Dynamic spatiotemporal ARCH models

Philipp Otto, Osman Doğan and Süleyman Taşpınar

ABSTRACT
Geo-referenced data are characterised by an inherent spatial dependence due to geographical proximity. In this paper, we introduce a dynamic spatiotemporal autoregressive conditional heteroscedasticity (ARCH) process to describe the effects of (i) the log-squared time-lagged outcome variable, the temporal effect, (ii) the spatial lag of the log-squared outcome variable, the spatial effect, and (iii) the spatiotemporal effect on the volatility of an outcome variable. We derive a generalised method of moments (GMM) estimator based on the linear and quadratic moment conditions. We show the consistency and asymptotic normality of the GMM estimator. After studying the finite-sample performance in simulations, the model is demonstrated by analysing monthly log-returns of condominium prices in Berlin from 1995 to 2015, for which we found significant volatility spillovers.

KEYWORDS
Spatial ARCH, GMM, volatility clustering, volatility, house price returns, local real-estate market

JEL C13, C23, P25, R31

HISTORY Received 13 February 2023; in revised form 31 May 2023

1. INTRODUCTION
In a standard autoregressive conditional heteroskedasticity (ARCH) model, the volatility is modelled as a linear function of the lagged squared outcome variable in order to account for the volatility clustering patterns observed in the outcome variable (Bollerslev et al., 1992; Engle, 1982; Engle & Bollerslev, 1986). However, when analysing geo-referenced time series, a further phenomenon occurs – observations close in space are more similar than observations farther away – known as Tobler’s first law of geography (Tobler, 1970). From a statistical perspective, this spatial dependence may occur in the means and the volatility of a random process (Otto et al., 2018). Thus, in this paper, we extend the standard ARCH model to spatiotemporal data by using some tools from spatial econometrics. In our suggested specification, the log-volatility term may depend on (i) the log-squared time-lagged outcome variable, (ii) the higher-order spatial lags of the log-squared outcome variable, (iii) the higher-order spatial lags of the log-squared time-lagged outcome variable, (iv) exogenous variables and (v) the unobserved heterogeneity across regions and time. The presence of higher-order spatial lags in our specification indicates that the log-volatility term of a
region may depend on the current and time-lagged outcome variables in the neighbouring locations in differing ways, depending on the specifications of the associated spatial weights matrices (e.g., different influences from different directions or directional dependence, cf. Gupta & Robinson, 2015; Merk & Otto, 2021). We refer to this extended model as the dynamic spatiotemporal ARCH model.

To introduce an estimation approach for our model, we transform the outcome equation so that it is in the form of log-squared terms. We then substitute the log-volatility equation into the transformed outcome equation to obtain an estimation equation for the log-squared outcome variable. The resulting specification is in the form of a higher-order spatial dynamic panel data model with disturbance terms that may not have a zero mean. We use an orthonormal and a deviation from the group-mean operator to wipe out the regional and time fixed effects from the estimation specification. For the estimation of the transformed model that is free of the regional and time fixed effects, we propose a generalised method of moments (GMM) estimator formulated with a set of linear and quadratic moment functions (Lee, 2007; Lee & Liu, 2010; Lee & Yu, 2014). We show that the resulting GMM estimator has the standard large sample properties irrespective of whether the number of time periods is large or finite. When the number of time periods is large, the precision matrix of our GMM estimator simplifies significantly, allowing us to determine a set of linear and quadratic moment functions that can lead to an efficient estimator. We provide such a set of best linear and quadratic moment functions and establish the asymptotic properties of the resulting best GMM estimator. In a Monte Carlo simulation study, we show that the proposed GMM estimator performs well in finite samples.

In the literature, Robinson (2009) introduces the log-square transformation to a cross-sectional spatial stochastic volatility model and considers a quasi maximum likelihood (QML) estimation approach to estimate the parameters of the transformed model. Additionally, the log-square transformation has been used in the literature for the estimation of (i) the cross-sectional spatial stochastic volatility models (Taşpinar et al., 2021) and (ii) the cross-sectional spatial ARCH/generalised ARCH (GARCH) models (Otto, 2019; Otto & Schmid, 2023; Sato & Matsuda, 2017; Sato & Matsuda, 2021). In a similar manner, we may alternatively consider the QML method rather than the GMM approach for our transformed model (Hølleland & Karlsen, 2020; Lee & Yu, 2010; Yu et al., 2008). However, compared to the QML estimator, our GMM estimation approach has the following advantages. First, the GMM approach has the computational advantage over the QML method, which involves the computation of the determinant of a Jacobian term at each iteration during the estimation process. The computational cost can be especially high when the number of cross-sectional units is large. LeSage and Pace (2009) provide some solutions based on various approximation methods to reduce the computation cost significantly. Second, it is well known that the QML estimator has an asymptotic bias, requiring a bias correction approach even when the number of time periods is large (Lee & Yu, 2010; Yu et al., 2008). Finally, the QML estimator may have poor finite sample properties since the distribution of the log-squared disturbance terms in the transformed model is approximated by a normal distribution. In the time series literature on volatility models, it has been documented that the QML estimator obtained in this way has poor finite sample properties (Jacquier et al., 1994; Kim et al., 1998; Sandmann & Koopman, 1998; Shephard, 1994).

In an empirical application, we use a monthly dataset of the real house price returns in Berlin at the postcode level over the period, January 1995 to December 2015 to test the effect of temporal, spatial and spatiotemporal lags of the log-squared returns on the volatility of the log-returns. That is, we analyse the volatility of the house price returns in a real-estate market on a small geographic area of around 900 km² (see also Billè et al. 2017; Bogin et al. 2019; Zhang & Yi, 2017). In Section 2, we show that our dynamic spatiotemporal ARCH model
implies a spatial dynamic panel data model for the log-squared returns. Therefore, the presence of spatial, temporal and spatiotemporal effects in the log-squared returns will provide the empirical evidence for our suggested specification. To motivate the presence of these effects on the log-squared returns, Figure 1 displays the average log-squared returns over Berlin’s postcode areas (left figure), the estimated temporal autocorrelation of the log-squared returns as a series of boxplots (centre figure) and the estimated spatiotemporal autocorrelation in terms of Moran’s I across the time horizon (right figure). The first figure shows a clustering pattern in the log-squared returns, indicating the presence of spatial dependence. From the ACF estimates, we can observe a clear temporal volatility clustering, while the spatiotemporal dependence is of a minor degree, irregularly fluctuating around zero.

By using a first-order version of our dynamic spatiotemporal ARCH process for the local house price returns, we separately identify temporal, spatial and spatiotemporal interaction effects in the log-squared returns. Our estimation results show that the temporal, spatial and spatiotemporal lags of the log-squared returns have statistically significant effects on the log-volatility, and for that, there is significant variation in the log-volatility of the real house price returns in Berlin over its zip codes. This finding is not surprising because it has been documented in the literature that the spatial dependence in house price variations might arise due to several factors such as migration, equity transfer, spatial arbitrage and spatial patterns in the determinants of house prices (Meen, 1999). These patterns can change with the local infrastructure (Chang & Diao, 2021). Recently, Holmes et al. (2017) and Bashar (2021) particularly focus on intra-city house prices and show significant temporal and spatial dependence in the growth rates. In contrast to these studies, we focus on analysing the log-volatility as a measure of the market risk.

The rest of the paper proceeds in the following way. In Section 2, we state our model specification and discuss its properties. In Section 3, we provide the details of the GMM estimation approach for our model, and formally establish its large sample properties. In Section 4, we investigate the finite sample properties of our suggested algorithm through an extensive Monte Carlo study. In Section 5, we provide the details of our empirical application on Berlin’s house price returns. In Section 6, we offer our concluding comments with some directions for future studies. Some technical results are relegated to the Appendix in the supplemental data online.

Figure 1. Indication of spatial, temporal and spatiotemporal volatility clustering. Left: Average log-squared house price return for each of the 190 postcode areas over the period from February 1995 to December 2015. The map shows a weak clustering effect, especially for the outer regions (indicating a positive spatial dependence in the volatility). Centre: Temporal ACF is depicted as a series of boxplots showing the estimated temporal autocorrelation of all 190 locations. Right: Spatiotemporal correlation in terms of the slope of a regression line between the log squared returns and their temporally lagged neighbours (Moran’s I of first spatiotemporal lag). There is no clear pattern with varying coefficients around zero, which may indicate a weak spatiotemporal dependence.
2. MODEL SPECIFICATION

The outcome variable \( y_{it} \) of region \( i \) at time \( t \) is modelled according to

\[
y_{it} = \beta_{it}^{t} \varepsilon_{it},
\]

(1)

\[
\log \beta_{it} = \sum_{i=1}^{p} \sum_{j=1}^{n} \rho_{ij} m_{ij} \log y_{jt}^2 + \gamma_0 \log y_{i,t-1}^2 + \sum_{i=1}^{p} \sum_{j=1}^{n} \delta_{ij} m_{ij} \log y_{j,t-1}^2
\]

\[
+ \lambda_t \mathbf{B}_0 + \mu + \epsilon_{it},
\]

(2)

for \( i = 1, 2, \ldots, n \) and \( t = 1, \ldots, T \). The spatial locations indexed by \( i = 1, \ldots, n \) are supposed to be on discrete regular (e.g., for image processes) or irregular lattice, also known as spatial polygons. The latter case is typically present in economics, e.g., regional legal units, districts, countries, etc. Here, \( h_{it} \) is considered as the volatility term in region \( i \) at time \( t \), and \( \varepsilon_{it} \) are independent and identically distributed random variables that has mean zero and unit variance. The log-volatility terms follow the process in (2), where \( \{m_{ij}\}_{i=1}^{p} \) for \( i, j = 1, \ldots, n \), are the non-stochastic spatial weights. Here, \( \rho \) is a finite positive integer, and \( \{m_{ij}\}_{i=1}^{p} \) are zero for \( i = 1, \ldots, n \). The spatial, temporal and spatiotemporal effects of the log-squared outcome variable on the log-volatility are measured by the unknown parameters \( \gamma_0 \), \( \{\rho_{ij}\}_{i=1}^{p} \), and \( \{\delta_{ij}\}_{i=1}^{p} \), respectively. In (2), \( \mathbf{X}_t \) is a \( k \times 1 \) vector of exogenous variables with the associated parameter vector \( \mathbf{B}_0 \), and the regional and time fixed effects are denoted by \( \mathbf{\mu} = (\mu_{i0}, \ldots, \mu_{n0})' \) and \( \mathbf{\alpha}_0 = (\alpha_{10}, \ldots, \alpha_{T0})' \). Both \( \mathbf{\mu}_0 \) and \( \mathbf{\alpha}_0 \) can be correlated with the exogenous variables in an arbitrary manner. We assume that the initial value vector \( \mathbf{Y}_0 = (y_{10}, \ldots, y_{n0})' \) is observable.

