Heriot-Watt University

Accountancy, Economics, and Finance Working Papers

Working Paper 2024-11

ENGLE-GRANGER REPRESENTATION IN SPATIAL AND SPATIO-TEMPORAL MODELS

Arnab Bhattacharjee, Jan Ditzen, and Sean Holly

September 2024

Keywords: Spatio-temporal dynamics, Error Correction Models, Weak and strong cross sectional dependence, US house prices, Spatial weight matrices, Common Correlated Effects estimator

JEL: C21, C22, C23, R3

Engle-Granger Representation in Spatial and Spatio-Temporal Models

Arnab Bhattacharjee[1], Jan Ditzen[2] and Sean Holly[3] *

June 2024

Abstract

The literature on panel models has made considerable progress in the last few decades, integrating non-stationary data both in the time and spatial domain. However, there remains a gap in the literature that simultaneously models non-stationarity and cointegration in both the time and spatial dimensions. This paper develops Granger representation theorems for spatial and spatio-temporal dynamics. In a panel setting, this provides a way to represent both spatial and temporal equilibria and dynamics as error correction models. This requires potentially two different processes for modelling spatial (or network) dynamics, both of which can be expressed in terms of spatial weights matrices. The first captures strong cross-sectional dependence, so that a spatial difference, suitably defined, is weakly cross-section dependent (granular) but can be non-stationary. The second is a conventional weights matrix that captures short-run spatio-temporal dynamics as stationary and granular processes. In large samples, cross-section averages serve the first purpose and we propose the mean group, common correlated effects estimator together with multiple testing of cross-correlations to provide the short-run spatial weights. We apply this model to house prices in the 375 MSAs of the US. We show that our approach is useful for capturing both weak and strong cross-section dependence, and partial adjustment to two long-run equilibrium relationships in terms of time and space.

JEL Codes: C21, C22, C23, R3.

Keywords: Spatio-temporal dynamics; Error Correction Models; Weak and strong cross sectional dependence; US house prices. Spatial Weight matrices; Common Correlated Effects Estimator.

*Corresponding authors: a.bhattacharjee@hw.ac.uk. [1]: Heriot-Watt University and National Institute of Economic & Social Research, UK. [2]: Free University of Bozen-Bolzano, Italy. [3]: Faculty of Economics, University of Cambridge, UK.
We thank participants at the "*Conference to Celebrate Prof. Hashem Pesaran's Achievements*" at the University of Southern California, USA and particularly Badi Baltagi, Anil Bera, Cheng Hsiao, George Kapetanios, Jeff Nugent and Hashem Pesaran for helpful comments and suggestions that helped shape the paper substantially. The usual disclaimer applies.

1 Introduction

The first introduction to economics of cointegration and its link to the error correction models of Phillips (1954, 1957) and Sargan (1964) in the time domain, was Granger (1981) and Granger and Weiss (1983) and then its formalisation and some testing procedures provided in Engle and Granger (1987). In the past, the most common empirical application focusing on spatial adjustments was in agricultural economics and the equilibrating processes (and the law of one price) in different markets for products. One of the earliest contributions to use the error correction approach was Ravallion (1986). He explored market integration of rice prices in Bangladesh. His approach extracted information on the nature of spatial price differentials. In particular, the dynamic approach in an ECM framework makes a clear distinction between short-run market integration, and integration as a long-run tendency, in the short-run adjustment processes. This was followed by a substantial literature on dynamic adjustment in agricultural markets; see, for example, Goodwin and Schroeder (1991), Gordon et al. (2004), Muwanga and Snyder (1997) and Asche et al. (2004). A more recent contribution is Kumar and Karak (2022) who used an error correction model to examine horizontal and vertical integration of wholesale and retail prices of wheat in the major markets of India. On confirming cointegration between the wholesale and retail prices of wheat, a vector error correction model was used to determine the speed of adjustment of wheat prices. The results revealed that price signals are transmitted across regions, indicating that price changes in one market are consistently related to price changes in other markets.

In another recent paper, von Cramon-Taubadel and Goodwin (2021) review recent developments in the analysis of price transmission in agricultural markets. Separated in time, form, and space (as well as in combinations of such factors) agricultural markets face transactions and storage costs as well as production and marketing costs. They argue that much of recent research on spatial market linkages has reflected methodological advances that have led to increasingly nonlinear time-series models. Advances in the theoretical and empirical literature over the last few decades have established that price relationships in the food chain are highly context specific. Furthermore, improvements in marketing, information, and transportation technologies have strengthened the links between prices in the food system.

The above literature provides a clear sense of cross-section dynamics in markets in a way that emphasises equilibrating forces of arbitrage at the spatial level. However, nonstationarity and error correction in the literature has focused almost exclusively on the temporal dimension. Likewise, the related literature on panel data cointegration and error correction models has considered solely time series nonstationary dynamics, but applied to many other contexts beyond agricultural economics; see, for example, Pesaran et al. (1999), Fingleton (1999, 2009) and Beenstock et al. (2012). This marks a clear departure of the current paper from the literature, in asking how nonstationarity (or nongranularity) and equilibrium can be understood in a unified spatial context and how such spatial dynamics can be modeled. Our research builds upon Müller and Watson

(2023) who represent a distinct view in clarifying the notion of spatial unit roots. Our contribution takes this approach further by developing spatial and spatio-temporal Granger representation theorems leading to error correction models integrating both the spatial and temporal dimensions.

The rest of the paper is organised as follows. Section 2 provides intuitive and applied context for spatial and spatio-temporal cointegration. Following this, Section 3 develops Granger representation theorems – first for the spatial domain and next for spatio-temporal settings. Section 4 provides an application of the proposed concepts and framework to understand nonstationary spatio-temporal dynamics in housing markets across MSAs in the USA and over time. Section 5 concludes.

2 Context: Conceptual empirical illustrations

Müller and Watson (2023) provided a framework to define and study spatial unit roots. In this short section, we conceptually and intuitively extend this framework using a couple of empirical application contexts. The applications themselves are hypothetical yet based upon previous literature, and the objective is to motivate what we mean by spatial (and spatio-temporal) cointegration and how error correction models may be useful for modeling spatial dynamics using cross-section and panel data.

2.1 Urban housing markets: Spatial cointegration

Our first example relates to a simple illustration of spatial nongranularity and cointegration in an urban housing market. There is remarkable similarity in residential sorting across many cities the world over (but particularly in Europe). The empirical regularity is that western neighbourhoods of cities are more affluent. A key explanation is that *"bad winds blow from the west"* (Meen, 2016; Heblich et al., 2021): predominantly winds come from the west and blow stench eastwards resulting in residential sorting. Thus, we expect one cluster of higher prices and larger houses in the west and another of lower sizes and prices in the east. This would likely result in co-trending of household incomes and better and more expensive housing.

This situation is conceptually very similar to temporal nonstationarity. One would therefore expect nongranularity (spatial strong dependence) reflected in a strong global relationship between prices and incomes. In addition, if there were a stable spatial equilibrium between the two, any local variation away from this relationship would generate a partial adjustment moving prices in that locality back towards equilibrium.

One can easily extend this argument to more clusters or multicentric cities. Also, this partial adjustment is intrinsically local rather than global (across the entire city). Then, spatial and temporal dynamics can be coincident and potentially lead to spatio-temporal ECM along the lines of Bhattacharjee et al.

(2022). Our developments in this paper formalises this argument in both the spatial and spatio-temporal contexts.

Beyond the urban scale, similar arguments may hold in larger (regional) scales as well. This could, for example, partly explain the phenomenon of "ripple effects" in the UK housing markets (Drake, 1995; Cook and Holly, 2000; Holly et al., 2011; Meen, 2016), whereby price (and income) shocks often arise in London and the South East of England and then spread out to other regions. Similar evidence of price gradients have also been observed elsewhere, not least across US regions (Chiang and Tsai, 2016). This partly motivates our application (in Section 4) to US metropolitan housing markets.

2.2 Firm panel data: Spatio-temporal Error Correction

To illustrate the potential for spatio-temporal cointegration and error correction, we follow Bhattacharjee et al. (2014) and consider the firm-level relationships between costs and sales. If there were an equilibrium profit margin for each firm, this would be reflected in a long-run (temporal) relationship between logarithms of cost and sales, and partial adjustment each period towards this equilibrium. This would imply a conventional and familiar temporal error correction model along the lines of Granger (1981), Granger (1983) and (Engle and Granger, 1987).

Now, consider in addition the potential for an equilibrium cross-firm relationship between costs and sales. This implies that, in equilibrium, costs and sales of each firm would move in line with each other in a way that their relative magnitudes remain in balance. Conceptually, this implies network (spatial) equilibrium market shares (Bhattacharjee et al., 2014). If such an equilibrium exists, costs and sales of each firm will remain tightly clustered around firm-specific proportions to total costs and sales within the sector. Whether such a network equilibrium exists or not is an empirical question, and our spatial and spatio-temporal Granger representation theorems precisely establishes conditions under which this happens.

If indeed there is such an equilibrium, analogous to temporal cointegration, it is possible that there is network (spatial) partial adjustment to this equilibrium. Once again, our theoretical work in the following section establishes precisely the conditions under which this may happen and the precise form of the resulting error correction model.

3 Spatial and Spatio-Temporal Engle-Granger Representation Theorems

First, we present a spatial Granger representation theorem and associated error correction model. This is followed by discussion of corresponding measures of spatial weights. Finally, we develop a spatio-temporal Granger representation theorem and discuss short run spatial dynamics.

3.1 Spatial Granger Representation and Error Correction Model

As a reminder, consider first the conventional time series error correction model as developed by Granger and Weiss (1983), Engle and Granger (1987) and Johansen (1995), among others.

