td>
<td>4.077%</td>
<td>4.491%</td>
</tr>
<tr>
<td>NN (y = W y)</td>
<td>4.248%</td>
<td>4.560%</td>
</tr>
<tr>
<td>CSA (C, eq 24a)</td>
<td>4.774%</td>
<td>5.249%</td>
</tr>
<tr>
<td>CSA (W, eq 24b)</td>
<td>4.812%</td>
<td>5.270%</td>
</tr>
<tr>
<td>STDM (C, eq 25a)</td>
<td>4.987%</td>
<td>4.902%</td>
</tr>
<tr>
<td>STDM (W, eq 25b)</td>
<td>4.986%</td>
<td>4.901%</td>
</tr>
<tr>
<td>STEM (C, eq 26a)</td>
<td>5.499%</td>
<td>4.819%</td>
</tr>
<tr>
<td>STEM (W, eq 26b)</td>
<td>5.472%</td>
<td>4.791%</td>
</tr>
<tr>
<td>STAR (C, eq 27a)</td>
<td>5.135%</td>
<td>5.172%</td>
</tr>
<tr>
<td>STAR (W, eq 27b)</td>
<td>5.138%</td>
<td>5.172%</td>
</tr>
</tbody>
</table>

Kok et al., (2017) show that “gradient boosting methods”, employing machine learning algorithms, significantly increase not only prediction precision over HP and other techniques, but also volatility (COD). Mayer et al. (2018) show that there indeed exist stability and accuracy trade-offs, especially for “gradient boosting methods” and advanced statistical techniques. This raises a significant trade-off issue for very large AVMs, as responsiveness and accuracy seem to go hand in hand. Hundreds of thousands of valuations jumping up to even small shocks in AVMs employed by government/wider-public-sector can have serious repercussions for an economy.
The spatiotemporal $W$ in comparison to $C$ improves MSE for in-sample and one-step-ahead prediction, while marginally reducing COD only for STDM and STEM. This means that only in these two models the spatiotemporal $W$ causes a slight divergence to the stability-accuracy trade-off highlighted by Mayer et al. (2018).

4.3. **Robustness checks and discussion**

We perform robustness checks by introducing a common spatially-delineated amenity to all the models in order to examine whether the effects of this variable are consistent with our methodological framework. We introduce a commonly used variable in HP studies, the distance to the city center, or “distance to center”. A number of variables in the previous models can be affected by the variation captured by “distance to center”, especially lot-size as land scarcity increases closer to city centers. Table 6 presents the introduction of distance to center to all the specifications at the NN levels of Table 2. We also report the results of a LR test on whether the extra constraint $\nu_1$ in equation 17 is superfluous when compared to equation 16. The distance to center coefficient $\gamma_1$ is positive and statistically significant, with comparable magnitudes, in both STEM and HP. The introduction of this variable in both HP and STEM did have a sizable impact on the magnitude of the lot-size coefficients, bringing their levels down towards those of CSA and STDM which did not change between Tables 2 and 6.

As predicted by our methodological framework and consistent with hypothesis 6, $\gamma_1$ is not statistically different to zero for econometric CSA specifications, which, in conjunction with the low magnitude of garage and lot-size coefficients, and the stability between Tables 2 and 6, implies minimized OVB. Minimized OVB is also a possibility for STDM specifications, as the hypotheses $\nu_1=0$ and $r_1=0$ cannot be rejected at the 99.9% level. Furthermore, the introduction of distance to center in STDM and CSA does not improve overall goodness of fit. For all other specifications, “distance to center” is statistically significant and improves overall goodness of fit.
Table 6: Models Including Distance Robustness checks