Squaring both sides of (1) and then taking the natural logarithm yields

\[
y_{it}^* = \beta_{it}^* + \epsilon_{it},
\]

(3)

where \( y_{it}^* = \log y_{it}^2 \), \( \beta_{it}^* = \log \beta_{it} \) and \( \epsilon_{it} = \log \epsilon_{it}^2 \). In vector form, we can express (3) and (2) as

\[
\mathbf{Y}_t^* = \mathbf{h}_t^* + \mathbf{e}_t^*,
\]

(4)

\[
\mathbf{h}_t^* = \sum_{i=1}^{p} \rho_{0i} M_i \mathbf{Y}^*_{i-1} + \sum_{i=1}^{p} \delta_{0i} M_i \mathbf{Y}^*_{i-1} + \mathbf{X}_t \mathbf{B}_0 + \mathbf{\mu} + \mathbf{\alpha}_0 \mathbf{1}_n,
\]

(5)

where \( \mathbf{M}_i = (m_{ij}) \) is the \( n \times n \) spatial weights matrices, \( \mathbf{Y}_t^* = (y_{1t}, \ldots, y_{nt})' \), \( \mathbf{h}_t^* = (h_{1t}, \ldots, h_{nt})' \), \( \mathbf{e}_t^* = (e_{1t}, \ldots, e_{nt})' \), \( \mathbf{X}_t = (x_{1t}, \ldots, x_{nt})' \), and \( \mathbf{1}_n \) is the \( n \times 1 \) vector of ones. The process in (5) indicates that \( \mathbf{h}_t^* \) depends on the high order spatial lags of \( \mathbf{Y}_t^* \) and \( \mathbf{Y}_{t-1}^* \). Substituting (5) into (4), we obtain

\[
\mathbf{Y}_t^* = \sum_{i=1}^{p} \rho_{0i} M_i \mathbf{Y}^*_{i-1} + \sum_{i=1}^{p} \delta_{0i} M_i \mathbf{Y}^*_{i-1} + \mathbf{X}_t \mathbf{B}_0 + \mathbf{\mu} + \mathbf{\alpha}_0 \mathbf{1}_n + \mathbf{e}_t^*,
\]

(6)

for \( t = 1, \ldots, T \). Also, our ensuing analysis does not require that the orders of the spatial lags of \( \mathbf{Y}_t^* \) and \( \mathbf{Y}_{t-1}^* \) are the same. That is, we can consider (6) with \( \sum_{i=1}^{g} \rho_{0i} M_i \mathbf{Y}^*_{i-1} \) and \( \sum_{i=1}^{g} \delta_{0i} M_i \mathbf{Y}^*_{i-1} \), where \( g \) can be different from \( p \). This transformed model indicates that our specification in (2) implies a high-order spatial dynamic panel data model for the log-squared outcome variable. However, our transformed model differs from a standard spatial dynamic panel data model in two ways. Firstly, the dependent variable, the spatial lag term and the spatiotemporal lag term are formulated in terms of the log-squared outcome variable. Secondly, the disturbance term \( \mathbf{e}_t^* \) consists of the log-squared original disturbance terms, i.e., \( \epsilon_{it}^* = \log \epsilon_{it}^2 \). If we assume that
$\varepsilon_{it} \sim N(0, 1)$, then it follows that $\varepsilon^*_t$ will have a log-chi squared distribution with one degree of freedom ($\log \chi^2_1$). In this case, the density of $\varepsilon^*_t$ can be determined as

$$p(\varepsilon^*_t) = \frac{1}{\sqrt{2\pi}} \exp\left( -\frac{1}{2} (\exp(\varepsilon^*_t) - \varepsilon^*_t)^2 \right), \quad -\infty < \varepsilon^*_t < \infty,$$

for $i = 1, \ldots, n$ and $t = 1, \ldots, T$. It is also known that $E(\varepsilon^*_t) = -\gamma - \log(2) \approx -1.2704$ with $\gamma$ being Euler's constant, and $\text{Var}(\varepsilon^*_t) = \pi^2/2 \approx 4.9348$ (Lee, 2012, pp. 379–380). Figure 2 shows the density plots of the log $\chi^2_1$ distribution and the corresponding normal distribution. The plot clearly illustrates that the log-chi squared distribution exhibits significant skewness with a long left tail. This analysis also suggests that the QML method will perform poorly, especially for small samples, because the normal distribution provides a poor approximation to the log $\chi^2_1$ distribution. In the next section, we suggest using a GMM estimator that does not rely on any distributional assumption for the original disturbance terms.

### 3. THE ESTIMATION APPROACH

The elements of $\varepsilon^*_t$ in (6) are i.i.d. across $i$ and $t$ but their mean may not be zero. Therefore, we add and subtract $E(\varepsilon^*_t)$ to obtain the following equation:

$$Y^*_t = \sum_{l=1}^{p} \rho_{0l} M_l Y^*_{t_l} + \gamma_0 Y^*_{t-1} + \sum_{l=1}^{p} \delta_{0l} M_l Y^*_{t_l-1} + X_t \beta_0 + \mu_0 + \alpha_0 1_n + \mu_e 1_n + U_t,$$

where $U_t = (u_{1t}, \ldots, u_{nt}) = \varepsilon^*_t - E(\varepsilon^*_t)$, and $\mu_e = E(\varepsilon^*_t)$. Let $\sigma^2_0 = E(u^2_{it})$. Then, it follows that the elements of $U_t$ are i.i.d. across $i$ and $t$ with mean zero and variance $\sigma^2_0$. We need to eliminate both fixed effects terms from the model in order to avoid the incidental parameter problem. To eliminate $\mu_0$ and $\mu_e 1_n$ from the model, we consider an orthonormal transformation based on the matrix decomposition of $J_T = (I_T - (1/T)1_T 1_T^T)$, where $I_T$ is the $T \times T$ identity matrix. Let $(F_{T,T-1}, (1/\sqrt{T})1_T)$ be the orthonormal eigenvector matrix of $J_T$, where $F_{T,T-1}$ is the $T \times (T-1)$ sub-matrix containing eigenvectors corresponding to the eigenvalues of one. Let
\( C = (c_1, \ldots, c_T) \) be an \( n \times T \) matrix, where \( c_t \) is an \( n \times 1 \) vector for \( t = 1, \ldots, T \). Using \( F_{T,T-1} \), we can transform \( C \) into a \( n \times (T - 1) \) matrix in the following way: 

\[
(c_1^*, \ldots, c_{T-1}^*) = (c_1, \ldots, c_T)F_{T,T-1},
\]

where \( c_j^* \) is the \( j \)th column of \( CF_{T,T-1} \) for \( j = 1, \ldots, T - 1 \). If we apply \( F_{T,T-1} \) to our model in (8) in a similar manner, we obtain

\[
Y_t^{**} = \sum_{l=1}^{p} \rho_l m_l Y_{t-l}^{**-1} + \sum_{l=1}^{p} \gamma_{l0} Y_{t-l}^{**-1} + \sum_{l=1}^{p} \gamma_{l0} Y_{t-l}^{**-1} + X_t^* \beta_0 + \alpha_{t0} 1_n + U_t^*,
\]

for \( t = 1, \ldots, T - 1 \), where

\[
(Y_1^{**}, \ldots, Y_{T-1}^{**}) = (Y_1, \ldots, Y_T) F_{T,T-1},
\]

\( (Y_0^{**-1}, \ldots, Y_{T-2}^{**-1}) = (Y_0, \ldots, Y_{T-1}^*) F_{T,T-1}, \) and

\( (X_{11}, \ldots, X_{1T}) F_{T,T-1}, \)

where \( X_{l1} \) is the \( l \)th column of \( X \) for \( l = 1, \ldots, k, (\alpha_1^*, \ldots, \alpha_{T-1}^*) = (\alpha_1, \ldots, \alpha_T) F_{T,T-1} \), and \( (U_1^*, \ldots, U_{T-1}^*) = (U_1, \ldots, U_T) F_{T,T-1} \). It is important to note that the transformation breaks the dynamic structure in our model, since \( Y_{t-1}^{**-1} \) is not equal to the time lag of \( Y_t^{**} \), i.e., \( Y_{t-1}^{**-1} \neq Y_{t-1}^{**} \).

Note that both \( \mu_0 \) and \( \mu_e 1_n \) are dropped from the model since

\[
(\mu_0, \ldots, \mu_0) F_{T,T-1} = \mu_0 1_T F_{T,T-1} = 0_{n \times (T-1)},
\]

\[
(\mu_e 1_n, \ldots, \mu_e 1_n) F_{T,T-1} = \mu_e 1_T F_{T,T-1} = 0_{n \times (T-1)}.
\]

Let \( N = n(T-1) \) and \( U_N = (U_1^*, \ldots, U_{T-1}^*)' \). Note that we can express \( (U_1^*, \ldots, U_{T-1}^*) = (U_1, \ldots, U_T) F_{T,T-1} \) as \( (U_1^*, \ldots, U_{T-1}^*)' = (F_{T,T-1} \otimes 1_n)(U_1, \ldots, U_T)' \).

Then, it follows that

\[
\mathbb{E}(U_N U_N') = \mathbb{E}((F_{T,T-1} \otimes 1_n)(U_1, \ldots, U_T)(F_{T,T-1} \otimes 1_n)) = \alpha_0^2 1_N,
\]

indicating that the elements of \( U_N \) are uncorrelated. Among the orthonormal transformations, Lee and Yu (2014) show that the forward orthogonal difference (the Helmert transformation) can be useful for the spatial dynamic panel data models. Thus, we have the following explicit forms for the transformed variables:

\[
Y_t^{**} = \left( \frac{T-t}{T-t+1} \right) \left( Y_t^{**} - \frac{1}{T-t} \sum_{b=t+1}^{T} Y_b^{**} \right),
\]

\[
Y_{t-1}^{**-1} = \left( \frac{T-t}{T-t+1} \right) \left( Y_{t-1}^{**} - \frac{1}{T-t} \sum_{b=t+1}^{T-1} Y_b^{**} \right)
\]

and the other variables are expressed similarly.