Theorem 1 (Granger Representation Theorem (Granger and Weiss, 1983)). *Consider two time series variables $\{y_t\}$ and $\{x_t\}$ indexed by discrete time points $t = 0, 1, 2, \dots$ jointly following a $VAR(1)$ model and admitting only one unit root. Then the two time series can be represented by an error correction model*

$$\begin{bmatrix} \Delta y_s \\ \Delta x_s \end{bmatrix} = (1-\rho) \begin{bmatrix} b \\ d \end{bmatrix} [c y_{t-1} - a x_{t-1}] + \eta_t,$$

where Δ is the first differencing operator.

Now, consider the bivariate spatial $VAR(1)$ model:

$$\begin{bmatrix} y_s \\ x_s \end{bmatrix} = \mathbf{\Phi} \begin{bmatrix} \mathbf{W} y_s \\ \mathbf{W} x_s \end{bmatrix} + \epsilon_s, \tag{1}$$

where $s \in S$ is a location, $\mathbf{\Phi}$ a (2×2) matrix representing spatial dynamics, and ϵ_s is an error white noise distributed over the spatial domain S. $\mathbf{W}$ is a spatial weights (interaction) matrix operator on $S \times S$ with zero diagonal elements. Hence $\mathbf{W} y_s$ and $\mathbf{W} x_s$ are spatial lags of y and x at location s respectively, capturing the average of neighbouring values where neighbourhood identity is represented by $\mathbf{W}$. Correspondingly, $\begin{bmatrix} \underline{\Delta} y_s \\ \underline{\Delta} x_s \end{bmatrix}$ denotes the spatial first difference $\begin{bmatrix} y_s - \mathbf{W} y_s \\ x_s - \mathbf{W} x_s \end{bmatrix}$.

Equation (1) is akin to the conventional and popular spatial autoregressive (or spatial lag) model (Anselin, 1988; Anselin et al., 1996; Baltagi et al., 1996) but with one key difference. Unlike traditional spatial econometric models that allow spatial stationary dynamics, here we do not assume the spatial granularity condition (Pesaran, 2006). Thereby, we enable $\mathbf{W}$ to model nonstationary (nongranular) or strong dependent spatial dynamics. This has important implications for specification of $\mathbf{W}$ which we discuss in the following Section 3.2.

Corollary 1.1 (Spatial Granger Representation Theorem). *If the spatial procesess y and x jointly have a single unit root, then they can be represented by an error correction model*

$$\begin{bmatrix} \underline{\Delta} y_s \\ \underline{\Delta} x_s \end{bmatrix} = (1-\lambda) \begin{bmatrix} \beta \\ \delta \end{bmatrix} [\gamma \mathbf{W} y_s - \alpha \, \boldsymbol{W} x_s] + \epsilon_s.$$

Proof. The line of proof follows Granger and Weiss (1983) and Engle and Granger (1987) closely. The characteristic equation for $\mathbf{\Phi}$ is:

$$|\mathbf{\Phi} - \lambda \mathbf{I}| = 0.$$

If both roots are unity, we have spurious regression. On the other hand, if both roots are less than 1 in absolute value, we have spatial stationary (cross section granular) processes (Müller and Watson, 2023). Our domain is intermediate with one unit root:

$$z_1 = 1, \quad z_2 = \lambda, \quad \lambda \neq 0, |\lambda| < 1.$$

Since $z_1 = 1$, both y and x are spatially $I(1)$ *a la* Müller and Watson (2023).

Since $\lambda \neq 0$, $\mathbf{\Phi}$ has full rank. Hence, we can write its SVD:

$$\mathbf{\Phi} = \mathbf{P} \begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix} \mathbf{Q},$$

where the matrix $\mathbf{P} = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$ holds the eigenvectors as columns, with $|\mathbf{P}| = 1$ without loss of generality, and $\mathbf{Q} = \mathbf{P}^{-1}$. Then, since

$$\alpha\delta - \beta\gamma = 1,$$

$$\mathbf{Q} = \mathbf{P}^{-1} = \begin{bmatrix} \delta & -\beta \\ -\gamma & \alpha \end{bmatrix}.$$

Putting elements together:

$$\mathbf{\Phi} = \begin{bmatrix} \alpha\delta - \lambda\beta\gamma & -\alpha\beta(1-\lambda) \\ \gamma\delta(1-\lambda) & -\gamma\beta + \lambda\alpha\delta \end{bmatrix}.$$

Simplifying the top left element (adding and subtracting $\gamma\beta$):

$$\begin{aligned} \alpha\delta - \lambda\beta\gamma = & \quad \alpha\delta - \gamma\beta + \gamma\beta - \lambda\beta\gamma \\ = & \quad 1 + \gamma\beta(1-\lambda). \end{aligned}$$

Likewise, adding and subtracting $\alpha\delta$ to the bottom right element

$$-\gamma\beta + \lambda\alpha\delta = 1 - \alpha\delta(1-\lambda).$$

Then,

$$\mathbf{\Phi} = \mathbf{I} + (1-\lambda) \begin{bmatrix} \beta \\ \delta \end{bmatrix} \begin{bmatrix} \gamma & -\alpha \end{bmatrix}.$$

Substituting into Eqn. (1):

$$\begin{bmatrix} \underline{\Delta} y_s \\ \underline{\Delta} x_s \end{bmatrix} = (1-\lambda) \begin{bmatrix} \beta \\ \delta \end{bmatrix} [\gamma \mathbf{W} y_s - \alpha \mathbf{W} x_s] + \epsilon_s,$$

where $\underline{\Delta}(.)_s = (\mathbf{I} - \mathbf{W})(.)_s$ and $[\gamma \mathbf{W} y_s - \alpha \mathbf{W} x_s]$ is a measure of the departure from equilibrium in the neighbourhood of location s. □

This is the spatial error correction model in the tradition of Granger and Weiss (1983) and Engle and Granger (1987):

$$\begin{bmatrix} \underline{\Delta} y_s \\ \underline{\Delta} x_s \end{bmatrix} = \begin{bmatrix} (1-\lambda)\beta \\ (1-\lambda)\delta \end{bmatrix} \begin{bmatrix} \gamma & -\alpha \end{bmatrix} \begin{bmatrix} \mathbf{W} y_s \\ \mathbf{W} x_s \end{bmatrix} + \epsilon_s. \tag{2}$$

Focussing on a single equation, we have:

$$\underline{\Delta} y_s = (1-\lambda)\beta \left[\gamma \mathbf{W} y_s - \alpha \mathbf{W} x_s\right] + \epsilon_{y,s}.$$

The above derivation follows closely that of Engle and Granger (1987). Here we considered a bivariate spatial process $\begin{bmatrix} y_s \\ x_s \end{bmatrix}$ which created the potential for at most one cointegrating vector corresponding to a single unit root. Extending to a higher dimensional vector of spatial processes, one can follow Johansen (1995) and consider potential for multiple cointegrating vectors. Conceptually, this constitutes a simple extension of analogous time series concepts, based on the above construction of spatial lags and differences. Then, one can also use this formulation and the framework of Johansen (1995) to test for cointegration and cointegrating rank. This would require aditional technical treatment which is retained for future research.

3.2 Choice of spatial weights

So far, we have left $\mathbf{W}$ unspecified. Next, we consider the choice of spatial weights. While short run dynamics will require a granular weights matrix along the lines of Anselin (1988), among others, modeling spatial long run equilibrium will require spatial weights with more specific structure, such as the common correlated effects of Pesaran (2006). First we consider spatial long-run equilibrium followed by short run stationary (granular) dynamics.

3.3 Equilibrating long-run W

The choice of $\mathbf{W}$ has to do with the nature of the spatial patterns in equilibrium. Looking towards a time series setting for inspiration, it is common to write an $AR(1)$ model:

$$\Delta y_t = \beta + \alpha \Delta y_{t-1} + \epsilon_t.$$

Then, setting $\Delta y_{t+1} = \Delta y_t = y^\star$ in equilibrium, we have

$$y^\star = \frac{\beta}{1-\alpha}.$$

Additional AR or DL terms lead to minor modifications, not any deep conceptual issues. The spatial context is similar to the above time series case:

$$\begin{aligned} \underline{\Delta} y_s &= \beta + \alpha W \underline{\Delta} y_s + \epsilon_s \\ (y_s - W y_s) &= \beta + \alpha W (y_s - W y_s) + \epsilon_s. \end{aligned}$$

However, there are several key points and distinctions to note. First, given the spatial weights matrix $\mathbf{W}$, $\mathbf{W}y_s$ (likewise $\mathbf{W}x_s$) has the interpretation of a spatial lag, meaning an average of y (respectively, x) values for the neighbouring locations. Importantly, this leads into a key distinction from the time series case.

Specifically, and second, a shock to y_s (analogously x_s) diffuses to neighbours represented by the weights matrix $\mathbf{W}$ and then to neighbours of neighbours, captured by $\mathbf{W}^2$, and so on. Thus, while in the time series case, information (effect of shocks) flow from the past to the present to the future, the natural way to structurally conceptualise the spatial case is exactly the opposite: from y to its spatial lag $\mathbf{W}y$ and further to its second order spatial lag $\mathbf{W}^2y$, and so on. Hence, the natural way to seek spatial equilibrium is to go wider into the network structure – $W^2, W^3, \dots$ and so on – that is to include locations in higher degrees of separation. This is in line with increasing domain asymptotics.

Third and finally, analogous to the temporal case, adding an additional DL term ($\underline{\Delta}x = x - \mathbf{W}x$ and its spatial lag) is straightforward. However, unlike time series, adding higher order spatial lags beyond spatial $ARDL(1,1)$ is not very instructive. In fact, the term "higher order spatial AR models" means something very different (Gupta and Robinson, 2015, 2018).

Then, following the reverse information flow discussed above and natural in this spatial context, we can set $\mathbf{W}^{d+1}(y_s - \mathbf{W}y_s) \approx \mathbf{W}^d(y_s - \mathbf{W}y_s) = \underline{\Delta}y^\star$. Then, in spatial equilibrium, we have:

$$y^\star = \lim_{d\to\infty} \mathbf{W}^d(y_s - \mathbf{W}y_s) = \frac{\beta}{1-\alpha},$$

analogous to the time series case apart from the above distinction. It is also natural to expect that $\lim_{d\uparrow\infty} W^d$ covers every location in the entire spatial domain. This is barring the possibility of disjoint markets, or several different equilibria in different regions, which can emerge in the case of multi-centric cities or disjoint social networks.