<table>
<thead>
<tr>
<th>Variables/Coefs</th>
<th>HP Eq23</th>
<th>CSA C NN8, Eq24</th>
<th>CSA W NN15, Eq24</th>
<th>STDM C NN6, Eq25</th>
<th>STDM W NN15, Eq25</th>
<th>STEM C NN15, Eq26</th>
<th>STEM W NN35, Eq26</th>
<th>STAR C NN6, Eq27</th>
<th>STAR W NN8, Eq27</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_1 ): age</td>
<td>-0.0189 ***</td>
<td>-0.0108 ***</td>
<td>-0.0109 ***</td>
<td>-0.0109 ***</td>
<td>-0.0108 ***</td>
<td>-0.0148 ***</td>
<td>-0.0147 ***</td>
<td>-0.0114 ***</td>
<td>-0.0108 ***</td>
</tr>
<tr>
<td>( \beta_2 ): ln_age</td>
<td>0.3146 ***</td>
<td>0.1738 ***</td>
<td>0.1818 ***</td>
<td>0.1762 ***</td>
<td>0.1805 ***</td>
<td>0.2051 ***</td>
<td>0.1991 ***</td>
<td>0.2134 ***</td>
<td>0.2078 ***</td>
</tr>
<tr>
<td>( \beta_3 ): ln_lot_size</td>
<td>0.0989</td>
<td>0.1045 ***</td>
<td>0.1059 ***</td>
<td>0.1025 ***</td>
<td>0.1069 ***</td>
<td>0.1345 ***</td>
<td>0.1381 ***</td>
<td>0.0693 ***</td>
<td>0.0647 ***</td>
</tr>
<tr>
<td>( \beta_4 ): ln_area</td>
<td>0.6678 ***</td>
<td>0.6626 ***</td>
<td>0.6501 ***</td>
<td>0.6534 ***</td>
<td>0.6441 ***</td>
<td>0.6794 ***</td>
<td>0.6784 ***</td>
<td>0.5643 ***</td>
<td>0.5465 ***</td>
</tr>
<tr>
<td>( \beta_5 ): Garage</td>
<td>0.3702 ***</td>
<td>0.2155 ***</td>
<td>0.2079 ***</td>
<td>0.2341 ***</td>
<td>0.2248 ***</td>
<td>0.2598 ***</td>
<td>0.2520 ***</td>
<td>0.2670 ***</td>
<td>0.2577 ***</td>
</tr>
<tr>
<td>( \beta_6 ): ln_bathrm</td>
<td>0.2152 ***</td>
<td>0.1958 ***</td>
<td>0.1883 ***</td>
<td>0.1979 ***</td>
<td>0.1896 ***</td>
<td>0.2273 ***</td>
<td>0.2163 ***</td>
<td>0.1438 ***</td>
<td>0.1311 ***</td>
</tr>
<tr>
<td>( \beta_7 ): Stories2+</td>
<td>0.1759 ***</td>
<td>0.0954 ***</td>
<td>0.0919 ***</td>
<td>0.0997 ***</td>
<td>0.0960 ***</td>
<td>0.1222 ***</td>
<td>0.1217 ***</td>
<td>0.1075 ***</td>
<td>0.1020 ***</td>
</tr>
<tr>
<td>( \gamma_1 ): Dist_Center</td>
<td>2.1E-05 ***</td>
<td>2.8E-07</td>
<td>2.4E-06</td>
<td>1.5E-05 ***</td>
<td>1.4E-05 ***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \mu_1 ): age</td>
<td>0.0062 ***</td>
<td>0.0082 ***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \mu_2 ): ln_age</td>
<td>-0.0780 ***</td>
<td>-0.1237 ***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \mu_3 ): ln_lot_size</td>
<td>-0.0826 ***</td>
<td>-0.0950 ***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \mu_4 ): ln_area</td>
<td>-0.5662 ***</td>
<td>-0.6240 ***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \mu_5 ): Garage</td>
<td>-0.0172</td>
<td>-0.0192</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \mu_6 ): ln_bathrm</td>
<td>-0.1107 **</td>
<td>-0.1352 **</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \mu_7 ): Stories2+</td>
<td>-0.0160</td>
<td>-0.0360 **</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( r_1 ): Dist_Center</td>
<td>4.2457 ***</td>
<td>1.4E-06</td>
<td>1.3E-06</td>
<td>-6.2E-06 ***</td>
<td>-6.0E-06 ***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>constant</td>
<td>4.2457 ***</td>
<td>0.5632 ***</td>
<td>-0.0197</td>
<td>4.1286 ***</td>
<td>4.1446 ***</td>
<td>-0.7056 ***</td>
<td>-0.8073 ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time fixed effects</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>( \rho )</td>
<td>0.9992 ***</td>
<td>0.9994 ***</td>
<td>0.8363 ***</td>
<td>0.9522 ***</td>
<td>1.1692 ***</td>
<td>1.2551 ***</td>
<td>0.5856 ***</td>
<td>0.6115 ***</td>
<td></td>
</tr>
<tr>
<td>LR ( \chi^2(1) ) test ( v_1=0 )</td>
<td>0.94</td>
<td>9.69 **</td>
<td>19.37 ***</td>
<td>116.51 ***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R(^2)</td>
<td>0.726</td>
<td>N/A</td>
<td>N/A</td>
<td>0.825</td>
<td>0.832</td>
<td>0.807</td>
<td>0.813</td>
<td>0.808</td>
<td>0.815</td>
</tr>
<tr>
<td>AIC/BIC</td>
<td>20371/20514</td>
<td>11498/11569</td>
<td>10488/10560</td>
<td>11139/11274</td>
<td>10208/10343</td>
<td>13172/13323</td>
<td>12476/12626</td>
<td>13050/13201</td>
<td>12290/12441</td>
</tr>
</tbody>
</table>