The transformed model in (9) includes the transformed time fixed effects. These terms can be eliminated by pre-multiplying the model with \( J_n = (1_n - (1/n) 1_n 1_n) \) to get

\[
J_n Y_t^{**} = \sum_{l=1}^{p} \rho_l J_n m_l Y_{t-l}^{**} + \gamma_{l0} J_n Y_{t-l}^{**-1} + \sum_{l=1}^{p} \gamma_{l0} J_n m_l Y_{t-l}^{**-1} + J_n X_t^* \beta_0 + J_n U_t^*,
\]

(10)

where we used the fact that \( J_n 1_n = 0_n \). Our GMM estimation approach is based on (10). It is clear that we need to determine instrumental variables (IVs) for the following terms:

\[
\{ M_j Y_{t}^{**} \}_{j=1}^{p} \text{ and } \{ M_j Y_{t-1}^{**-1} \}_{j=1}^{p} \text{ for } t = 1, \ldots, T - 1.
\]

That is, we need IVs for the following variables:

\[
J_n \{ M Y_t^{**}, Y_{t-1}^{**-1}, M Y_{t-1}^{**-1} \},
\]

(11)

where \( M Y_t^{**} = (M_1 Y_t^{**}, \ldots, M_p Y_t^{**}) \) and \( M Y_{t-1}^{**-1} = (M_1 Y_{t-1}^{**-1}, \ldots, M_p Y_{t-1}^{**-1}) \). Let \( F_{T-1} \) be the \( \sigma \)-algebra generated by \( \{ Y_0, \ldots, Y_{T-1} \} \) conditional on \( \{ X_1, \ldots, X_T, \mu_0, \alpha_0 \} \). Then, we can formulate the theoretical linear IVs based on the expectation of (11) conditional on \( F_{T-1} \).

We use \( \rho_0 = (\rho_{t0}, \ldots, \rho_{p0})' \) and \( \delta_0 = (\delta_{t0}, \ldots, \delta_{p0})' \) to denote the true parameter values and \( \rho = (\rho_1, \ldots, \rho_p)' \) and \( \delta = (\delta_1, \ldots, \delta_p)' \) to denote arbitrary parameter values.
Let \( S(\rho) = (I_n - \sum_{i=1}^p \rho_i M_i) \), \( A(\rho, \delta, \gamma) = S^{-1}(\rho)(\gamma I_n + \sum_{i=1}^p \delta_i M_i) \) and \( A = A(\rho_0, \delta_0, \gamma_0) \). Then, the reduced form of (9) can be expressed as

\[
Y_{t-1}^* = AY_{t-1}^{ss-1} + S^{-1}(X_t^*\beta_0 + \alpha_0^* 1_n + U_t^*). \tag{12}
\]

Let \( Z_{t}^* = (Y_{t-1}^{ss-1}, \cdots, X_{t}^*) \) be the \( n \times k_z \) matrix, where \( k_z = p + k + 1 \), and \( Z_N = (Z_1^*, \cdots, Z_{T-1}^*)' \). Then, using (9), we have

\[
M_1 Y_t^* = G_t(Z_t^* \eta_0 + \alpha_0^* 1_n) + G_t U_t^*, \quad r = 1, \ldots, p, \tag{13}
\]

where \( \eta_0 = (\gamma_0, \delta_0, \beta_0' )' \) and \( G_t = M_t S^{-1} \). We can use (13) to determine IVs for \( (M_1 Y_{t-1})_{j=1}^p \). In the case of \( Y_{t-1}^{ss-1} \), we can use all strictly exogenous variables \( X_s^* \) for \( s = 1, \ldots, T - 1 \), and the time lag variables \( Y_0^*, \ldots, Y_{t-1}^* \) as IVs. Similarly, we can use \( M_i X_s^* \) for \( s = 1, \ldots, T - 1 \), and \( M_i Y_s^* \) for \( s = 0, 1, \ldots, t - 1 \) as IVs for \( M_i Y_{t-1}^{ss-1} \). Let \( Q_t \) be the \( n \times k_q \) matrix of IVs for \( t = 1, \ldots, T - 1 \), where \( k_q \geq k + 2p + 1 \). For example, we may choose \( Q_t \) as

\[
(Y_{t-1}^*, M_2 Y_{t-1}^*, X_{t}^*, M_1 X_s^*, M_2^2 X_s^*), \tag{14}
\]

where \( M_2 Y_{t-1}^* = (M_2^2 Y_{t-1}^*, \ldots, M_1 M_p Y_{t-1}^*, \ldots, M_1 M_p M_1 Y_{t-1}^*, \ldots, M_1 M_p M_2 Y_{t-1}^*, \ldots, M_1 M_p M_2^2 Y_{t-1}^*) \), and \( M_2^2 X_s^* \) is defined similarly. Additionally, it is worth noting that in the absence of any external variables in the model, we can construct linear IVs by considering only the variable \( Y_{t-1}^* \) and its spatial lag terms. Denote \( Q_N = (Q_1^*, \ldots, Q_{T-1}^*)' \), \( J_N = I_{T-1} \otimes J_n \) and \( S_N(\rho) = I_{T-1} \otimes S(\rho) \). Then, the linear moment conditions based on \( Q_N \) can be formulated as

\[
Q_N^* J_N U_N(\theta), \tag{15}
\]

where \( \theta = (\rho', \eta', \beta')' \), \( U_N(\theta) = (U_1^* (\theta), \ldots, U_{T-1}^* (\theta))' \) and \( U_t^* (\theta) = S(\rho)Y_t^* - Z_t^* \eta - \alpha_t^* 1_n \). Note that the transformed time fixed effects \( \alpha_t^* = (\alpha_1^*, \ldots, \alpha_{T-1}^*)' \) will be eliminated in the moment function because \( U_N(\theta) \) is pre-multiplied by \( J_N \).

Following Lee (2007) and Lee and Yu (2014), we also consider the quadratic moment functions for estimation. The quadratic moment functions are based on the idea that the vector \( P_t J_s U_t^* \) can be uncorrelated with \( J_s U_t^* \) for an \( n \times n \) matrix \( P_t \) satisfying \( tr(J_t P_t J_s) = 0 \), while it may be correlated with \( G_t U_t^* \) in (13). Let \( P_N = I_{T-1} \otimes P_b \) and assume that there are \( m \) such quadratic moment matrices. Then, the quadratic moment functions can be expressed as

\[
U_N(\theta) J_N P_d N U_N(\theta), \tag{16}
\]

for \( l = 1, 2, \ldots, m \). Combining the linear and quadratic moment functions, we obtain the following vector of moment functions,

\[
g_N(\theta) = \begin{pmatrix}
U_N(\theta) J_N P_{1N} J_N U_N(\theta) \\
\vdots \\
U_N(\theta) J_N P_{mN} J_N U_N(\theta) \\
Q_N J_N U_N(\theta)
\end{pmatrix}. \tag{17}
\]

Let \( \text{vec}(P) \) be the vectorisation of the square matrix \( P \), \( \text{vecD}(P) \) be the column vector formed from the diagonal elements of \( P \) and \( P' = P + P' \). Define \( \Omega_N = (1/N) E(\tilde{g}_N(\theta_0) g_N(\theta_0)') \). Then, using
Lemma 1 in the Appendix, it can be shown that \(^2\)

\[
\Omega_N = \plim_{n \to \infty} \frac{\Delta_{mN}}{N} \left( \frac{1}{\sigma_0^2} \mathbf{Q}'_N J_N Q_N \right) + \lim_{n \to \infty} \frac{\mu_4 - 3 \sigma_0^4}{N} \left( \mathbf{w}_m^N \mathbf{w}_m^N \right) \left( \frac{1}{\sigma_0^2} \mathbf{Q}'_N \right) \left( \frac{1}{\sigma_0^2} \mathbf{Q}_N \right),
\]

(18)

where \(\mu_4\) is the fourth moment of \(u_it\), \(\mathbf{w}_m^N = (\text{vec}_D(J_N P_{1N}^t J_N), \ldots, \text{vec}_D(J_N P_{mN}^t J_N))\) and

\[
\Delta_{mN} = (\text{vec}(J_N P_{1N}^t), \ldots, \text{vec}(J_N P_{mN}^t J_N))',
\]

\[
\times (\text{vec}(J_N P_{1N}^t J_N), \ldots, \text{vec}(J_N P_{mN}^t J_N)).
\]

Let \(\hat{\Omega}_N\) be a consistent estimator of \(\Omega_N\), i.e., \(\hat{\Omega}_N - \Omega_N = o_p(1)\). Then, the optimal GMM estimator is defined as

\[
\hat{\theta}_N = \arg\min_{\theta \in \mathcal{G}_N}(\theta) \hat{\Omega}_N^{-1} g_N(\theta).
\]

(19)

To investigate the asymptotic properties of \(\hat{\theta}_N\), we require the following assumptions.

Assumption 1. The disturbance terms \(u_{it}\), for \(i = 1, 2, \ldots, n\), and \(t = 1, 2, \ldots, T\), are i.i.d. across \(i\) and \(t\) with mean zero, variance \(\sigma_0^2\) and \(\mathbb{E}(u_it^{1+\kappa}) < \infty\) for some \(\kappa > 0\).

Assumption 2. The spatial weights matrices \(\{M_i\}_{i=1}^q\) are uniformly bounded in both row and column sums in absolute value.

Assumption 3. (i) \(S(\mathbf{r})\) is invertible for all \(\mathbf{r} \in \Lambda\), where \(\Lambda\) is a compact parameter space, and \(\mathbf{r}_0\) is in the interior of \(\Lambda\). (ii) \(S^{-1}(\mathbf{r})\) is uniformly bounded in both row and column sums in absolute value.