Note further that, the above setting allows spatial variation in the offset of y_s at location s from its neighbours $\mathbf{W}y_s$. For example, the metropolitan area of Manhattan can have a different premium relative to its neighbourhood, as compared with downtown Miami, for example. This premium can potentially converge to a stable location-specific value in the long run. But this is a question for temporal equilibrium, to which we return later in a panel context.

The standard connected case implies that $\mathbf{W}$ is such that every pair of locations is connected, potentially across multiple degrees of separation. This would hold automatically if $\mathbf{W}$ represents a weighted average of all locations, which is the central common correlated effects case conventionally used to model strong dependence (Pesaran, 2006). In this common correlated effects case,

$$\mathbf{W} = \frac{1}{n}1\,1' \approx \frac{1}{n-1}\left[1\,1' - I\right].$$

Clearly, the above social network type weights matrix $\mathbf{W}$ induces a connected network, but is also idempotent, such that suitably normalised $\mathbf{W} = \mathbf{W}^2 = \mathbf{W}^3 = \dots$.

3.3.1 Spatial short run dynamics

One can add stationary higher order $\underline{\Delta}$ terms in y and x to ensure white noise errors ϵ_s. However, while including a single spatial-differenced term is natural, as discussed above, there is no apparent intuitive interpretation of higher order DL terms. In particular, the ARDL(1,0) specification is common in time series which, translated into the spatial context, implies adding to Equation (1) one additional spatial distributed lag term:

$$SDL(0): \quad \begin{bmatrix} 0 & \theta_1 \\ \theta_2 & 0 \end{bmatrix} \begin{bmatrix} y_s \\ x_s \end{bmatrix}.$$

Then, we obtain the familiar ECM form which can be written in this spatial context as:

$$\underline{\Delta} y_s = \beta_0 + \beta_1 \underline{\Delta} x_s + (1-\lambda)\beta\gamma \left[\mathbf{W} y_s - \frac{\alpha}{(1-\lambda)\beta\gamma} \mathbf{W} x_s \right] + \epsilon_{y,s},$$

where β_1 represents the short-run dynamic adjustment and $-(1-\lambda)\beta\gamma$ the partial adjustment to the spatial equilibrium relationship $\left[y_s - \frac{\alpha}{(1-\lambda)\beta\gamma} x_s \right]$.

Note that, the above spatial DL term is additional to the core ECM in Corollary 1.1 (and similarly Theorem 2). Hence, while there is a well-defined common correlated effects choice for the long run equilibrating spatial weights matrix $\mathbf{W}$, there is more freedom in the choice of spatial weights for the short run dynamics. We return to a discussion of this choice in Section 3.4 in the context of the spatio-temporal error correction model.

3.4 Spatio-Temporal Granger Representation and Large Panels

Now, let us consider panel data:

$$\begin{aligned} z_{it} &= \begin{pmatrix} y_{it} & x_{it} \end{pmatrix}' \quad i = 1, \ldots, n; \quad t = 1, \ldots T \\ \mathbf{z}_t &= \begin{pmatrix} \mathbf{y}_t & \mathbf{x}_t \end{pmatrix}'_{(n \times 2)}. \end{aligned}$$

Consider the spatio-temporal stochastic process with first order autoregressive dynamics in both dimensions:

$$\begin{aligned} z_{it} &= \Phi \mathbf{w}_i \mathbf{z}_t + \epsilon_{it} \\ z_{i,t-1} &= \Phi \mathbf{w}_i \mathbf{z}_{t-1} + \epsilon_{i,t-1} \\ z_{it} &= \Omega z_{i,t-1} + \eta_{it} \\ \mathbf{w}_i \mathbf{z}_t &= \Omega \mathbf{w}_i \mathbf{z}_{t-1} + \mathbf{w}_i \eta_t, \end{aligned}$$

where $\eta_t = \begin{pmatrix} \eta_{1t} & \eta_{2t} & \ldots & \eta_{nt} \end{pmatrix}'$, $\mathbf{w}_i$ is the i-th row of a $n \times n$ spatial weights matrix W, and the errors ϵ_{it} and η_{it} are idiosyncratic and independent white noise processes.

Theorem 2 (Spatio-Temporal Granger Representation Theorem)**.** *Assume that* $\mathbf{z}$ *is cointegrated along both the spatial dimension (as discussed before) and the conventional temporal dimension. Then, the twice differenced process*

$$\Delta\underline{\Delta}\mathbf{z} \equiv \underline{\Delta}\Delta\mathbf{z}$$

admits an error correction representation with: (a) potential partial adjustment to two equilibrium relationships – a spatial equilibrium and a temporal equilibrium; and (b) potential strong dependence modeled by common correlated effects.

Proof. Spatial cointegration implies that $\mathbf{\Phi}$ has one (unit) root lying on the unit circle and another root within the circle. Then, as shown above:

$$\mathbf{\Phi} = \mathbf{I} + (1-\lambda)\begin{bmatrix}\beta \\ \delta\end{bmatrix}\begin{bmatrix}\gamma & -\alpha\end{bmatrix}.$$

Likewise, following Granger and Weiss (1983) and Engle and Granger (1987):

$$\mathbf{\Omega} = \mathbf{I} + (1-\rho)\begin{bmatrix}b \\ d\end{bmatrix}\begin{bmatrix}c & -a\end{bmatrix}.$$

Then,

$$\begin{aligned}
\underline{\Delta}\mathbf{z}_{it} =& (\mathbf{\Phi}-\mathbf{I})\mathbf{w}_i\mathbf{z}_t + \epsilon_{it} \\
\Delta\underline{\Delta}\mathbf{z}_{it} =& (\mathbf{\Phi}-\mathbf{I})\Delta\mathbf{w}_i\mathbf{z}_t + \Delta\epsilon_{it} \\
=& (\mathbf{\Phi}-\mathbf{I})\left[(\mathbf{\Omega}-\mathbf{I})\mathbf{w}_i\mathbf{z}_{t-1} + \mathbf{w}_i\eta_t\right] + \Delta\epsilon_{it} \\
=& (\mathbf{\Phi}-\mathbf{I})(\mathbf{\Omega}-\mathbf{I})\mathbf{w}_i\mathbf{z}_{t-1} + \left[(\mathbf{\Phi}-\mathbf{I})\mathbf{w}_i\eta_t + \Delta\epsilon_{it}\right] \\
=& \left[(\mathbf{\Phi}-\mathbf{I})\mathbf{\Omega}\mathbf{w}_i\mathbf{z}_{t-1} + (\mathbf{\Phi}-\mathbf{I})\mathbf{w}_i\eta_t\right] \\
& + \left[(\mathbf{\Omega}-\mathbf{I})\mathbf{\Phi}\mathbf{w}_i\mathbf{z}_{t-1} + (\mathbf{\Omega}-\mathbf{I})\epsilon_{i,t-1}\right] \\
& + (\mathbf{\Omega}\mathbf{\Phi}-\mathbf{I})\mathbf{w}_i\mathbf{z}_{t-1} \\
& + \left[(\mathbf{\Phi}-\mathbf{I})\mathbf{w}_i\eta_t + \Delta\epsilon_{it} - (\mathbf{\Phi}-\mathbf{I})\mathbf{w}_i\eta_t - (\mathbf{\Omega}-\mathbf{I})\epsilon_{i,t-1}\right] \\
=& (\mathbf{\Phi}-\mathbf{I})\mathbf{w}_i\mathbf{z}_t + (\mathbf{\Omega}-\mathbf{I})z_{i,t-1} + (\mathbf{\Omega}\mathbf{\Phi}-\mathbf{I})\mathbf{w}_i\mathbf{z}_{t-1} \\
& + \left[\epsilon_{it} - \mathbf{\Omega}\epsilon_{i,t-1}\right] \\
=& (1-\lambda)\begin{bmatrix}\beta \\ \delta\end{bmatrix}\begin{bmatrix}\gamma & -\alpha\end{bmatrix}\begin{bmatrix}\mathbf{w}_i\mathbf{y}_t \\ \mathbf{w}_i\mathbf{x}_t\end{bmatrix} + (1-\rho)\begin{bmatrix}b \\ d\end{bmatrix}\begin{bmatrix}c & -a\end{bmatrix}\begin{bmatrix}y_{i,t-1} \\ x_{i,t-1}\end{bmatrix} \\
& + (\mathbf{\Omega}\mathbf{\Phi}-\mathbf{I})\mathbf{w}_i\mathbf{z}_{t-1} + \left[\epsilon_{it} - \mathbf{\Omega}\epsilon_{i,t-1}\right],
\end{aligned}$$

where the first term represents partial adjustment to the spatial equilibrium, the second is the conventional Engle-Granger partial adjustment to the temporal equilibrium and the third term has spatial strong dependence interpretation; see also Bhattacharjee et al. (2022). □

Two important implications follow. First, in large panels the third term $(\mathbf{\Omega}\mathbf{\Phi}-\mathbf{I})\mathbf{w}_i\mathbf{z}_{t-1}$ has a spatial strong dependence interpretation (Bhattacharjee et al., 2022). Hence, a natural choice for the long-run weights matrix $\mathbf{W}$

is a cross-section weighted average. Its validity and adequacy is underlined by Pesaran (2006), for example, with the popular choice being the common correlated effects. Further, this term also has another interpretation as capturing the interaction between space and time dynamics. To see this, consider again spatial cointegration in the previous section. Together unit root and orthonormalisation placed two constraints on the elements of $\mathbf{\Phi}$, leaving two free parameters which are then linked to the long run spatial equilibrium relationship and partial adjustment to it. Likewise, the four elements of $\mathbf{\Omega}$ are fixed by one unit root, orthonormality, long run temporal equilibrium and partial adjustment to this equilibrium (Granger and Weiss, 1983). Our model puts the two together, so here we have elements of a (4×4) matrix $\mathbf{\Phi} \otimes \pi\mathbf{\Omega}$, where π measures the interaction contribution of temporal dynamics relative to spatial dynamics (normalised to unity). This parameter π leads to the third term additional to one-dimensional space or time dynamics. Hence, the third term can also be interpreted as measuring the strength of space-time interactions.