***p< 0.001; **p<0.01; *p<0.05
It is noteworthy that STAR is the only specification that has a negative coefficient for “distance to center”. When this is combined with \( \nu \neq 0 \), it suggests a biased \( r_1 \) coefficient dominated by the negative \( \nu \), as shown in equation 16 and consistent with hypothesis 7b. This is further confirmed by running the STAR specifications with \( \gamma_1 \) and \( \nu_1 \), instead of \( r_1 \). For both STAR C and W, the distance to center coefficient becomes positive and statistically significant at the 99.9% level. Its magnitude (1.4E-05) is almost identical to the STEM \( \gamma_1 \) and comparable to HP. \( \nu_1 \) is indeed negative with a larger magnitude (-2.88E-05) than \( \gamma_1 \) and significant at the 99.9% level.

We employ again an iterative estimation to determine the effects of introducing the distance to center variable across the full range of NNs, and examine hypothesis 7. Table 7 provides for every specification precise statistical testing on the NN range for which \( \gamma_1 \), \( r_1 \), and \( \nu_1 \) are lower, equal, or higher to zero. This allows us to deduce for different specifications the values of \( \zeta \) range in equation 18, within which the common spatially-delineated amenities might hold (\( \xi_{h\lambda} = \xi_{c\lambda} \)). We can also deduce the more restrictive range of possible minimization of spatially-delineated OVB. For CSA specifications, this is the NN range where \( \gamma_1 = 0 \), but for STAR and STDM this is the range where both \( r_1 = 0 \) and \( \nu_1 = 0 \). Looking at Table 7, the STEM coefficient is, as expected, positive and significant across the NN range.

In order to achieve a minimized spatially-delineated OVB, the NNs need to range between: 1 – 20 for CSA C; 1 – 35 for CSA W; 4 – 10 for STDM C; 5 – 15 for STDM W; 2 – 4 for STAR C; and 2 – 5 for STAR W. The superiority of the econometric CSA becomes obvious as OVB minimization across a wide range of NNs is combined with the lowest data and computational requirements of all specifications.

When the common spatially-delineated amenities assumption does not hold (\( \nu \neq 0 \)), running both STAR and STDM specifications with \( \gamma_1 \) and \( \nu_1 \), instead of \( r_1 \), returns a positive and significant \( \gamma_1 \) consistent with STEM and HP for most of the NNs. The spatiotemporal distance matrix \( W \) seems to provide the advantage of increasing the range of NNs across specifications when compared to its simpler spatiotemporal nearest neighbor matrix \( C \).
version. We can also test whether both STAR applicability conditions come into effect, rendering it preferable to STDM for any number of NNs \((v_1=0, r_1=0\) and \(\rho < 1\) for STAR at NNs where STDM \(\rho \geq 1\)). In our data and the range of NNs we tested \((1 – 1000)\), these conditions do not come into effect. In any case, STAR exhibits a worse fit to the data than STDM and CSA specifications do across the NN range. The fit to the data deteriorates as the number of NNs increases above 15 in all cases. It is also noted that there can be a number of nonlinearities not captured by the very limited number of variables in this exercise.

Table 7: Distance to center coefficients for a range of NNs at \(p < 0.001\) significance level

<table>
<thead>
<tr>
<th>Model</th>
<th>Coef</th>
<th>NN range → coef &lt; 0</th>
<th>NN range → coef = 0</th>
<th>NN range → coef &gt; 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSA C</td>
<td>(\gamma_1)</td>
<td>1 – 20</td>
<td>25 – 1000</td>
<td></td>
</tr>
<tr>
<td>CSA W</td>
<td>(\gamma_1)</td>
<td>1 – 35</td>
<td>40 – 1000</td>
<td></td>
</tr>
<tr>
<td>STDM C</td>
<td>(r_1) (v_1)</td>
<td>11 – 500</td>
<td>4 – 70</td>
<td>1 – 80 – 1000</td>
</tr>
<tr>
<td>STDM C</td>
<td>(r_1) (v_1)</td>
<td>2 – 10, 1000</td>
<td>1 – 1</td>
<td></td>
</tr>
<tr>
<td>STDM W</td>
<td>(r_1) (v_1)</td>
<td>16 – 1000</td>
<td>5 – 250</td>
<td>1 – 4, 500 – 1000</td>
</tr>
<tr>
<td>STEM C</td>
<td>(\gamma_1)</td>
<td>1 – 1000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STEM W</td>
<td>(\gamma_1)</td>
<td>1 – 1000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STAR C</td>
<td>(r_1) (v_1)</td>
<td>5 – 240</td>
<td>2 – 4, 250 – 500</td>
<td>1, 1000</td>
</tr>
<tr>
<td>STAR C</td>
<td>(r_1) (v_1)</td>
<td>5 – 1000</td>
<td>1 – 4</td>
<td></td>
</tr>
<tr>
<td>STAR W</td>
<td>(r_1) (v_1)</td>
<td>6 – 1000</td>
<td>2 – 5</td>
<td></td>
</tr>
<tr>
<td>STAR W</td>
<td>(r_1) (v_1)</td>
<td>8 – 1000</td>
<td>1 – 7</td>
<td></td>
</tr>
</tbody>
</table>