Assumption 4. (i) \(X_t\) is non-stochastic with \(\sum_{t=1}^{T} \sum_{i=1}^{n} |x_{it}|^{1+\epsilon} < \infty\) for some \(\epsilon > 0\), where \(x_{it}\) is the \((i, t)\)th element of the \(l\)th regressor for \(l = 1, \ldots, k\). Moreover, \(\lim_{n \to \infty} (1/N)X_N J_N X_N\) exists and is non-singular, where \(X_N = (X_1^t, \ldots, X_{T-1}^t)\). (ii) \(\mu_0\) and \(\alpha_0\) are non-stochastic with \(\sup_n (1/n) \sum_{i=1}^{n} |\mu_0|^2 + \epsilon < \infty\) and \(\sup_n (1/T) \sum_{i=1}^{T} |\alpha_0|^2 + \epsilon < \infty\).

Assumption 5. (i) \(Y_0^t = \sum_{b=0}^{\bar{b}} A^t A^{-1} (X_{-t} \beta_0 + \mu_0 + \alpha_{-b} 1_n + \mu_1 1_n + U_{-b})\), where \(\bar{b}\) can be finite or infinite. (ii) \(\sum_{b=0}^{\infty} \text{abs}(A^t)\) is uniformly bounded in both row and column sums in absolute value, where the \((i, j)\)th element of \(\text{abs}(A)\) is given by \(|A_{ij}|\) and \(A_{ij}\) is the \((i, j)\)th element of \(A\).

Assumption 6. \(\mathbb{E}(Q_{4i} | F_{t-1}) = Q_{4i}\) and \(\mathbb{E}(|q_{it}|^{1+\epsilon}) < \infty\), where \(q_{it}\) is the \((i, t)\)th element of the \(l\)th column of \(Q_{4i}\). Moreover, \(\plim_{n \to \infty} (1/N)Q_N J_N (Z_N, L_N)\) and \(\plim_{n \to \infty} (1/N)Q_N N Q_N\) have full column ranks, where \(L_N = (L_1, \ldots, L_{T-1})'\) with \(L_r = (L_{1r}, \ldots, L_{pr})\) and \(L_{r,t} = G_s(Z_{st}^* \eta_0 + \alpha_0^* 1_n)\) for \(r = 1, 2, \ldots, p\).

Assumption 1 specifies the distribution of the elements of \(U_t\) for \(t = 1, 2, \ldots, T\). The moment condition in this assumption is required to show the asymptotic distribution of our set of moment functions. Assumptions 2 and 3 are standard assumptions adopted in the literature to limit the degree of spatial correlation at a manageable degree, e.g., among others, see Kelejian and Prucha (2010), Lee (2004). Assumption 4 provides the regularity conditions for \(X_N, \mu_0\) and \(\alpha_0\). The first part of Assumption 5 specifies \(Y_0^t\), and the remaining parts are required to limit dependence over time and across cross-sectional units (see Lee and Yu (2014) for the details). The sufficient conditions for the first part of Assumption 3, and the second part of 5 can be determined. Let \(|| \cdot ||\) be any matrix norm. Then, the following respective conditions will be sufficient
to ensure these parts: (i) $\left\| \sum_{j=1}^{p} \rho_j M_j \right\| < 1$ and (ii) $\left\| A(\rho, \delta, \gamma) \right\| < 1$. Note that

$$\left\| \sum_{j=1}^{p} \rho_j M_j \right\| \leq |\rho_1| \cdot |M_1| + \ldots + |\rho_p| \cdot |M_p| \leq \left( \sum_{j=1}^{p} |\rho_j| \right) \times \max_{1 \leq j \leq p} |M_j|.$$ 

Thus, a relatively restrictive condition for (i) is $\left( \sum_{j=1}^{p} |\rho_j| \right) \times \max_{1 \leq j \leq p} |M_j| < 1$. Similarly, we have

$$||A(\rho, \delta, \gamma)|| \leq ||S^{-1}(\rho)|| \times \left\| \gamma_1 + \sum_{j=1}^{p} \delta_j M_j \right\|$$

$$= \left\| I_n + \left( \sum_{j=1}^{p} \rho_j M_j \right) + \left( \sum_{j=1}^{p} \rho_j M_j \right)^2 + \ldots \right\| \times \gamma_1 + \sum_{j=1}^{p} \delta_j M_j$$

$$\leq \frac{1}{1 - \tau_1} \times \left( |\gamma| + \left( \sum_{j=1}^{p} |\delta_j| \right) \times \max_{1 \leq j \leq p} |M_j| \right),$$

where $\tau_1 = (\sum_{j=1}^{p} |\rho_j|) \times \max_{1 \leq j \leq p} |M_j| < 1$ is guaranteed by the first condition. This result suggests that a relatively restrictive condition for (ii) is

$$\frac{1}{1 - \tau_1} \times \left( |\gamma| + \left( \sum_{j=1}^{p} |\delta_j| \right) \times \max_{1 \leq j \leq p} |M_j| \right) < 1.$$

When the spatial weights matrices are row normalised these relatively restrictive conditions can be further simplified. For example, if we use the matrix row sum norm, we will get the following sufficient conditions: (i) $\left( \sum_{j=1}^{p} |\rho_j| \right) < 1$ and (ii) $\left( \sum_{j=1}^{p} |\rho_j| + |\gamma| + \sum_{j=1}^{p} |\delta_j| \right) < 1$.

Assumption 6 provides the regularity conditions for the IV matrix $Q_\gamma$. The first part states that $Q_\gamma$ is pre-determined in the sense that $\mathbb{E}(Q_\gamma | F_{t-1}) = Q_\gamma$. The moment condition in this assumption is required for the application of a CLT to the set of our moment functions (see the CLT given in Lemma 3 in the Appendix). The full column rank condition in Assumption 6 gives the identification condition based on the linear moment function in our setting. See Appendix, Section B, for the details on the identification condition in our setting.

Let $\partial g_N(\theta)/\partial \theta = (\partial g_N(\theta)/\partial \rho, (\partial g_N(\theta)/\partial \rho))$. In Section C of the Appendix, we show that $1/N(\partial g_N(\theta))/\partial \theta = D_{1N} + D_{2N} + O_p(N^{-1/2})$, where $D_{1N} = O(1)$ and $D_{2N} = O(T^{-1})$. The following result gives the limiting distribution of $\hat{\theta}_N$ under the large $T$ and finite $T$ cases.

**Theorem 1.** Under Assumptions 1–6, we have the following results,

(1) When $T$ is finite and $n \to \infty$, we have

$$\sqrt{n}(\hat{\theta}_N - \theta_0) \overset{d}{\to} \mathcal{N}\left(0_{k+\hat{p}}, \lim_{n \to \infty} \frac{1}{T-1} (D_{1N} + D_{2N}) \Omega_N^{-1}(D_{1N} + D_{2N})^{-1}\right).$$

(2) When $T \to \infty$ and $n \to \infty$, we have

$$\sqrt{N}(\hat{\theta}_N - \theta_0) \overset{d}{\to} \mathcal{N}(0_{k+\hat{p}}, \lim_{n,T \to \infty} (D_{1N}' \Omega_N^{-1}D_{1N})^{-1}).$$
Our estimator defined in (19) requires a consistent estimator of $\Omega_N$. We can use a plug-in estimator of $\Omega_N$ based on an initial GMM estimator, or alternatively, we can formulate a 2SLS estimator based on $Q_N$. Let $Y_N = (Y_{1s}, \ldots, Y_{Ts})'$ and $Z_N = (Z_{1s}, \ldots, Z_{Ts})'$. Then, the 2SLS estimator is

$$\hat{\theta}_N = ((\mathbb{I}_N Y_N, Z_N)M_Q(\mathbb{I}_N Y_N, Z_N))^{-1}(\mathbb{I}_N Y_N, Z_N)'M_Q Y_N,$$

(22)

where $M_Q = J_N Q_N^{-1} Q_N ' J_N$ and $\mathbb{I}_N Y_N = (M_1 N Y_N, \ldots, M_p N Y_N)$ with $M_{ij} = I_{t-1} \otimes M_j$ for $j = 1, 2, \ldots, p$. Our Theorem 1 suggests that

$$\sqrt{n}(\hat{\theta}_N - \theta_0) \xrightarrow{d} N(0, \sigma_0^2 \text{plim}_{n \to \infty} \left( \frac{1}{N} (L_N, Z_N)' M_Q (L_N, Z_N) \right)^{-1}).$$

(23)

In the case of both the initial GMM and 2SLS estimators, the linear IV matrix can be $Q_N = (Y_{t-1}^*, MY_{t-1}^*, MY_{t-1}^2, X_{t-1}^*, MX_{t-1}^*, MX_{t-1}^2)$ for $t = 1, 2, \ldots, T - 1$. We may consider the following quadratic moment matrices for the initial GMM estimator:

$$P_j = \left(M_j - \frac{\text{tr}(M_j J_J)}{n-1} J_J \right)$$
and

$$P_{j+p} = \left(M_j^2 - \frac{\text{tr}(M_j^2 J_J)}{n-1} J_J \right)$$
for $j = 1, 2, \ldots, p$.

Then, the initial GMM estimator is given by $\hat{\theta}_N = \arg\min_{\theta \in \Theta} g_N(\theta)' g_N(\theta)$ where

$$g_N(\theta) = \begin{bmatrix} U_N(\theta) J_N P_{1N} J_N U_N(\theta) \\ \vdots \\ U_N(\theta) J_N P_{pN} J_N U_N(\theta) \\ Q_N ' J_N U_N(\theta) \end{bmatrix}.$$ 

(24)

We can use $\hat{\theta}_N$ to formulate the plug-in estimator of $\Omega_N$, which requires the estimators of $\sigma_0^2$ and $\mu_4$. Let $\tilde{V}_t = S(\tilde{p}_N) Y_{ts}^* - Z_{ts}^* \tilde{\eta}_N$. Then, we can estimate $\sigma_0^2$ by $\tilde{\sigma}_N^2 = \frac{1}{N} \sum_{t=1}^{T-1} \tilde{V}_t J_N \tilde{V}_t$. Let $\Delta \tilde{V}_t = S(\tilde{p}_N) \Delta Y_{ts}^* - \Delta Z_{ts}^* \tilde{\eta}_N$. Then, following Lee and Yu (2014), we can estimate $\mu_4$ by $\tilde{\mu}_4 = \frac{1}{2N} \sum \sum (1/2) [J_4 \Delta \tilde{V}_t]_i^4 - 3 \tilde{\sigma}_0^4$, where $[J_4 \Delta \tilde{V}_t]_i$ is the $i$th element of $J_4 \Delta \tilde{V}_t$.