Second, in line with the time series literature, additional distributed lag terms in both x and y can be added to enrich short run dynamics and aid modelling and interpretation. This provides additional flexibility in modelling short run dynamics. Effectively, the weights matrix has two roles here: (a) to model the long run spatial equilibrium $\begin{bmatrix}\gamma & -\alpha\end{bmatrix}\begin{bmatrix}\mathbf{w}_i\mathbf{y}_t \\ \mathbf{w}_i\mathbf{x}_t\end{bmatrix}$; and (b) to model short run dynamics as in $\Delta\underline{\Delta}\mathbf{y}_{it}$ and $\Delta\underline{\Delta}\mathbf{x}_{it}$. While, as discussed above, the first role lends itself naturally to the choice of cross-section averages, the weights matrix representing the short-run dynamics in additional distributed lag terms can, and in principle, should be different from the above long-run $\mathbf{W}$. This is because it is used to model the stationary dynamics, and its specification can be based on theory, alternate heuristics or even spatial weights estimated from the data.

To see this more clearly, consider a potentially alternate choice $\mathbf{W_0}$. Then,

$$\begin{aligned}\Delta\underline{\Delta}\mathbf{y}_{it} =& \Delta y_{it} - \mathbf{w}_i\Delta y_t \\ =& \left[\Delta y_{it} - \mathbf{w}_{0i}\Delta y_t\right] - \mathbf{w}_i\Delta y_t + \mathbf{w}_{0i}\Delta y_t\end{aligned}$$

The first term is the double difference based on the alternate short run dynamics weights matrix $\mathbf{W_0}$, the second term is encompassed in the above common correlated effects term (plus error), and hence the only new component is the third term. However, note that because of partial adjustment to the spatial equilibrium,

$$\mathbf{w}_{0i}\Delta y_t = \frac{\alpha}{\gamma}\mathbf{w}_{0i}\Delta x_t + \text{error terms.}$$

Hence, all this requires is adding a spatial distributed lag in x on the right hand side. This implies a free choice for $\mathbf{W_0}$. For example, along the lines of Bhattacharjee et al. (2022), the cross-section correlation weights (Bailey et al., 2016) can be used here.

This suggests an empirical model of the following form:

$$\begin{aligned}\Delta\underline{\Delta}y_{i,t} =& \beta_{i,0} + \beta_{i,1}\Delta\underline{\Delta}x_{i,t} - \phi_i\left(y_{i,t-1} - \kappa_i x_{i,t-1}\right) \\ &- \lambda_i\left(\boldsymbol{w}_i\boldsymbol{y}_t - \gamma_i\boldsymbol{w}_i\boldsymbol{x}_t\right) + \psi_{i,1}\boldsymbol{w}_i\boldsymbol{y}_{t-1} + \psi_{i,2}\boldsymbol{w}_i\boldsymbol{x}_{t-1} + e_{i,t},\end{aligned} \quad (3)$$

where $\boldsymbol{w}_i$ denotes the i-th row of the common correlated effects or some other strong dependence weights matrix $\boldsymbol{W}$ representing the factor structure, while $\begin{bmatrix}\underline{\Delta}y_{it}\\ \underline{\Delta}x_{it}\end{bmatrix} = \begin{bmatrix}y_{it} - \boldsymbol{w}_{0i}\boldsymbol{y}_t\\ x_{it} - \boldsymbol{w}_{0i}\boldsymbol{x}_t\end{bmatrix}$ with $\boldsymbol{w}_{0i}$ denoting the i-th row of a weak dependence $\boldsymbol{W}_0$ that can be chosen freely, for example from the first stage where $\hat{\boldsymbol{W}}_0$ is obtained from multiple testing of residual cross-correlation. An alternative would be to obtain it from geographical spatial weight matrices such as those based on neighbours (contiguity) or distances.

To equation (3), we add and substract $\lambda_i\left(\boldsymbol{w}_i\boldsymbol{y}_{t-1} - \gamma_i\boldsymbol{w}_i\boldsymbol{x}_{t-1}\right)$ and $\phi_i\left(\boldsymbol{w}_i\boldsymbol{y}_{t-1} - \kappa_i\boldsymbol{w}_i\boldsymbol{x}_{t-1}\right)$. Since both terms can be expressed as linear combinations of $\boldsymbol{w}_i\boldsymbol{y}_{t-1}$ and $\boldsymbol{w}_i\boldsymbol{x}_{t-1}$, we obtain the equivalent one-way differenced ECM similar to Bhattacharjee et al. (2022):

$$\begin{aligned}\Delta\underline{\Delta}y_{i,t} = \beta_{i,0} + \beta_{i,1}\Delta\underline{\Delta}x_{i,t} - \phi_i\left\{\left(y_{i,t-1} - \boldsymbol{w}_i\boldsymbol{y}_{t-1}\right) - \kappa_i\left(x_{i,t-1} - \boldsymbol{w}_i\boldsymbol{x}_{t-1}\right)\right\}\\ - \lambda_i\left(\boldsymbol{w}_i\Delta\boldsymbol{y}_t - \gamma_i\boldsymbol{w}_i\Delta\boldsymbol{x}_t\right) + \psi^{\star}_{i,1}\boldsymbol{w}_i\boldsymbol{y}_{t-1} + \psi^{\star}_{i,2}\boldsymbol{w}_i\boldsymbol{x}_{t-1} + e^{\star}_{i,t}.\end{aligned} \tag{4}$$

Note that, equations (3) and (4) are not different models, but alternate representations of the same spatio-temporal ECM. Obviously the two equations have different error terms and their stochastic behaviour will be different. Which of these is more useful for estimation and inference depends on weak/strong dependence and stationarity/nonstationarity of the error terms, and is therefore largely an empirical question. This is one of the issues on which we place special emphasis in our empirical application.

For the remainder we differentiate between short and long run spatial dynamics. The spatial weights are defined as $\boldsymbol{w}_{i,k}y_{i,t}$, where $k = [S, L]$ with S representing short run and L long run spatial weights. Hence, the spatial lag for variable $y_{i,t}$ is then $\boldsymbol{w}_{i,k}y_{i,t}$ while the spatial first difference is $\underline{\Delta}_k y_{i,t} = y_{i,t} - \boldsymbol{w}_{i,k}y_{i,t}$, where $k = [S, L]$.

In terms of spatial weight matrices we consider four options. First, (a) $\hat{\boldsymbol{w}}_i$ are the cross-correlations obtained from the multiple testing and will be discussed in the next section. Further, we define (b) $\boldsymbol{w}_{i,CSA} = \boldsymbol{w}_{CSA} = 1/N(\mathbf{1}'\mathbf{1} - I_N)$ as the cross-sectional averages, with $\mathbf{1}$ a Nx1 vector of ones. Note, $\boldsymbol{w}_{CSA}\boldsymbol{y}_t \approx \frac{1}{N}\sum_{j=1}^{N} y_{j,t}$. Finally, two geographical weight matrices are considered: a (c) contiguity ($\boldsymbol{w}_{i,c}; \underline{\Delta}_c$) and a (d) distance ($\boldsymbol{w}_{i,d}; \underline{\Delta}_d$) based weighting matrix.

Then, Equations (3) and (4) can be specified as:

$$\begin{aligned}\Delta\underline{\Delta}_S y_{i,t} =& \beta_{i,0} + \beta_{i,1}\Delta\underline{\Delta}_S x_{i,t}\\ &- \phi_i\left(y_{i,t-1} - \kappa_i x_{i,t-1}\right) - \lambda_i\left(\boldsymbol{w}_{i,L}\boldsymbol{y}_t - \gamma_i\boldsymbol{w}_{i,L}\boldsymbol{x}_t\right)\\ &+ \psi_{i,1}\boldsymbol{w}_{i,S}\boldsymbol{y}_{t-1} + \psi_{i,2}\boldsymbol{w}_{i,S}\boldsymbol{x}_{t-1} + e_{i,t}\end{aligned} \tag{5}$$

$$\begin{aligned}\Delta\underline{\Delta}_S y_{i,t} =& \beta_{i,0} + \beta_{i,1}\Delta\underline{\Delta}_S x_{i,t}\\ &- \phi_i\left(\underline{\Delta}_L y_{i,t-1} - \kappa_i\underline{\Delta}_L x_{i,t-1}\right) - \lambda_i\left(\boldsymbol{w}_{i,L}\Delta\boldsymbol{y}_t - \gamma_i\boldsymbol{w}_{i,L}\Delta\boldsymbol{x}_t\right)\\ &+ \psi^{\star}_{i,1}\boldsymbol{w}_{i,S}\boldsymbol{y}_{t-1} + \psi^{\star}_{i,2}\boldsymbol{w}_{i,S}\boldsymbol{x}_{t-1} + e^{\star}_{i,t}.\end{aligned} \tag{6}$$

4 An application – US MSA level house prices

Like regions in the UK and elsewhere around the globe, there is also evidence of "ripple effects" in house prices across US regions (Chiang and Tsai, 2016). We apply our approach to the modelling of US house prices[1] at the level of the Metropolitan Statistical Area (MSA).[2] The sample period is quarterly data from 1975q1 to 2021q4. We use data on 375 MSAs, excluding three that are located in Alaska and Hawaii.