There is an issue that can potentially require further research, as STDM specifications at very high NNs essentially become the difference-from-the-mean models implied in equation 19. At 1000 NNs, the \(v_1\) of STDM C is not different to zero, which may imply that the common spatially-delineated amenities assumption holds when compared to the mean (see equation 19). In this case, STAR or a parsimonious STDM might be appropriate. This might also be reflected in STAR C and STDM C \(r_1\) coefficients turning positive at 1000NNs. Therefore, in contexts of spatially very dense observations and of intentional use of very high number of NNs, STAR could potentially be preferable to STDM. One instance we can envisage for such an application is a massive AVM, in which the modelers, for unknown reasons, have no interest in using limited comparables per observation, but
want to pursue difference-from-the-mean models. This would be unproductive for most housing AVMs, as model fit to data deteriorates with increasing numbers of NNs.

Summarizing, the econometric CSA offers a number of advantages, such as OVB minimization across a wide range of NNs, excellent fit to the data, and price prediction accuracy, combined with the lowest data and computational requirements of all specifications. Parsimonious STDM is also shown to share most of these advantages, albeit a lower NN range for OVB minimization and potentially higher computational requirements. STDM and econometric CSA are shown to be distinctly different from the Spatial Error Models and the STDM version previously derived by Pace (1998), Pace and Gilley (1998), Pace et al., (1998) and Pace et al. (2000). Even though STAR can potentially handle a very high number of comparables, its applicability over econometric CSA and STDM is limited.

5. Conclusions

Typical spatial econometric Hedonic Pricing (HP) approaches introduce spatially weighted terms and arbitrarily assert influence due to spatial proximity. A key contribution of this paper is to further develop the theoretical and methodological framework that explicitly integrates spatial (and temporal) econometric models to HP, based on the practice of the comparable sales approach (CSA) and nearest neighbors (NNs). We distinguish between comparability weights, which inform the comparable selection process, and housing attribute weights, which can only be given by the HP coefficients. We argue that the housing attribute distance/similarity metrics that currently dominate the real estate literature (McCluskey and Borst 2017) should not be central to the CSA; rather, proximity in space and time should be prominent.

An important contribution of this paper is the demonstration, through hypothesis testing and robustness checks, that the econometric CSA specifications and potentially a Spatiotemporal Durbin Model (STDM) reduce the spatially-delineated omitted variable
bias (OVB) typically plaguing HP. The preferred specification of the econometric CSA combines an excellent fit to the data, high price prediction accuracy, and OVB minimization, with the lowest data and computational requirements of all specifications. From the combination of CSA and HP we also derive the typical spatiotemporal autoregressive (STAR) and spatiotemporal error model (STEM), providing specific applicability conditions for the former. However, it is shown that none of the STAR or STEM specifications offer the advantages of the econometric CSA and STDM, at least for our dataset.

Our analysis reinforces the common real estate practice of selecting a small number of comparables in grid CSA (Ratcliffe 1972), which may not be the case for AVMs, where one can introduce hundreds or thousands of comparables. We take LeSage and Pace (2014) a step further to show that non-marginal increases to the number of spatiotemporal NNs, or comparables, deteriorate model fit and price predictions. This is on top of the potential bias, as a high number of comparables is shown to nullify the minimization of spatially-delineated OVB in econometric CSA and STDM. We also show that even though only STAR can potentially handle a very high number of comparables without overestimation issues, there are significant limitations in recovering unbiased HP coefficients of spatially-delineated amenities.

Potential limitations of our approach may involve nonlinearities, endogeneity, and/or spatial heterogeneity, which, along with further testing of the methodology across different markets and contexts, may offer new research opportunities. Avenues for future research also include development of housing-similarity constraints/metrics (Lindenthal, 2017) and innovative approaches of addressing the housing sub-market issue. As a spatially-delineated amenity can be observed to vary between comparables when all other local amenities remain constant, our framework can potentially provide opportunities in improving HP valuation of environmental externalities by integrating it to quasi-experimental approaches and boundary discontinuity design (Greenstone and Gayer, 2009; Chay and Greenstone, 2005).
References


LeSage, J.P., 1999. The theory and Practice of Spatial Econometrics, Department of Economics, University of Toledo, Texas, US.