Our set of moment functions in (16) depends on the IV matrix $Q_N$ and the quadratic moment matrices $Q_{IN}$ for $l = 1, 2, \ldots, m$. The asymptotic efficiency of $\hat{\theta}_N$ should be considered in choosing the IV and quadratic moment matrices. The best set of IV and quadratic moment matrices is the set that leads to the most efficient GMM estimator. When $T$ is large, the precision matrix $\hat{\theta}_N$ takes a simple form allowing for determining the best set of IV and quadratic moment matrices. When $T$ is large, the proof of Theorem 1 indicates that

$$\frac{1}{N} \frac{\partial g_N(\theta_0)}{\partial \theta} = D_{1N} + O_p(N^{-1/2}),$$

where

$$D_{1N} = \frac{1}{N} \left( \begin{array}{cc} \sigma_0^2 C_N & 0_{m \times k} \\ Q_N ' J_N L_N & Q_N ' J_N Z_N \end{array} \right).$$

(25)
Then, (21) shows that the precision matrix of $\sqrt{N}(\hat{\theta}_N - \theta_0)$ is

$$
D_{1N}'\Omega_N^{-1}D_{1N} = \frac{1}{N}
\begin{pmatrix}
C_N'\left(\Delta_{mN} + \frac{\mu_4 - 3\sigma_0^4}{\sigma_0^4}\omega_{mN}'\omega_{mN}\right)^{-1}C_N & 0_{p \times k_z} \\
0_{k_z \times p} & 0_{k_z \times k_z}
\end{pmatrix}
+ \frac{1}{N\sigma_0^2}(L_N, Z_N)'M_\Omega(L_N, Z_N).
$$

(26)

Since the above precision matrix has the same form as the one given in Lee and Yu (2014), we use their approach to determine the best set of quadratic moment matrices. We will choose the best quadratic matrices by maximising $C_N\left(\Delta_{mN} + \frac{\mu_4 - 3\sigma_0^4}{\sigma_0^4}\omega_{mN}'\omega_{mN}\right)^{-1}C_N$. As shown in Lee and Yu (2014), these matrices are

$$
P^*_j = \left(G_j - \frac{\text{tr}(G_j\Sigma^*_j)}{n}J_n\right) + c\left(\text{Diag}(J_nG_jJ_n) - \frac{\text{tr}(G_jJ_n)}{n}I_n\right), \quad j = 1, 2, \ldots, p,
$$

(27)

where $c = \left(\frac{n}{n-2}\right)^2\left(\frac{1}{n/(n-2) + (\eta_4 - 3)/2} - \frac{n-2}{n}\right)$ and $\eta_4 = \mu_4/\sigma_0^4$.

In the case of the best linear moment function, we should consider the conditional mean $E(M_{Y_t}, Z_t^{**}|F_{t-1})$. The conditional mean of $Y_t^{**}$ can be determined from $Y_{t-1}^{**} = \zeta_t\left(Y_{t-1} - \frac{1}{T-t}\sum_{i=t}^{T-1}Y_i\right)$, where $\zeta_t = \left(\frac{T-t}{T-t+1}\right)^{1/2}$. Using Lemma 4 in the Appendix, $E(Y_{t-1}^{**}|F_{t-1})$ can be approximated by

$$
H_t = \zeta_t \left(I_n - \frac{1}{T-t}\sum_{b=1}^{T-t}A^b\right)Y_{t-1} - \frac{1}{T-t}\sum_{r=t}^{T-1}\left(\sum_{b=0}^{T-r-1}A^b\right)S^{-1}(X_r\beta_0 + \alpha_{r,0}1_n)
$$

$$
-\zeta_t\left(\frac{1}{(T-t)(t-1)}\sum_{r=t}^{T-1}\sum_{b=0}^{T-r-1}A^b\right)S^{-1}\sum_{i=1}^{T-r-1}(SY_i - Z_i\eta_0 - \alpha_{i,0}1_n),
$$

where $Z_i^* = (Y_{i-1}^{**}, M_{Y_i}^{**}, X_i)$. Thus, the best theoretical IV $J_nE(Y_{t-1}^{**}|F_{t-1})$ can be approximated by $J_nH_t$. Similarly, the best IVs for $J_nZ_{t}^{**}$ can be taken as $J_nK_t$, where $K_t = (H_t, M_{H_t}, X_t)$. Using (13), the best IV for $J_nM_{Y_t}^{**}$ is $J_nG_t(K_t\delta_0 + \alpha_{r,0}1_n)$ for $r = 1, 2, \ldots, p$. Overall, we may use $J_nQ^*_t$ as the IV matrix for $J_n(M_{Y_t}^{**}, Z_t^{**})$, where

$$
Q^*_t = \left(G_t(K_t\delta_0 + \alpha_{r,0}1_n), \ldots, G_p(K_t\delta_0 + \alpha_{r,0}1_n), K_t\right), \quad t = 1, 2, \ldots, T - 1.
$$

(28)

The feasible version of $Q^*_t$ can be obtained by substituting consistent estimators of the unknown parameters into (28). Note that if $M_11_n = 1_n$, i.e., when $M_t$ is row-normalised, we have $J_nM_t = J_nM_01_n$. This property suggests that the time fixed effects will be dropped from $J_nQ^*_t$ since $J_n1_n = 0_n$. However, if $M_t$'s are not row normalised, then we also need an estimator of the time fixed effects to get a feasible version of $Q^*_t$. Let $\hat{\delta}_t = S(\delta_N)Y_t - Z_t\hat{\eta}_N$ be an estimator of $\mu_0 + \mu_e1_n + \alpha_{i,0}1_n$. Under the normalisation assumption of the form $1'_n(\mu_0 + \mu_e1_n) = 0_n$, we can estimate the time fixed effects by $\hat{\alpha}_{t,N} = \frac{1}{n}1'_n\hat{\delta}_t$ for $t = 1, 2, \ldots, T$. The following theorem provides our result on the best GMM estimator formulated with the feasible versions of $Q^*_t$ and $P^*_j$ for $j = 1, 2, \ldots, p$. 
Theorem 2. Let \( \hat{Q}_t \) be the feasible version of \( Q_t^\ast \) for \( t = 1, 2, \ldots, T - 1 \), and \( \hat{P}_j^\ast \) be the feasible version of \( P_j^\ast \) for \( j = 1, 2, \ldots, p \). Consider the set of moment functions \( g_N(\theta) \) formulated with \( \hat{Q}_t \) and \( \hat{P}_j^\ast \). Then, the feasible best GMM estimator defined by \( \hat{\theta}_N^\ast = \arg\min_{\theta} \ell_N(\theta) \hat{\Omega}_N^{-1} g_N(\theta) \) has the following asymptotic distribution

\[
\sqrt{N}(\hat{\theta}_N^\ast - \theta_0) \overset{d}{\rightarrow} N(0_{k+r}, \Sigma_N^{s-1}),
\]

where

\[
\Sigma_N^{s} = \lim_{n,T \to \infty} \left( C_N^s / N \begin{bmatrix} 0_{p \times k} \end{bmatrix} \right) + \plim_{n,T \to \infty} \frac{1}{N\sigma_0^2}(L_N, Z_N)J_N(L_N, Z_N),
\]

with

\[
C_N^s = \begin{pmatrix} \text{tr}(G_1N^s_1N^p_1N^a_1N^s_1N) & \cdots & \text{tr}(G_1N^s_1N^pN^a_1N^pN) \\ \vdots & \ddots & \vdots \\ \text{tr}(G_1N^s_1N^pN^a_1N^pN) & \cdots & \text{tr}(G_1N^a_1N^pN^a_1N^pN) \end{pmatrix}.
\]

Proof. See Section D of the Appendix in the supplemental data online.

4. A MONTE CARLO STUDY

In this section, we investigate the finite sample properties of the best GMM estimator provided in Theorem 2. To that end, we consider \( y_{it} = h_{it}^{1/2}\varepsilon_{it} \), and the following cases for \( h_{it} \):

\[
M_1: \log h_{it} = \sum_{j=1}^n \rho_0 m_{ij} \log y_{jt}^2 + \gamma_0 \log y_{jt-1}^2 + \sum_{j=1}^n \delta_0 m_{ij} \log y_{jt-1}^2 + \mathbf{x}'_it \mathbf{b}_0 + \mu_{i0} + \alpha_{i0}
\]

\[
M_2: \log h_{it} = \sum_{j=1}^n \rho_0 m_{ij} \log y_{jt}^2 + \gamma_0 \log y_{jt-1}^2 + \sum_{j=1}^n \delta_0 m_{ij} \log y_{jt-1}^2 + \mathbf{x}'_it \mathbf{b}_0 + \mu_{i0},
\]

\[
M_3: \log h_{it} = \sum_{l=1}^L \sum_{j=1}^n \rho_{0l} m_{ij} \log y_{jt}^2 + \gamma_0 \log y_{jt-1}^2 + \sum_{l=1}^L \sum_{j=1}^n \delta_{0l} m_{ij} \log y_{jt-1}^2
\]

\[+ \mathbf{x}'_it \mathbf{b}_0 + \mu_{i0} + \alpha_{i0},\]

where \( \mu_{i0} \)‘s and \( \alpha_{i0} \)‘s are i.i.d. \( N(0, 1) \), and \( \mathbf{x}_it \sim \text{i.i.d. } N(0_{2 \times 1}, \mathbf{I}_2) \) with \( \mathbf{b}_0 = (0.5, 1)' \). For the first two models, denoted \( M_1 \) and \( M_2 \), we consider two different temporal dependence structures – a weakly temporal dependent model and a strongly persistent model. Specifically, we set \( (\rho_0, \gamma_0, \delta_0)' = \{(0.2, 0.2, -0.2)', (0.2, 0.8, -0.2)'\} \) in \( M_1 \) and \( M_2 \), respectively. Moreover, \( M_2 \) considers the case without temporal fixed effects, i.e., \( \alpha_{i0} = 0 \) for all \( t \). In \( M_3 \), including higher-order spatial lags, we set \( (\rho_{10}, \rho_{20}, \gamma_0, \delta_{10}, \delta_{20}) = (0.6, 0.2, 0.1, 0.01, 0.01)' \). That is, we have very weak temporal and spatiotemporal effects. In all cases, we consider row-normalised queen contiguity spatial weights matrices, where \( \mathbf{M}_1 \) has positive weights for the first-lag neighbours and \( \mathbf{M}_2 \) for the second-lag neighbours. Furthermore, we consider two distributions to generate the disturbance terms: (i) \( \varepsilon_{it} \sim N(0, 1) \) and (ii) \( \varepsilon_{it} \sim t_3 \), where \( t_3 \) is the Student’s \( t \) distribution with 3 degrees of freedom. We set \( (n, T) = \{(64, 5), (64, 20), (100, 10), (100, 40)\} \), and the number of
repetitions to 1000 in all cases. Note that we also include the case with a very small time horizon of \( T = 5 \), which is sometimes encountered in spatiotemporal applications. Thus, we considered 12 different model specifications in total.