For the error correction model, we use house prices deflated by the consumer price index at the MSA level, dependent on real per capita personal income at the MSA level. We use a version of the panel dataset employed by Bailey et al. (2016) and Aquaro et al. (2021), but extended up to 2021q4. This was further augmented with population and per capita real income data by Yang (2021). In case of spatial weights for the short term dynamics obtained from cross-correlation $\hat{\boldsymbol{w}}_i$ we carry out estimation using the following steps:

1. Estimate a cross-section averages augmented ECM panel model to obtain the cross-correlations:

$$\begin{aligned}\Delta y_{i,t} =& \beta_{i,0} + \beta_{i,1} y_{i,t-1} + \beta_{i,2} \Delta x_{i,t} + \beta_{i,3} x_{i,t-1} \\ &+ \sum_{l=0}^{p_x} \gamma_{x,i,l} \bar{x}_{t-l} + \sum_{l=0}^{p_y} \gamma_{y,i,l} \bar{y}_{t-l} + \epsilon_{i,t}\end{aligned}$$

2. Obtain the cross-correlation matrix from the residuals $\rho_{i,j} = \frac{1}{N} \sum_{t=1}^{T} \hat{\epsilon}_{i,t} \hat{\epsilon}_{j,t}$:

$$\tilde{\boldsymbol{W}} = \begin{pmatrix} \hat{\rho}_{1,1} & \hat{\rho}_{1,2} & \cdots & \hat{\rho}_{1,N} \\ \hat{\rho}_{2,1} & \hat{\rho}_{2,2} & \cdots & \hat{\rho}_{2,N} \\ \vdots & & \ddots & \vdots \\ \hat{\rho}_{N,1} & \cdots & \cdots & \hat{\rho}_{N,N} \end{pmatrix}$$

3. Use multiple testing (Bailey et al., 2016, 2019) to obtain significant cross-correlations with $\rho_{i,j} > c_p = \phi^{-1} \left(1 - \frac{p/2}{n^\delta}\right)$ which then gives $\hat{\boldsymbol{W}}$.[3] Row standardize $\hat{\boldsymbol{W}}$.

4. Calculate spatial lags as $\sum_{s=1}^{N} \hat{w}_{i,s} y_{i,t}$ and $\sum_{s=1}^{N} \hat{w}_{i,s} x_{i,t}$.

[1] For some other approaches see Holly et al. (2010) and Holly et al. (2011).

[2] Metropolitan statistical areas are delineated by the U.S. Office of Management and Budget (OMB) and usually consist of a core city with a large population and its surrounding region, which may include several adjacent counties. The area defined by the MSA is typically marked by significant social and economic interaction. People living in outlying rural areas, for example, may commute considerable distances to work, shop, or attend social activities in the urban center.

[3] In application contexts, δ is not known *a priori*, so we need to admit some uncertainty. For our baseline specification we set $\delta = 0.7$, the nominal size of the test is set to $p = 0.05$ and ϕ^{-1} is the inverse of the cumulative distribution of a standard normal variable. Interpretation of results are robust to varying $\delta = [0.5, 1, 2, 4]$; please see Appendix for these additional results.

5. Calculate spatio (-temporal) first differences as $\Delta\underline{\Delta}y_{i,t} = y_{i,t} - y_{i,t-1} - \mathbf{w}_i\mathbf{y}_t + \hat{\mathbf{w}}_i\mathbf{y}_{t-1}$, $\underline{\Delta}y_{i,t} = y_{i,t} - \hat{\mathbf{w}}_i\mathbf{y}_t$ and same for $\Delta\underline{\Delta}x_{i,t}$ and $\underline{\Delta}x_{i,t}$.

6. Estimate the following models based on Equations (3) and (4):

$$\begin{aligned}
\Delta\underline{\Delta}_S y_{i,t} =& \beta_{i,0} + \beta_{i,1}\Delta\underline{\Delta}_S x_{i,t} \\
&- \phi_i\left(y_{i,t-1} - \kappa_i x_{i,t-1}\right) - \lambda_i\left(\boldsymbol{w}_{i,L}\boldsymbol{y}_t - \gamma_i\boldsymbol{w}_{i,L}\boldsymbol{x}_t\right) \\
&+ \psi_{i,1}\boldsymbol{w}_{i,S}\boldsymbol{y}_{t-1} + \psi_{i,2}\boldsymbol{w}_{i,S}\boldsymbol{x}_{t-1} + e_{i,t} \\
\Delta\underline{\Delta}_S y_{i,t} =& \beta_{i,0} + \beta_{i,1}\Delta\underline{\Delta}_S x_{i,t} \\
&- \phi_i\left(\underline{\Delta}_L y_{i,t-1} - \kappa_i\underline{\Delta}_L x_{i,t-1}\right) - \lambda_i\left(\boldsymbol{w}_{i,L}\Delta\boldsymbol{y}_t - \gamma_i\boldsymbol{w}_{i,L}\Delta\boldsymbol{x}_t\right) \\
&+ \psi^{\star}_{i,1}\boldsymbol{w}_{i,S}\boldsymbol{y}_{t-1} + \psi^{\star}_{i,2}\boldsymbol{w}_{i,S}\boldsymbol{x}_{t-1} + e^{\star}_{i,t}
\end{aligned}$$

   where $\boldsymbol{w}_{i,S} = \hat{\boldsymbol{w}}_i$ and $\boldsymbol{w}_{i,L} = \hat{\boldsymbol{w}}_{CSA}$. Further cross-sectional averages can be added to both regressions.

Alternatively we set $\boldsymbol{w}_{i,S} = \hat{\boldsymbol{w}}_{i,c}$ for the contiguity based spatial weight matrix or $\boldsymbol{w}_{i,S} = \hat{\boldsymbol{w}}_{i,d}$ for a distance based weight matrix. In this case the short run spatial weight matrix is known, no multiple testing is needed and hence only steps (4) - (6) are required.

Implicit in what we are doing is to provide a temporal and spatial cointegrating relationship between real house prices and real per capita incomes at the MSA level. Although we do not provide any formal proofs of cointegration, we adopt the DOLS approach of Banerjee et al. (1986) and Stock and Watson (1993) where a statistically significant λ is a measure of temporal cointegration and a statistically significant ϕ a measure of spatial cointegration.

The results are reported in Table 1. The first 3 columns provide the current best practice estimates of a temporal error correction model with common correlated effects for panel data. The results, however, are quite mixed with the estimate of the long relationship (κ) between real house prices and real incomes varying considerably, ranging between 0.043 and 0.889, depending upon the number of lags used. Weak cross sectional dependence of residuals is rejected at the 5% level except for column 1. However, the error correction model in column 1 produces counterintuitive findings with respect to temporal cointegration. Overall these findings are not satisfactory, reflecting large heterogeneity across MSAs in the USA, with the implication that we are likely to find better insights on equilibrium, cointegration and partial adjustment in our spatial and spatio-temporal models. Columns 4 and 5 provide estimates for a spatial error correction model on its own. These results are also unsatisfactory. This reflects yet again the potential of the model developed here, modelling jointly potential cointegration at both the temporal and spatial dimensions.

Finally, in columns 6 to 9 we report estimates of the spatial and temporal error correction models together. The lags are either 0 or 3 and the weighting matrices ($\hat{w}$) from the spatial weights or $1/N$ the cross sectional averages. Further, house prices and incomes are cointegrated only spatially but not temporarily. This is in sharp contrast with housing markets in the UK, where

Bhattacharjee et al. (2022) found evidence of cointegration across both the spatial and temporal dimensions. A key observation is that there is very large heterogeneity across MSAs in the unit-specific estimates of the temporal long run coefficient, such that the standard error of the mean group estimator is very large. The estimated value of λ suggests that at the spatial level the response of immediate neighbours to a shock to a MSA is very close to unity, suggesting that arbitrage opportunities are negated strongly in housing markets. This is in line with the law of one price intuition provided by the agricultural economics literature discussed earlier. By contrast the results for the UK suggest a much more integrated domain. It would take 78 UKs to match the size of the US. The sheer vastness of the US continent suggests that it is better to think of spatial equilibrium in house prices but not necessarily at the temporal dimension.

Some support for "ripple effects" is found in the observation of strong cross section dependence and spatial cointegration. Indeed, detailed MSA-level analysis reflects higher correlation in incomes in California with cross-section average incomes across all MSAs, but the same is not observed for house prices. This indicates that house price shocks are more likely to originate in more affluent MSAs in California and spread across the country. Likewise, in the construction of our cross-correlation spatial weights, we find strongest correlations between affluent MSAs in California and Florida, which is in line with intuition. Finally, the cross-correlation spatial weights are located in only 3.5% or all cross-correlations, which also highlights the strength of the methods employed. Overall, spatial nonstationary and stationary dynamics are very rich and support the theory and model development in this paper.

In Table 2 we also explore whether a distance measure of contiguity[4] provides a better way to model spatial closeness. Although the results are similar to Table 1, the null of weak dependence is rejected in all cases. As discussed, robustness with regard to choice of δ is reported as additional results in the Appendix.

5 Conclusion

Unit roots, cointegration and error correction are well-understood and useful concepts in temporal stochastic processes. However, equivalent concepts and methods for spatial and network contexts are generally lacking, even if there are some developments on spatial unit roots and strong dependence. Our first contribution is to develop a spatial Granger representation theorem, together with an error correction model, to precisely locate potential for cointegration in spatial processes.

Second, we consider panel data and develop Granger representation and error correction across two dimensions: space and time. This provides a way to simultaneously model spatial and temporal processes in a panel framework as a joint error correction mechanism. Idiosyncratic shocks to a spatial unit are weakly translated into shocks to neighbours while a factor-driven shock is

[4]Using longitude and latitude for each MSA we calculate distances using a variant of the haversine formula of Vincenty (1975) programmed by Austin Nichols in Stata.

strongly transmitted across the spatial domain or network. This provides the potential for two error correction mechanisms – one temporal and one spatial – with partial adjustment towards potentially two different long run equilibrium relationships. Short run dynamics can be modelled similar to time series contexts. We find that there is a key distinction between spatial spillover processes in weak and strong dependence. While strong dependence spatial weights can be captured by cross-sectional averages (or common correlated effects), the analyst can make a choice of spatial weights to model short run stationary (weak dependence) dynamics.

Third, applying our framework and spatio-temporal error correction models to panel data on housing markets across metropolitan areas of the USA, but we find little evidence of equilibria and cointegration across both dimensions. The sheer size of the US suggests that integration of housing markets may only be a local, spatial feature of the data. Our work suggests substantial further avenues for theory, model development and applications.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Time	Time	Time	Spatio	Spatio	Spatio-Temporal	Spatio-Temporal	Spatio-Temporal	Spatio-Temporal
Dep. Var.	$\Delta y_{i,t}$	$\Delta y_{i,t}$	$\Delta y_{i,t}$	$\underline{\Delta} y_{i,t}$	$\underline{\Delta} y_{i,t}$	$\Delta\underline{\Delta} y_{i,t}$	$\Delta\underline{\Delta} y_{i,t}$	$\Delta\underline{\Delta}_{CSA} y_{i,t}$	$\Delta\underline{\Delta}_{CSA} y_{i,t}$
β	0.287	0.378	0.274	0.204	0.209	0.351	0.284	0.294	0.294
	(0.014)***	(0.015)***	(0.014)***	(0.015)***	(0.016)***	(0.014)***	(0.014)***	(0.014)***	(0.014)***
ψ_1						9.235	1.559	0.061	-0.002
						(3.130)***	(1.290)	(0.016)***	(0.002)
ψ_2						2.729	1.062	-0.057	0.010
						(2.324)	(1.008)	(0.015)***	(0.002)***
Temporal ECM									
ϕ	-0.043	-0.078	-0.042			-0.093	-0.039	-0.044	-0.044
	(0.002)***	(0.003)***	(0.002)***			(0.005)***	(0.002)***	(0.002)***	(0.002)***
κ	0.043	0.889	0.150			1.361	0.875	-0.008	-0.008
	(0.004)***	(0.236)***	(0.771)			(0.557)**	(0.534)	(0.831)	(0.831)
Spatial ECM									
λ				-375.123	-375.115	0.124	0.986	-0.019	-0.019
				(0.033)***	(0.035)***	(0.005)***	(0.015)***	(0.015)	(0.015)
γ				0.205	0.209	0.789	0.252	0.132	0.132
				(0.015)***	(0.016)***	(0.286)***	(0.021)***	(0.329)	(0.329)
N	375	375	375	375	375	375	375	375	375
T	185	184	181	185	182	184	184	184	184
CD	-1.361	1006.329	-3.938	358.630	237.026	954.987	0.763	1.332	1.332
p	0.174	0.000	0.000	0.000	0.000	0.000	0.445	0.183	0.183
CSA	lrh lry	lrh lry	lrh lry	lrh lry	lrh lry				
Lags (p_{CSA})	2	0	3	0	3				
$\boldsymbol{w}_S$				$\hat{\boldsymbol{w}}_i$	$\hat{\boldsymbol{w}}_i$	$\hat{\boldsymbol{w}}_i$	$\hat{\boldsymbol{w}}_i$	$\boldsymbol{w}_{CSA}$	$\boldsymbol{w}_{CSA}$
$\boldsymbol{w}_L$				$\boldsymbol{w}_{CSA}$	$\boldsymbol{w}_{CSA}$	$\boldsymbol{w}_{CSA}$	$\boldsymbol{w}_{CSA}$	$\boldsymbol{w}_{CSA}$	$\boldsymbol{w}_{CSA}$
EQ						(5)	(6)	(5)	(6)

$^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table 1: Spatial Temporal Error Correction Model based on Eq. (5) and (6). CSA indicates variables added as cross-section averages.
Column (1) - (3): $\Delta y_{i,t} = \beta_{i,0} + \beta_{i,1}\Delta x_{i,t} - \phi_i\left(y_{i,t-1} - \kappa_i x_{i,t-1}\right) + \sum_{l=0}^{p_{CSA}} \boldsymbol{\gamma}_i \bar{\boldsymbol{Z}}_{i,t} + e_{i,t}$.
Column 4 & 5: $\underline{\Delta}_S y_{i,t} = \beta_{i,0} + \beta_{i,1}\underline{\Delta}_S x_{i,t} - \lambda_i\left(\boldsymbol{w}_{i,L}\boldsymbol{y}_t - \gamma_i \boldsymbol{w}_{i,L}\boldsymbol{x}_t\right) + \sum_{l=0}^{p_{CSA}} \boldsymbol{\gamma}_i \bar{\boldsymbol{Z}}_{i,t} + e_{i,t}$.
Column 6 & 8: $\Delta\underline{\Delta}_S y_{i,t} = \beta_{i,0} + \beta_{i,1}\Delta\underline{\Delta}_S x_{i,t} - \phi_i\left(y_{i,t-1} - \kappa_i x_{i,t-1}\right) - \lambda_i\left(\boldsymbol{w}_{i,L}\boldsymbol{y}_t - \gamma_i \boldsymbol{w}_{i,L}\boldsymbol{x}_t\right) + \psi_{i,1}\boldsymbol{w}_{i,S}\boldsymbol{y}_{t-1} + \psi_{i,2}\boldsymbol{w}_{i,S}\boldsymbol{x}_{t-1} + e_{i,t}$.
Column 7 & 9: $\Delta\underline{\Delta}_S y_{i,t} = \beta_{i,0} + \beta_{i,1}\Delta\underline{\Delta}_S x_{i,t} - \phi_i\left(\underline{\Delta}_L y_{i,t-1} - \kappa_i \underline{\Delta}_L x_{i,t-1}\right) - \lambda_i\left(\boldsymbol{w}_{i,L}\Delta\boldsymbol{y}_t - \gamma_i \boldsymbol{w}_{i,L}\Delta\boldsymbol{x}_t\right) + \psi^{\star}_{i,1}\boldsymbol{w}_{i,S}\boldsymbol{y}_{t-1} + \psi^{\star}_{i,2}\boldsymbol{w}_{i,S}\boldsymbol{x}_{t-1} + e^{\star}_{i,t}$.
where $\underline{\Delta}_k y_{i,t} = y_{i,t} - \boldsymbol{w}_{i,k} y_{i,t}, k = [S, L]$ defines the first spatial difference for the short or long run relationship. The spatial weights and spatial first differences are either constructed using cross-correlations ($\hat{\boldsymbol{w}}_i; \underline{\Delta}$), contiguity ($\boldsymbol{w}_{i,c}; \underline{\Delta}_c$), distance ($\boldsymbol{w}_{i,d}; \underline{\Delta}_d$) or (cross-section) averages ($\boldsymbol{w}_{i,csa} = 1/N(\boldsymbol{1}'\boldsymbol{1} - I_N); \underline{\Delta}_{csa}$). Δ is the first difference in the time dimension.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Time	Time	Time	Spatio	Spatio	Spatio-Temporal	Spatio-Temporal	Spatio-Temporal	Spatio-Temporal
Dep. Var.	$\Delta y_{i,t}$	$\Delta y_{i,t}$	$\Delta y_{i,t}$	$\underline{\Delta}_{c,i} y_{i,t}$	$\underline{\Delta}_{c,i} y_{i,t}$	$\Delta\underline{\Delta}_{c,i} y_{i,t}$	$\Delta\underline{\Delta}_{c,i} y_{i,t}$	$\Delta\underline{\Delta}_{CSA} y_{i,t}$	$\Delta\underline{\Delta}_{CSA} y_{i,t}$
β	0.287	0.378	0.274	0.248	0.252	0.351	0.286	0.294	0.294
	(0.014)***	(0.015)***	(0.014)***	(0.015)***	(0.015)***	(0.014)***	(0.014)***	(0.014)***	(0.014)***
ψ_1						21.643	2.651	0.061	-0.002
						(2.808)***	(0.673)***	(0.016)***	(0.002)
ψ_2						-3.487	1.930	-0.057	0.010
						(2.186)	(0.681)***	(0.015)***	(0.002)***
Temporal ECM									
ϕ	-0.043	-0.078	-0.042			-0.092	-0.043	-0.044	-0.044
	(0.002)***	(0.003)***	(0.002)***			(0.005)***	(0.002)***	(0.002)***	(0.002)***
κ	0.043	0.889	0.150			-0.858	-2.982	-0.008	-0.008
	(0.004)***	(0.236)***	(0.771)			(1.692)	(3.271)	(0.831)	(0.831)
Spatial ECM									
λ				-375.034	-375.037	0.124	0.981	-0.019	-0.019
				(0.014)***	(0.014)***	(0.005)***	(0.015)***	(0.015)	(0.015)
γ				0.248	0.253	-0.278	0.248	0.132	0.132
				(0.015)***	(0.015)***	(0.642)	(0.021)***	(0.329)	(0.329)
N	375	375	375	375	375	375	375	375	375
T	185	184	181	185	182	184	184	184	184
CD	-1.361	1006.329	-3.938	81.730	61.330	955.233	3.820	1.332	1.332
p	0.174	0.000	0.000	0.000	0.000	0.000	0.000	0.183	0.183
CSA	lrh lry	lrh lry	lrh lry	lrh lry	lrh lry				
Lags (p_{CSA})	2	0	3	0	3				
$\boldsymbol{w}_S$				$\boldsymbol{w}_{c,i}$	$\boldsymbol{w}_{c,i}$	$\boldsymbol{w}_{c,i}$	$\boldsymbol{w}_{c,i}$	$\boldsymbol{w}_{CSA}$	$\boldsymbol{w}_{CSA}$
$\boldsymbol{w}_L$				$\boldsymbol{w}_{CSA}$	$\boldsymbol{w}_{CSA}$	$\boldsymbol{w}_{CSA}$	$\boldsymbol{w}_{CSA}$	$\boldsymbol{w}_{CSA}$	$\boldsymbol{w}_{CSA}$
EQ						(5)	(6)	(5)	(6)