The results of our Monte Carlo simulation study are reported in Tables 1–3. To evaluate the estimation performance, we report the average bias across all replications and the mean absolute errors (MAE). For all simulation settings, our theoretical findings are supported in the finite sample case. More precisely, when \( n \) and \( T \) increase, our suggested GMM estimator reports a smaller bias and MAE in all cases. Comparing the performance with respect to the error distribution, we see slightly lower MAEs in the heavy-tailed case. These differences are insignificant in almost all cases (\( \alpha = 0.05 \)). Unsurprisingly, when \( T = 5 \), the GMM estimator reports relatively larger bias and MAE for all parameters, as the information about the spatial structure is effectively based on only \( T - 1 = 4 \) repeated observations of the spatial field. However, as \( T \) increases, the estimator performs satisfactorily in all cases. Overall, these results indicate that our suggested GMM estimator has good finite sample properties in terms of bias and MAE.

5. REAL-WORLD EXAMPLE: INTRA-CITY HOUSING MARKET RISK

The real-estate market is undoubtedly a financial market with the most apparent spatial and temporal dependence. The location of a property, along with size and condition, is an essential price-determining influence. Hence, there are pronounced spatial spillover effects in real-estate prices, in addition to the natural temporal dependence. Furthermore, taxes may significantly affect the market, such as property taxes or real-estate transfer taxes. In general, however, taxes appear to play a subordinate role in purchasing decisions – with one exception, namely, if the transfer taxes change, some sales could be shifted for a certain period. If, for example, the land transfer tax increases by one percentage point and one wants to buy a property in January, it is profitable to conclude the purchase contract already in December. This results in a shift in property sales from January to December, and thus more sales in December and fewer sales in January than expected. However, does this also impact the risk of the real-estate market?

For the empirical analysis, we use monthly log-returns of the average sales prices of all condominium sales in all postcode regions of Berlin from January 1995 to December 2015 (see Figure 3, left). The relative price per square metre is determined for each postcode area from an average of 6.31 sales per month. The specific location of the German capital, Berlin, in the centre of another federal state, Brandenburg, makes it a very intriguing example. The surrounding area of Berlin is very well connected to the city centre by public transport and infrastructure, such that exogenous effects such as tax changes in Brandenburg may have an impact on Berlin and vice versa. Every real-estate purchase in Germany is subject to the real-estate transfer tax, which must be paid once at the time of purchase. The amount of tax depends on the purchase price. Until 1 September 2006, a unified tax rate of 3.5% was applied in Germany as a whole. Afterwards, each federal state could set its tax rate, and there were gradual increases in all federal states. Specifically, Berlin increased the tax rates from 3.5% to 4.5% on 1 January 2007, from 4.5% to 5% on 1 April 2012, and from 5% to 6% on 1 January 2014. The shifting effects described above can also be observed for Berlin, as is shown in Figure 3. In addition, an end-of-year effect is clearly visible due to other accounting and tax reasons. This motivates why we would expect different market risks at the end and beginning of a year. To estimate these temporal effects, we consider a model without temporal fixed effects (i.e., \( \alpha_t = 0 \) for all \( t \) in (2)) and model the temporal effects by including yearly and monthly indicator variables as regressors.

We estimate a first-order version of our model in (2). We specify the spatial weights matrix (row-standardised) based on neighbourhood relations, where all adjacent neighbours are equally weighted. In Table 4 and Figure 4, we present the estimated parameters of our dynamic spatiotemporal ARCH model and the estimated volatility, respectively. The included covariates were
Table 1. Average bias and mean absolute errors (MAEs) of the estimated parameters of Model $M_1$.

|          | $n = 64$ | $n = 100$ | $n = 64$ | $n = 100$ |
|----------|----------|-----------|----------|-----------|
|          | $T = 5$  | $T = 20$  | $T = 10$ | $T = 40$  |
| Bias     |          |           |          |           |
| $\rho_0$ = 0.2 | 0.0316   | 0.0034    | 0.0092   | 0.0041    |
| $\gamma_0$ = 0.2 | -0.0118  | 0.0001    | -0.0001  | -0.0008   |
| $\delta_0$ = -0.2 | 0.0000   | -0.0013   | -0.0017  | -0.0004   |
| $\beta_{00}$ = 0.5 | -0.0062  | -0.0033   | -0.0040  | -0.0023   |
| $\beta_{10}$ = 1 | -0.0238  | -0.0056   | -0.0046  | -0.0007   |
| MAE      |          |           |          |           |
| $\rho_0$ = 0.2 | 0.2895   | 0.1142    | 0.1379   | 0.0590    |
| $\gamma_0$ = 0.2 | 0.0820   | 0.0266    | 0.0345   | 0.0139    |
| $\delta_0$ = -0.2 | 0.1867   | 0.0612    | 0.0792   | 0.0321    |
| $\beta_{00}$ = 0.5 | 0.1327   | 0.0527    | 0.0662   | 0.0292    |
| $\beta_{10}$ = 1 | 0.1357   | 0.0517    | 0.0637   | 0.0287    |

For $\varepsilon_{it} \sim N(0, 1)$

|          | $n = 64$ | $n = 100$ | $n = 64$ | $n = 100$ |
|----------|----------|-----------|----------|-----------|
|          | $T = 5$  | $T = 20$  | $T = 10$ | $T = 40$  |
| Bias     |          |           |          |           |
| $\rho_0$ = 0.2 | 0.0685   | -0.0001   | 0.0020   | 0.0027    |
| $\gamma_0$ = 0.2 | -0.0088  | -0.0004   | -0.0011  | 0.0005    |
| $\delta_0$ = -0.2 | -0.0054  | 0.0007    | 0.0058   | -0.0009   |
| $\beta_{00}$ = 0.5 | -0.0025  | -0.0020   | -0.0026  | -0.0003   |
| $\beta_{10}$ = 1 | -0.0164  | -0.0026   | -0.0020  | -0.0035   |
| MAE      |          |           |          |           |
| $\rho_0$ = 0.2 | 0.3038   | 0.1183    | 0.1442   | 0.0653    |
| $\gamma_0$ = 0.2 | 0.0862   | 0.0253    | 0.0345   | 0.0134    |
| $\delta_0$ = -0.2 | 0.1905   | 0.0606    | 0.0818   | 0.0348    |
| $\beta_{00}$ = 0.5 | 0.1512   | 0.0581    | 0.0715   | 0.0324    |
| $\beta_{10}$ = 1 | 0.1448   | 0.0573    | 0.0714   | 0.0318    |

For $\varepsilon_{it} \sim t_3$
Table 2. Average bias and mean absolute errors (MAEs) of the estimated parameters of Model $M_2$.

|        | $i.i.d. \varepsilon_{it} \sim N(0, 1)$ | $i.i.d. \varepsilon_{it} \sim t_3$ |
|--------|----------------------------------------|------------------------------------|
|        | $n = 64$ | $n = 100$ | $n = 64$ | $n = 100$ |
|        | $T = 5$  | $T = 20$  | $T = 10$  | $T = 40$  | $T = 5$  | $T = 20$  | $T = 10$  | $T = 40$  |
| Bias   |          |          |          |          |          |          |          |          |
| $\rho_0 = 0.2$ | 0.0839 | 0.0214 | 0.0389 | 0.0119 | 0.0982 | 0.0204 | 0.0339 | 0.0035 |
| $\gamma_0 = 0.8$ | −0.0539 | −0.0014 | −0.0037 | −0.0012 | −0.0525 | −0.0025 | −0.0072 | −0.0013 |
| $\delta_0 = −0.2$ | −0.0036 | 0.0109 | 0.0017 | −0.0019 | −0.0158 | −0.0036 | −0.0038 | 0.0033 |
| $\beta_{00} = 0.5$ | −0.0221 | −0.0031 | −0.0050 | 0.0001 | −0.0107 | −0.0039 | −0.0059 | −0.0014 |
| $\beta_{10} = 1$ | −0.0462 | −0.0054 | −0.0080 | −0.0018 | −0.0295 | −0.0091 | −0.0100 | 0.0002 |
| MAE    |          |          |          |          |          |          |          |          |
| $\rho_0 = 0.2$ | 0.3766 | 0.1096 | 0.1409 | 0.0582 | 0.3657 | 0.1161 | 0.1617 | 0.0623 |
| $\gamma_0 = 0.8$ | 0.2560 | 0.0372 | 0.0679 | 0.0167 | 0.2382 | 0.0383 | 0.0678 | 0.0157 |
| $\delta_0 = −0.2$ | 0.4409 | 0.1184 | 0.1771 | 0.0627 | 0.4225 | 0.1238 | 0.1733 | 0.0643 |
| $\beta_{00} = 0.5$ | 0.1519 | 0.0526 | 0.0624 | 0.0296 | 0.1608 | 0.0553 | 0.0669 | 0.0321 |
| $\beta_{10} = 1$ | 0.1788 | 0.0532 | 0.0671 | 0.0276 | 0.1873 | 0.0581 | 0.0736 | 0.0308 |
Table 3. Average bias and mean absolute errors (MAEs) of the estimated parameters of Model $M_3$.