$^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table 2: Contiguity based weight matrix $\boldsymbol{w}_c$. Spatial Temporal Error Correction Model based on Eq. (5) and (6). CSA indicates variables added as cross-section averages. For Notes see Table 1.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Time	Time	Time	Spatio	Spatio	Spatio-Temporal	Spatio-Temporal	Spatio-Temporal	Spatio-Temporal
Dep. Var.	$\Delta y_{i,t}$	$\Delta y_{i,t}$	$\Delta y_{i,t}$	$\underline{\Delta}_{d,i} y_{i,t}$	$\underline{\Delta}_{d,i} y_{i,t}$	$\Delta\underline{\Delta}_{d,i} y_{i,t}$	$\Delta\underline{\Delta}_{d,i} y_{i,t}$	$\Delta\underline{\Delta}_{CSA} y_{i,t}$	$\Delta\underline{\Delta}_{CSA} y_{i,t}$
β	0.287	0.378	0.274	0.294	0.293	0.408	0.292	0.294	0.294
	(0.014)***	(0.015)***	(0.014)***	(0.009)***	(0.009)***	(0.015)***	(0.014)***	(0.014)***	(0.014)***
ψ_1						149.788	4.093	0.061	-0.002
						(6.878)***	(0.799)***	(0.016)***	(0.002)
ψ_2						-157.142	-0.225	-0.057	0.010
						(4.721)***	(0.798)	(0.015)***	(0.002)***
Temporal ECM									
ϕ	-0.043	-0.078	-0.042			-0.005	-0.043	-0.044	-0.044
	(0.002)***	(0.003)***	(0.002)***			(0.005)	(0.002)***	(0.002)***	(0.002)***
κ	0.043	0.889	0.150			0.666	0.518	-0.008	-0.008
	(0.004)***	(0.236)***	(0.771)			(0.639)	(0.374)	(0.831)	(0.831)
Spatial ECM									
λ				-375.010	-375.010	0.423	0.978	-0.019	-0.019
				(0.003)***	(0.003)***	(0.010)***	(0.015)***	(0.015)	(0.015)
γ				0.294	0.293	1.130	0.254	0.132	0.132
				(0.009)***	(0.009)***	(0.069)***	(0.021)***	(0.329)	(0.329)
N	375	375	375	375	375	375	375	375	375
T	185	184	181	185	182	184	184	184	184
CD	-1.361	1006.329	-3.938	126.865	126.072	470.579	1.966	1.332	1.332
p	0.174	0.000	0.000	0.000	0.000	0.000	0.049	0.183	0.183
CSA	lrh lry	lrh lry	lrh lry	lrh lry	lrh lry				
Lags (p_{CSA})	2	0	3	0	3				
$\boldsymbol{w}_S$				$\boldsymbol{w}_{d,i}$	$\boldsymbol{w}_{d,i}$	$\boldsymbol{w}_{d,i}$	$\boldsymbol{w}_{d,i}$	$\boldsymbol{w}_{CSA}$	$\boldsymbol{w}_{CSA}$
$\boldsymbol{w}_L$				$\boldsymbol{w}_{CSA}$	$\boldsymbol{w}_{CSA}$	$\boldsymbol{w}_{CSA}$	$\boldsymbol{w}_{CSA}$	$\boldsymbol{w}_{CSA}$	$\boldsymbol{w}_{CSA}$
EQ						(5)	(6)	(5)	(6)

$^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table 3: Distance based weight matrix $\boldsymbol{w}_d$. Spatial Temporal Error Correction Model based on Eq. (5) and (6). CSA indicates variables added as cross-section averages. For Notes see Table 1.

References

Anselin, Luc, *Spatial Econometrics: Methods and Models*, Springer: New York, 1988.

_ **, Anil K. Bera, Raymond Florax, and Mann J. Yoon**, "Simple diagnostic tests for spatial dependence," *Regional Science and Urban Economics*, 1996, *26* (1), 77–104.

Aquaro, Michele, Natalia Bailey, and M. Hashem Pesaran, "Estimation and inference for spatial models with heterogeneous coefficients: An application to US house prices," *Journal of Applied Econometrics*, 2021, *36* (1), 18–44.

Asche, Frank, Daniel V. Gordon, and Rognvaldur Hannesson, "Tests For Market Integration and the Law of One Price: The Market For Whitefish in France," *Marine Resource Economics*, 2004, *19* (2), 195–210.

Bailey, Natalia, M Hashem Pesaran, and L. Vanessa Smith, "A multiple testing approach to the regularisation of large sample correlation matrices," *Journal of Econometrics*, 2019, *208* (2), 507–534.

_ **, Sean Holly, and M. Hashem Pesaran**, "A Two-Stage Approach to Spatio-Temporal Analysis with Strong and Weak Cross-Sectional Dependence," *Journal of Applied Econometrics*, 2016, *31* (1), 249–280.

Baltagi, Badi H., Seuck H. Song, and Won Koh, "Testing panel data regression models with spatial error correlation," *Journal of Econometrics*, 1996, *117* (1), 123–150.

Banerjee, A., J. J. Dolado, F. D. Hendry, and G. W. Smith, "Exploring equilibrium relations in econometrics through state models: Some monte carlo evidence," *Oxford Bulletin of Economics and Statistics*, 1986, *48*, 253–278.

Beenstock, Michael, Dan Feldman, and Daniel Felsenstein, "Testing for Unit Roots and Cointegration in Spatial Cross-Section Data," *Spatial Economic Analysis*, 2012, *7* (2), 203–222.

Bhattacharjee, Arnab, Chris Higson, and Sean Holly, "Asymmetric price adjustment, sticky costs and operating leverage over the business cycle," Technical Report, Working Paper No. 1402, Spatial Economics and Econometrics Centre, Heriot-Watt University 2014.

_ **, Jan Ditzen, and Sean Holly**, "Spatial and Spatio-Temporal Error Correction, Networks and Common Correlated Effects," in Alexander Chudik, Cheng Hsiao, and Allan Timmermann, eds., *Advances in Econometrics*, Vol. 43B, Emerald Publishing Limited, 2022, chapter Essays in, pp. 37–60.

Chiang, Ming-Chu and I-Chun Tsai, "Ripple effect and contagious effect in the US regional housing markets," *The Annals of Regional Science*, 2016, *56*, 55–82.

Cook, Steven and Sean Holly, "Statistical properties of UK house prices: An analysis of disaggregated vintages," *Urban Studies*, 2000, *37* (11), 2045–2051.

Drake, L, "Testing for convergence between UK regional house prices.," *Regional Studies*, 1995, *29*, 357–366.

Engle, Robert F. and C. W. J. Granger, "Co-Integration and Error Correction: Representation, Estimation, and Testing," *Econometrica*, 1987, *55* (2), 251.

Fingleton, Bernard, "Spurious Spatial Regression: Some Monte Carlo Results with a Spatial Unit Root and Spatial Cointegration," *Journal of Regional Science*, 1999, *39* (1), 1–19.

_ , "Spatial Autoregression," *Geographical Analysis*, 2009, *41* (4), 385–391.

Goodwin, B.K. and T.C. Schroeder, "Co-integration Tests and Spatial Price Linkages in Regional Cattle Markets," *American Journal of Agricultural Economics*, 1991, *73*, 452–64.

Gordon, D.V., J.E. Hobbs, and W.A. Kerr, "A Test for Price Integration in the EC Lamb Market," *Journal of Agricultural Economics*, 2004, *44*, 126–34.

Granger, C. W. J., "Co-Integrated Variables and Error-Correcting Models," Technical Report, UCSD Discussion Paper 83-13, University of San Diego CA 1983.

Granger, C.W.J., "Some properties of time series data and their use in econometric model specification," *Journal of Econometrics*, 1981, *16* (1), 121–130.

_ **and A.A. Weiss**, "Time Series Analysis of Error-Correction Models," in Samuel Karlin, Takeshi Amemiya, and Leo A. Goodman, eds., *Studies in Econometrics, Time Series, and Multivariate Statistics*, Academic Press, 1983, pp. 255–278.

Gupta, Abhimanyu and Peter M. Robinson, "Inference on higher-order spatial autoregressive models with increasingly many parameters," *Journal of Econometrics*, 2015, *186* (1), 19–31.

_ **and** _ , "Pseudo maximum likelihood estimation of spatial autoregressive models with increasing dimension," *Journal of Econometrics*, 2018, *202* (1), 92–107.

Heblich, Stephan, Alex Trew, and Yanos Zylberberg, "East-Side Story: Historical Pollution and Persistent Neighborhood Sorting," *Journal of Political Economy*, 2021, *129* (5), 1508–1552.

Holly, Sean, M. Hashem Pesaran, and Takashi Yamagata, "A spatio-temporal model of house prices in the USA," *Journal of Econometrics*, 2010, *158* (1), 160–173.

_ , _ , and _ , "The spatial and temporal diffusion of house prices in the UK," *Journal of Urban Economics*, jan 2011, *69* (1), 2–23.

Johansen, Søren, "A Statistical Analysis of Cointegration for I(2) Variables," *Econometric Theory*, 1995, *11* (1), 25–59.

Kuiper, W. Erno and Matthew T. G. Meulenberg, "Cointegration and prediction analysis of market supply in the Dutch pig-farming industry," *European Review of Agricultural Economics*, 06 1997, *24* (2), 285–312.

Kumar, Paul Ranjit and Tanmoy Karak, "Asymmetric Price Transmission: A Case of Wheat in India," *Agriculture*, 2022, *12* (3).

Meen, Geoff, "Spatial housing economics: A survey," *Urban Studies*, 2016, *74* (10), 1987–2003.

Müller, Ulrich K. and Mark W. Watson, "Spatial Unit Roots," Technical Report, Princeton University 2023.

Muwanga, Gertrude S. and Donald L. Snyder, "Market integration and the law of one price: Case study of selected feeder cattle markets," 1997.

Pesaran, M. Hashem, "Estimation and inference in large heterogeneous panels with a multifactor error structure," *Econometrica*, 2006, *74* (4), 967–1012.