| Bias       | $n=64$     | $n=100$   | $n=64$     | $n=100$   |
|------------|------------|-----------|------------|-----------|
|            | $T=5$      | $T=20$    | $T=10$     | $T=40$    |
| $\rho_{10}$ = 0.6 | 0.0929     | 0.0112    | 0.0214     | 0.0026    |
| $\rho_{20}$ = 0.2 | 0.0685     | 0.0122    | 0.0179     | 0.0041    |
| $\gamma_0$ = 0.1 | -0.0174    | -0.0007   | -0.0044    | -0.0004   |
| $\delta_{10}$ = 0.01 | -0.0158    | -0.0066   | -0.0059    | -0.0014   |
| $\delta_{20}$ = 0.01 | -0.0390    | -0.0059   | -0.0116    | -0.0023   |
| $\beta_{00}$ = 0.5 | -0.0143    | -0.0006   | -0.0065    | -0.0022   |
| $\beta_{10}$ = 1  | -0.0184    | -0.0064   | -0.0133    | -0.0046   |
| $\rho_{10}$ = 0.6 | 0.2090     | 0.0794    | 0.0986     | 0.0458    |
| $\rho_{20}$ = 0.2 | 0.2786     | 0.1183    | 0.1375     | 0.0650    |
| $\gamma_0$ = 0.1 | 0.0774     | 0.0254    | 0.0342     | 0.0136    |

| MAE        | $n=64$     | $n=100$   | $n=64$     | $n=100$   |
|------------|------------|-----------|------------|-----------|
|            | $T=5$      | $T=20$    | $T=10$     | $T=40$    |
| $\rho_{10}$ = 0.6 | 0.2257     | 0.0864    | 0.1077     | 0.0479    |
| $\rho_{20}$ = 0.2 | 0.2923     | 0.1271    | 0.1482     | 0.0659    |
| $\gamma_0$ = 0.1 | 0.0794     | 0.0261    | 0.0356     | 0.0135    |
| $\delta_{10}$ = 0.01 | 0.1397     | 0.0514    | 0.0654     | 0.0277    |
| $\delta_{10}$ = 0.01 | 0.1917     | 0.0646    | 0.0780     | 0.0338    |
| $\beta_{00}$ = 0.5 | 0.1359     | 0.0559    | 0.0625     | 0.0293    |
| $\beta_{10}$ = 1  | 0.1386     | 0.0534    | 0.0660     | 0.0306    |
selected by stepwise excluding regressors, such that the Bayesian information criterion is mini-
mised. The overall market dynamic measured by the total number of real-estate transactions
has the most significant effect on the volatility (see Table 4). The more transactions, the
lower the log-volatility. In addition, we observe lower volatilities at the end of each year,
and significantly higher volatilities in January and February. These effects decrease from January
to February (0.2152 to 0.1465). The March effects were already insignificant. Further, we
observe two periods of significantly lower risks compared to the remaining periods, namely
1997–2001 and 2011–2013. The anticipated tax effects are not significant, though. Thus, we
could not find evidence that the legal changes in the taxation framework affect the volatility
of log-returns.

As expected, we also see significant spatial and spatiotemporal spill-over effects. The spatial
ARCH parameter $\hat{p} = 0.4032$ is on a moderate level. That is, an increase in the log-squared
return in one location instantaneously increases the log-volatility in the adjacent regions. Further,
the temporal dependence is composed of a purely temporal lag ($\hat{\gamma} = 0.1913$) and a spatiotem-
poral lag ($\hat{\delta} = -0.0737$), which in total indicate a moderate temporal persistence.

The log-volatility $h_t$ can be regarded as an indicator of the risk associated with house prices. It
enables the identification of areas with high volatility and facilitates closer monitoring of these
areas in the future. To get an estimate of $h_t$, we consider
$$
\hat{h}_t^* = Y^* - \hat{\gamma}_N Y^*_{t-1} - \hat{\delta} X_t \hat{\beta}_N
$$
as an estimate of the sum of the fixed effects $(\mu_0 + \alpha_0 \mathbf{1}_n)$. Then, under the normalisation restriction $\mathbf{1}_n^\prime \mu_0 = 0$, we can consider
$$
\hat{\alpha}_{tN} = \frac{1}{n} \mathbf{1}_n^\prime \hat{h}_t^*
$$
as an estimator of $\alpha_0$. This, in turn, suggests that we can use
$$
\hat{\mu}_N = \frac{1}{T} \sum_{t=1}^T \left( \hat{h}_t^* - \frac{1}{n} \mathbf{1}_n^\prime \hat{h}_t^* \mathbf{1}_n \right)
$$
as an estimator of $\mu_0$. Then, using $\hat{\theta}_N$, $\hat{\mu}_N$ and $\hat{\alpha}_{tN}$ in (5), we
can formulate an estimate of $h_t^*$. Finally, we can recover an estimate of $h_t$ using the relationship
$h_t = \exp(\hat{h}_t^*)$. Figure 4 presents the estimated volatilities. When analysing the estimated
volatility, we clearly see temporal patterns with reduced risks during the two above-mentioned
periods. More interestingly, there are several regions of higher volatility, mostly located at the

Figure 3. Overview of the data set. Left: Average house prices for each of the 190 postcode areas over
the period from January 1995 to December 2015. Right: Total number of real-estate transactions in all
postcode areas of Berlin. The vertical bars indicate the time points of changes in the real-estate transfer
taxes in Berlin or the surrounding state of Brandenburg (red: increase from 3.5% to 4.5% in Berlin;
green: increase from 3.5% to 5% in Brandenburg; blue: increase from 4.5% to 5% in Berlin; yellow:
increase from 5% to 6% in Berlin).

Note: Readers of the print article can view the figures in colour online at https://doi.org/10.1080/
17421772.2023.2254817
outer postcode regions in the north and northwest, as shown in the top panels of Figure 4, where the averaged volatility estimates are depicted. Moreover, the average volatility across all postcodes changes over time, as shown in the first figure of the top panels of Figure 4. Finally, to illustrate the estimated hits for one selected location, we show the results for Berlin-Tempelhof (postcode of the closed airport Berlin-Tempelhof) in April 2012, the month after the increase of the real-estate transfer taxes from 4.5% to 5%. We do not see different patterns in the volatility estimates after the airport was closed, marked by the dashed black line.

Table 4. Estimated parameters of the dynamic ARCH process and diagnostic measures of the residuals.

| Parameter | Estimate | Standard error | T-statistics |
|-----------|----------|----------------|--------------|
| **Regressive effects** | | | |
| Total number of transactions | $\beta_1$ | -0.1356 | 0.0184 | -7.3509 |
| **End-of-year effects** | | | |
| December | $\beta_2$ | -0.0516 | 0.0465 | -1.1098 |
| January | $\beta_3$ | 0.2152 | 0.0729 | 2.9527 |
| February | $\beta_4$ | 0.1465 | 0.0555 | 2.6379 |
| **Yearly effects** | | | |
| 1997 | $\beta_5$ | -0.1447 | 0.0603 | -2.4002 |
| 1998 | $\beta_6$ | -0.2112 | 0.0735 | -2.8739 |
| 1999 | $\beta_7$ | -0.2003 | 0.0708 | -2.8292 |
| 2000 | $\beta_8$ | -0.1136 | 0.0585 | -1.9413 |
| 2001 | $\beta_9$ | -0.1194 | 0.0600 | -1.9897 |
| 2011 | $\beta_{10}$ | -0.1244 | 0.0674 | -1.8466 |
| 2012 | $\beta_{11}$ | -0.1161 | 0.0684 | -1.6974 |
| 2013 | $\beta_{12}$ | -0.1204 | 0.0684 | -1.7606 |
| **Tax effects** | | | |
| Month before tax increase | $\beta_{13}$ | -0.0559 | 0.1120 | -0.4988 |
| **Spatiotemporal effects** | | | |
| Spatial interaction (contiguity-based) | $\rho$ | 0.4032 | 0.1450 | 2.7801 |
| Temporal interaction (first time lag) | $\gamma$ | 0.1913 | 0.0051 | 37.4971 |
| Spatiotemporal interaction | $\delta$ | -0.0737 | 0.0313 | -2.3547 |
| **Model diagnostics** | | | |
| BIC | 82572.28 | | |
| N | 190 | | |
| T | 239 | | |
| Percentage of locations with significantly (temp.) autocorrelated errors ($\alpha = 5\%$) | 18 | | |
| Percentage of time points with significantly (spatially) positive autocorrelated errors (Moran’s I, $\alpha = 5\%$) | 0 | | |
6. CONCLUSION

In this paper, we introduced a dynamic spatiotemporal ARCH model that allows for unobserved heterogeneity over time and space. The model can be used to describe the spatiotemporal clustering effect in the volatility of a random process. As typically observed for spatial data, the model allows for instantaneous spill-over effects across space, which is the main difference to multivariate time-series GARCH models. In the latter case, spatial interactions would only occur after one time lag. In addition to these instantaneous spatial effects, the model includes temporal and spatiotemporal autoregressive effects of the log-squared returns in the log-volatility equation.

Figure 4. Estimated conditional volatility $\hat{h}_{it}$. Top left: The spatially averaged volatilities levels are displayed over time $\frac{1}{n} \sum_{i=1}^{n} \hat{h}_{it}$. The vertical grey line marks the Euro introduction in January 2002. Top right: The temporally averaged volatilities levels are displayed on the map $\frac{1}{T} \sum_{t=1}^{T} \hat{h}_{it}$. Bottom left: $\hat{h}_{it}$ of one selected region over time (Berlin Tempelhof, postcode of the former airport, marked by a red cross in the bottom right plot). Bottom right: $\hat{h}_{it}$ at one selected time point (April 2012, first month after the increase of real-estate transfer taxes from 4.5% to 5%, shown by the dashed red line in the bottom left plot).
While the temporal effect measures the dependence between the current and past observation of the same spatial unit, the spatiotemporal coefficients describe the dependence between the observation in one location and its past observations at neighbouring locations.

For our suggested dynamic spatiotemporal ARCH model, we obtain an estimation equation by applying a log-square transformation together with an orthonormal and a deviation from the group-mean operator to eliminate the fixed effects. We introduced a GMM estimation approach based on a set of linear and quadratic moment functions of the transformed process. We establish the consistency and asymptotic normality of our suggested GMM estimator under fairly general assumptions for large and finite $T$ cases. Moreover, when the number of time periods is large, we present an optimal set of moment functions that leads to an efficient estimator.

We investigated the finite-sample performance of our suggested estimator in a series of Monte Carlo simulations under different model settings and error distributions. Overall, the simulation results are in line with our theoretical claims. In an empirical application, we illustrated the use of our model for the log-returns of the intra-city real-estate prices in Berlin over the period 1995 to 2015. Our estimation results show that the spatial, temporal and spatiotemporal lags of the log-squared returns have statistically significant effects on the log-volatility. This leads to temporal and spatial spillover effects. We showed that the average volatility of log-returns over space and time varies significantly. Finally, our model allows us to estimate the market risk in terms of the volatility in each location and time point. In the future, it would also be interesting to conduct this analysis of house price risks based on sophisticated hedonic price indices in order to eliminate further price-determining factors that may appear unbalanced in the current sample based on average sales values.

In future studies, our model can be extended in a number of ways. First, we considered additive time and space fixed effects in the log-volatility equation. Instead of this additive structure, a log-volatility equation that includes interactive fixed effects can be studied. Second, the spatial and spatiotemporal lags in the log-volatility equation can be formulated with time-varying spatial weights matrices. Finally, we can also allow for potential endogeneity in the spatial weights instead of exogenous spatial weights. All of these extensions can be explored in future studies.

**ACKNOWLEDGEMENTS**

This paper is based on the preprint arXiv:2202.13856

**DISCLOSURE STATEMENT**

No potential conflict of interest was reported by the author(s).

**NOTES**

1 Note that the matrix equation $ABC = D$, where $D, A, B,$ and $C$ are suitable matrices, can be expressed as $\text{vec}(D) = (C' \otimes A)\text{vec}(B)$, where $\text{vec}(B)$ denotes the vectorisation of the matrix $B$ (Abadir & Magnus, 2005, p. 282). This property can be applied to $(U_1, U_2, \ldots, U_{T-1}) = (U_1, U_2, \ldots, U_T)F_{T,T-1}$ by setting $D = (U_1^*, U_2^*, \ldots, U_{T-1}^*)$, $C = F_{T,T-1}$, $B = (U_1, U_2, \ldots, U_T)$ and $A = I_n$.

2 In applying Lemma 1 in the Appendix in the supplemental data online, we use the fact that $\text{tr}(A'B) = \text{vec'}(A)\text{vec}(B) = \text{vec'}(B)\text{vec}(A)$, where $A$ and $B$ are any two $N \times N$ matrices.
The explicit forms of $D_{1N}$ and $D_{2N}$ are given in Section C of the Appendix.

Note that when $t = 1$, we may simply use $H_1 = c_1\left((1_n - \frac{1}{T-1}\sum_{h=1}^{T-1}A_h^b)Y_0^n - \frac{1}{T-1}\sum_{r=1}^{T-1}\left(\sum_{h=0}^{r-1}A_h^b\right)S^{-1}(X_r\beta_0 + \alpha_r1_n)\right)$.

Note that when $T$ is large, $\tilde{\mu}_0 = (\mu_0 + \mu_\epsilon1_n)$ can be estimated by $\hat{\mu}_N = \frac{1}{T}\sum_{t=1}^{T}\left(\hat{\theta}_t - \frac{1}{n}1_n^\prime\hat{\theta}_t1_n\right)$.

ORCID

Philipp Otto http://orcid.org/0000-0002-9796-6682
Osman Doğan http://orcid.org/0000-0001-7324-9454
Süleyman Taşpinar http://orcid.org/0000-0001-7372-995X

REFERENCES

Abadir, K. M., & Magnus, J. R. (2005). Matrix algebra. Cambridge University Press.
Bashar, O. H. (2021). An intra-city analysis of house price convergence and spatial dependence. The Journal of Real Estate Finance and Economics, 63(4), 525–546. https://doi.org/10.1007/s11146-020-09799-w
Billè, A. G., Benedetti, R., & Postiglione, P. (2017). A two-step approach to account for unobserved spatial heterogeneity. Spatial Economic Analysis, 12(4), 452–471. https://doi.org/10.1080/17421772.2017.1286373
Bogin, A., Doerner, W., & Larson, W. (2019). Local house price dynamics: New indices and stylized facts. Real Estate Economics, 47(2), 365–398. http://doi.org/10.1111/reec.2019.47.issue-2
Bollerslev, T., Chou, R. Y., & Kroner, K. F. (1992). Arch modeling in finance: A review of the theory and empirical evidence. Journal of Econometrics, 52(1), 5–59. https://doi.org/10.1016/0304-4076(92)90064-X
Chang, Z., & Diao, M. (2021). Inter-city transport infrastructure and intra-city housing markets: Estimating the redistribution effect of high-speed rail in Shenzhen, China. Urban Studies, 59(4), 870–889. https://doi.org/10.1177/00420980211017811.
Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica, 50(4), 987–1007. https://doi.org/10.2307/1912773
Engle, R. F., & Bollerslev, T. (1986). Modelling the persistence of conditional variances. Econometric Reviews, 5(1), 1–50. https://doi.org/10.1080/07474938608800095
Gupta, A., & Robinson, P. M. (2015). Inference on higher-order spatial autoregressive models with increasingly many parameters. Journal of Econometrics, 186(1), 19–31. https://doi.org/10.1016/j.jeconom.2014.12.008
Hølleland, S., & Karlsen, H. A. (2020). A stationary spatio-temporal GARCH model. Journal of Time Series Analysis, 41(2), 177–209. https://doi.org/10.1111/jtsa.12498
Holmes, M. J., Otero, J., & Panagiotidis, T. (2017). A pair-wise analysis of intra-city price convergence within the Paris housing market. The Journal of Real Estate Finance and Economics, 54(1), 1–16. https://doi.org/10.1007/s11146-015-9542-z
Jacquier, E., Polson, N. G., & Rossi, P. E. (1994). Bayesian analysis of stochastic volatility models. Journal of Business & Economic Statistics, 12(4), 371–389. https://doi.org/10.1198/073500102753410408
Kelejian, H. H., & Prucha, I. R. (2010). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. Journal of Econometrics, 157(1), 53–67. https://doi.org/10.1016/j.jeconom.2009.10.025
Kim, S., Shephard, N., & Chib, S. (1998). Stochastic volatility: Likelihood inference and comparison with ARCH models. The Review of Economic Studies, 65(3), 361–393. https://doi.org/10.1111/1467-937X.00050
Lee, L.-f. (2004). Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica, 72*(6), 1899–1925. https://doi.org/10.1111/j.1468-0262.2004.00558.x

Lee, L.-f. (2007). GMM and 2SLS estimation of mixed regressive, spatial autoregressive models. *Journal of Econometrics, 137*(2), 489–514. https://doi.org/10.1016/j.jeconom.2005.10.004

Lee, L.-f., & Liu, X. (2010). Efficient GMM estimation of high order spatial autoregressive models with autoregressive disturbances. *Econometric Theory, 26*(1), 187–230. https://doi.org/10.1017/S0266466609090653

Lee, L.-f., & Yu, J. (2010). A spatial dynamic panel data model with both time and individual fixed effects. *Econometric Theory, 26*(2), 564–597. https://doi.org/10.1017/S0266466609100099

Lee, M. P. (2012). *Bayesian statistics: An introduction* (4th ed.). John Wiley and Sons Ltd.

LeSage, J. P., & Pace, R. K. (2009). *Introduction to spatial econometrics (statistics: A series of textbooks and monographs)*. Chapman and Hall/CRC.

Meen, G. (1999). Regional house prices and the ripple effect: A new interpretation. *Housing Studies, 14*(6), 733–753. https://doi.org/10.1080/02673039982524

Merk, M. S., & Otto, P. (2021). Directional spatial autoregressive dependence in the conditional first- and second-order moments. *Spatial Statistics, 41*, 100490. https://doi.org/10.1016/j.spa.2020.100490

Otto, P. (2019). spGARCH: An R-package for spatial and spatiotemporal ARCH and GARCH models. *The R Journal, 11*(2), 401–420. https://doi.org/10.32614/RJ-2019-053

Otto, P., & Schmid, W. (2023). A general framework for spatial GARCH models. *Statistical Papers, 64*, 1721–1747. https://doi.org/10.1007/s00362-022-01357-1

Otto, P., Schmid, W., & Garthoff, R. (2018). Generalised spatial and spatiotemporal autoregressive conditional heteroscedasticity. *Spatial Statistics, 26*, 125–145. https://doi.org/10.1016/j.spasta.2018.07.005

Robinson, P. M. (2009). Large-sample inference on spatial dependence. *Econometrics Journal, 12*, S68–S82. https://doi.org/10.1111/j.1368-423X.2008.00264.x

Sandmann, G., & Koopman, S. J. (1998). Estimation of stochastic volatility models via Monte Carlo maximum likelihood. *Journal of Econometrics, 87*(2), 271–301. https://doi.org/10.1016/S0304-4076(98)00016-5

Sato, T., & Matsuda, Y. (2017). Spatial autoregressive conditional heteroskedasticity models. *Journal of the Japan Statistical Society, 47*(2), 221–236. https://doi.org/10.14490/jjss.47.221

Sato, T., & Matsuda, Y. (2021). Spatial extension of generalized autoregressive conditional heteroskedasticity models. *Spatial Economic Analysis, 16*(2), 148–160. https://doi.org/10.1080/17421772.2020.1742929

Shephard, N. (1994). Partial non-Gaussian state space. *Biometrika, 81*(1), 115–131. https://doi.org/10.1093/biomet/81.1.115

Taşpınar, S., Doğan, O., Chae, J., & Bera, A. K. (2021). Bayesian inference in spatial stochastic volatility models: An application to house price returns in Chicago. *Oxford Bulletin of Economics and Statistics, 83*(5), 1243–1272. https://doi.org/10.1111/obes.12425

Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography, 46*, 234–240. https://doi.org/10.2307/143141

Yu, J., de Jong, R., & Fei Lee, L. (2008). Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both $n$ and $t$ are large. *Journal of Econometrics, 146*(1), 118–134. https://doi.org/10.1016/j.jeconom.2008.08.002

Zhang, L., & Yi, Y. (2017). Quantile house price indices in Beijing. *Regional Science and Urban Economics, 63*, 85–96. https://doi.org/10.1016/j.regsciurbeco.2017.01.002