_ , Yongcheol Shin, and Ron P. Smith, "Pooled mean group estimation of dynamic heterogeneous panels," *Journal of the American Statistical Association*, 1999, *94* (446), 621–634.

Phillips, A.W.H., "Stabilization Policy in a Closed Economy," *Economic Journal*, 1954, *64* (254), 290–323.

_ , "Stabilization Policy and the Time Form of Lagged Responses," *Economic Journal*, 1957, *67* (266), 265–277.

Ravallion, Martin, "Testing Market Integration," *American Journal of Agricultural Economics*, 1986, pp. 102–109.

Sargan, Denis, "Wages and Prices in the United Kingdom: A Study in Econometric Methodology," in Peter Edward Hart, Gordon Mills, and John King Whitaker, eds., *Econometric Analysis for National Economic Planning.*, Butterworth, 1964, pp. 25–63.

Sekhar, C.S.C., "Agricultural market integration in India: An analysis of select commodities," *Food Policy*, 2012, *37* (3), 309–322.

Stock, James H and Mark W. Watson, "A simple estimator of cointegrating vectors in higher order integrated systems," *Econometrica*, 1993, *61*, 783–820.

Vincenty, T., "Direct and Inverse Solutions of Geodesics on the Ellipsoid with application of nested equations," *Survey Review*, 1975, *23* (176), 88–93.

von Cramon-Taubadel, Stephan and Barry K. Goodwin, "Price Transmission in Agricultural Markets," *Annual Review of Resource Economics*, 2021, *13*, 65–84.

Yang, Cynthia Fan, "Common factors and spatial dependence: an application to US house prices," *Econometric Reviews*, 2021, *40* (1), 14–50.

A Appendix

In this section we present results where we allowed δ to vary during the multiple testing approach when obtaining significant cross-correlations using $\rho_{i,j} > c_p = \phi^{-1}\left(1 - \frac{p/2}{n^{\delta}}\right)$. δ is varied between $\delta = [0.5, 1, 2, 4]$. Tables A1 and A2 present the results. We also report the share of non-zero elements in $\hat{\boldsymbol{W}}$. We note that an increase in δ increases the sparsity in $\hat{\boldsymbol{W}}$, implying a lower percentage of non-zero elements. We further note that the estimate for β and γ also decreases in size for the Spatial ECMs (Column 2,3, respectively 7 and 8). Results for the spatial temporal error correction models (Columns 4,5, respectively 9 and 10) are less affected.

	$\delta = 0.5$					$\delta = 1$				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Time	Spatio	Spatio	Spatio-Temporal	Spatio-Temporal	Time	Spatio	Spatio	Spatio-Temporal	Spatio-Temporal
Dep. Var.	$\Delta y_{i,t}$	$\underline{\Delta} y_{i,t}$	$\underline{\Delta} y_{i,t}$	$\Delta\underline{\Delta} y_{i,t}$	$\Delta\underline{\Delta} y_{i,t}$	$\Delta y_{i,t}$	$\underline{\Delta} y_{i,t}$	$\underline{\Delta} y_{i,t}$	$\Delta\underline{\Delta} y_{i,t}$	$\Delta\underline{\Delta} y_{i,t}$
β	0.287	0.283	0.284	0.342	0.278	0.287	0.111	0.113	0.357	0.291
	(0.014)***	(0.016)***	(0.017)***	(0.014)***	(0.014)***	(0.014)***	(0.013)***	(0.013)***	(0.014)***	(0.014)***
ψ_1				7.743	1.465				5.474	0.400
				(3.136)**	(1.098)				(2.454)**	(0.786)
ψ_2				4.196	0.599				1.620	1.418
				(2.355)*	(1.019)				(2.302)	(0.799)*
Temporal ECM										
ϕ	-0.043			-0.103	-0.041	-0.043			-0.084	-0.034
	(0.002)***			(0.006)***	(0.002)***	(0.002)***			(0.005)***	(0.002)***
κ	0.043			1.671	0.967	0.043			6.867	0.401
	(0.004)***			(0.713)**	(0.510)*	(0.004)***			(6.270)	(0.275)
Spatial ECM										
λ		-375.205	-375.191	0.135	0.984		-375.072	-375.071	0.115	0.986
		(0.045)***	(0.049)***	(0.005)***	(0.015)***		(0.025)***	(0.027)***	(0.004)***	(0.015)***
γ		0.284	0.285	-0.046	0.253		0.112	0.113	-3.796	0.262
		(0.016)***	(0.017)***	(0.466)	(0.021)***		(0.013)***	(0.013)***	(5.352)	(0.021)***
N	375	375	375	375	375	375	375	375	375	375
T	185	185	182	184	184	185	185	182	184	184
CD	-1.361	56.571	38.674	932.316	0.916	-1.361	1180.543	742.948	973.236	0.522
p	0.174	0.000	0.000	0.000	0.360	0.174	0.000	0.000	0.000	0.602
CSA	lrh lry	lrh lry	lrh lry			lrh lry	lrh lry	lrh lry		
Lags (p_{CSA})	2	0	3			2	0	3		
$\boldsymbol{w}_S$		$\hat{\boldsymbol{w}}_i$	$\hat{\boldsymbol{w}}_i$	$\hat{\boldsymbol{w}}_i$	$\hat{\boldsymbol{w}}_i$		$\hat{\boldsymbol{w}}_i$	$\hat{\boldsymbol{w}}_i$	$\hat{\boldsymbol{w}}_i$	$\hat{\boldsymbol{w}}_i$
$\boldsymbol{w}_L$		$\boldsymbol{w}_{CSA}$	$\boldsymbol{w}_{CSA}$	$\boldsymbol{w}_{CSA}$	$\boldsymbol{w}_{CSA}$		$\boldsymbol{w}_{CSA}$	$\boldsymbol{w}_{CSA}$	$\boldsymbol{w}_{CSA}$	$\boldsymbol{w}_{CSA}$
EQ				(5)	(6)				(5)	(6)
Pct Non-Zero	3.67					0.83				

$^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table A1: Estimated spatial weight matrix, $\delta = 0.5$ and $\delta = 1$. Spatial Temporal Error Correction Model based on Eq. (5) and (6). CSA indicates variables added as cross-section averages. For Notes see Table 1.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	$\delta = 2$					$\delta = 3$				
	Time	Spatio	Spatio	Spatio-Temporal	Spatio-Temporal	Time	Spatio	Spatio	Spatio-Temporal	Spatio-Temporal
Dep. Var.	$\Delta y_{i,t}$	$\underline{\Delta} y_{i,t}$	$\underline{\Delta} y_{i,t}$	$\Delta\underline{\Delta} y_{i,t}$	$\Delta\underline{\Delta} y_{i,t}$	$\Delta y_{i,t}$	$\underline{\Delta} y_{i,t}$	$\underline{\Delta} y_{i,t}$	$\Delta\underline{\Delta} y_{i,t}$	$\Delta\underline{\Delta} y_{i,t}$
β	0.287	0.030	0.030	0.374	0.299	0.287	0.001	0.001	0.378	0.302
	(0.014)***	(0.008)***	(0.008)***	(0.015)***	(0.014)***	(0.014)***	(0.002)	(0.002)	(0.015)***	(0.014)***
ψ_1				1.199	-0.322				0.527	-0.044
				(1.038)	(0.195)*				(1.016)	(0.049)
ψ_2				-0.766	0.340				-0.433	0.099
				(1.219)	(0.328)				(0.285)	(0.134)
Temporal ECM										
ϕ	-0.043			-0.076	-0.028	-0.043			-0.077	-0.027
	(0.002)***			(0.003)***	(0.001)***	(0.002)***			(0.003)***	(0.001)***
κ	0.043			1.027	0.196	0.043			0.921	0.225
	(0.004)***			(0.280)***	(0.209)	(0.004)***			(0.231)***	(0.186)
Spatial ECM										
λ		-375.026	-375.026	0.103	0.990		-375.000	-375.000	0.100	0.990
		(0.011)***	(0.011)***	(0.004)***	(0.015)***		(0.001)***	(0.001)***	(0.004)***	(0.015)***
γ		0.030	0.030	3.028	0.267		0.001	0.001	2.987	0.268
		(0.008)***	(0.008)***	(1.990)	(0.021)***		(0.002)	(0.002)	(1.992)	(0.021)***
N	375	375	375	375	375	375	375	375	375	375
T	185	185	182	184	184	185	185	182	184	184
CD	-1.361	2959.744	1698.124	1003.469	-0.398	-1.361	3349.426	2004.215	1005.399	-0.661
p	0.174	0.000	0.000	0.000	0.691	0.174	0.000	0.000	0.000	0.509
CSA	lrh lry	lrh lry	lrh lry			lrh lry	lrh lry	lrh lry		
Lags (p_{CSA})	2	0	3			2	0	3		
$\boldsymbol{w}_S$		$\hat{\boldsymbol{w}}_i$	$\hat{\boldsymbol{w}}_i$	$\hat{\boldsymbol{w}}_i$	$\hat{\boldsymbol{w}}_i$		$\hat{\boldsymbol{w}}_i$	$\hat{\boldsymbol{w}}_i$	$\hat{\boldsymbol{w}}_i$	$\hat{\boldsymbol{w}}_i$
$\boldsymbol{w}_L$		$\boldsymbol{w}_{CSA}$	$\boldsymbol{w}_{CSA}$	$\boldsymbol{w}_{CSA}$	$\boldsymbol{w}_{CSA}$		$\boldsymbol{w}_{CSA}$	$\boldsymbol{w}_{CSA}$	$\boldsymbol{w}_{CSA}$	$\boldsymbol{w}_{CSA}$
EQ				(5)	(6)				(5)	(6)
Pct Non-Zero	0.09					0.01				

$^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table A2: Estimated spatial weight matrix, $\delta = 2$ and $\delta = 3$. Spatial Temporal Error Correction Model based on Eq. (5) and (6). CSA indicates variables added as cross-section averages. For Notes see Table 1.