unibz

Fakultät für Wirtschaftswissenschaften
Facoltà di Economia
Faculty of Economics and Management

BEMPS –
Bozen Economics & Management
Paper Series

NO 76/ 2021

Spatial and Spatio-temporal Error Correction, Networks and Common Correlated Effects

Arnab Bhattacharjee, Jan Ditzen,
Sean Holly

Spatial and Spatio-temporal Error Correction, Networks and Common Correlated Effects.

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

December 23, 2020

Abstract

We provide a way to represent spatial and temporal equilibria in terms of error correction models in a panel setting. This requires potentially two different processes for 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 nonstationary. 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 corrrelated effects estimator together with multiple testing of cross-correlations to provide the short-run spatial weights. We apply this model to the 324 local authorities of England, and show that our approach is useful for modelling weak and strong cross-section dependence, together with partial adjustments to two long-run equilibrium relationships and short-run spatio-temporal dynamics, and provides exciting new insights.
JEL Codes: C21, C22, C23, R3.
Keywords: Spatio-temporal dynamics; Error Correction Models; Weak and strong cross sectional dependence.

1 Introduction

The (temporal) error correction representation theorem of Engle-Granger clarifies the role of long run equilibrium, partial adjustment to disequilibrium and short-run dynamics. Here, we develop error correction models, first for a spatial or network framework, and then for a spatio-temporal framework, to study short and long-run dynamics of spatial or network panel data over time and space.

*Corresponding author: a.bhattacharjee@hw.ac.uk. [1]: Heriot-Watt University and National Institute of Economic & Social Research, UK. [2]: Free University of Bozen-Bolzano, Italy, and Center for Energy Economics Research and Policy (CEERP), Heriot-Watt University, Edinburgh, UK. [3]: Faculty of Economics, University of Cambridge, UK.

Both representations offer simple interpretation as error correction models analogous to the temporal case. Specifically, in a spatio-temporal setting and under some simplifying assumptions, there is partial adjustment to two equilibrium relationships: one in the time dimension and the other in the cross-sectional dimension.[1] This simplification relies on two homogeneity assumptions. The first is the familiar pooled mean group assumption (Pesaran et al., 1999) of a homogeneous temporal equilibrium relationship across all the panel units, and the second is an analogous cross-sectional equilibrium homogeneous over time.[2] However, the spatial and temporal equilibria are closely connected. Because the model is discrete, disequilibrium in the spatial dimension requires adjustments over time.

Furthermore, our work highlights how the spatial equilibrium can be modelled using, for example, common correlated effects (Pesaran, 2006) which incorporates strong cross-section dependence. Much of the existing literature treats strong dependence, modelled using common correlated effects (CCE) or cross-section averages, as nuisance parameters. Our spatio-temporal ECM shows how these strong dependence effects can be structurally interpreted.

Once strong dependence is adequately modelled, remaining weak dependence at the local, spatial level[3] means partial adjustment to long run equilibrium, for which the current literature provides several estimates of spatial weights matrices. In general, this weak dependence spatial weights matrix is not fully identified (Bhattacharjee and Jensen-Butler, 2013), but it can be estimated under alternate identifying assumptions (Bhattacharjee and Jensen-Butler, 2013; Bhattacharjee and Holly, 2013; Bailey et al., 2016). Applying our spatio-temporal model to house prices at the local authority level in England, we find evidence of temporal and spatial cointegration, as well as substantial short run dynamics which we model by multiple testing on cross-section correlations (Bailey et al., 2016).

The mathematical notation for the remainder is as follows: lowercase letters refer to scalars ($y_{i,t}$), bold lowercase ($\mathbf{y}_t$) to vectors and bold upper case ($\mathbf{W}$) to matrices. We denote temporal first difference by Δ and spatial difference by $\underline{\Delta} = \mathbf{I} - \mathbf{W}$, where $\mathbf{I}$ is the identity matrix and $\mathbf{W}$ is a spatial (or network) weights matrix. Then, $\mathbf{Wy}$ denotes the spatial lag of $\mathbf{y}$. The remainder of the paper is organised as follows. In Section 2, we develop a spatio-temporal Engle-Granger representation and the corresponding error correction model, and then rstimate this in a spatio-temporal setting in Section 3. We develop an

[1]The cross-sectional dimension is notionally located either in a spatial context or within a network. Hereafter when we refer to the spatial dimension we mean a network as well. Networks can be spatial but there are many networks between households, peers and firms that are not spatial in a geographical sense. However, the network architecture itself can be viewed as connections in an abstract spatial domain.

[2]The familiar pooled mean group assumption (Pesaran et al., 1999) restricts the temporal long-run equilibrium relationship to be homogeneous across all the cross-sectional units. Likewise, the corresponding spatial long-run homogeneity assumption requires the equilibrium relationship across the spatial domain to remain the same throughout the period under study.

[3]The term "local" is with respect to the weights matrix and does not necessarily imply geographically proximate locations.

application to UK house prices in Section 4 and Section 5 concludes.

2 Spatial and Spatio-Temporal Error Correction Models

In this section we describe a spatio-temporal error correction model. First, we define the spatial, second, the temporal equilibrium and finally a combination of the two. We start with a model with one lag of the dependent variable and the contemporaneous and the first lag of the explanatory variable. In addition to the above temporal lags, the model includes a spatial lag of the dependent and independent variables:

$$\begin{aligned} y_{i,t} &= \beta_{i,0} + \beta_{i,1} x_{i,t} + \beta_{i,2} x_{i,t-1} + \alpha_i y_{i,t-1} \\ &\quad + \pi_i \sum_{j=1, i \neq j}^{N} w_{ij} x_{j,t} + \rho_i \sum_{j=1, i \neq j}^{N} w_{ij} y_{j,t} + e_{i,t} \end{aligned} \tag{1}$$

$$e_{i,t} = \gamma_i' f_t + \epsilon_{i,t} \tag{2}$$

for $i = 1, ..., N$ and $t = 1, ..., T$. The error component $e_{i,t}$ contains common factors f_t and their loadings γ_i which introduces potential strong cross-sectional dependence and temporal nonstationarity.[4] The idiosyncratic random component $\epsilon_{i,t}$ is an error term with finitely summable autocovariances.

This is a first order spatio-temporal autoregressive distributed lag model. All coefficients are assumed to be heterogeneous across cross-sectional units. Similar to the temporal lags $x_{i,t-1}$ and $y_{i,t-1}$, $\sum_{j=1,j\neq i}^{N} w_{ij} y_{j,t}$ and $\sum_{j=1,j\neq i}^{N} w_{ij} x_{j,t}$ are the spatial lags of $\mathbf{y}$ and $\mathbf{x}$ based on a spatial weights matrix $\mathbf{W}$. For the moment, we are not explicit about the spatial weights, and since they can represent both spatial weak and strong dependence, we do not require them to satisfy the spatial granularity condition of Pesaran (2006).

Both the variables $\mathbf{x}$ and $\mathbf{y}$ are potentially cointegrated across time and space. Therefore an error correction model which takes the time and the spatial or network dimension into account can be used to represent the short and long run relationships. For convenience we re-write the model with the spatial interactions in matrix form:

$$\mathbf{y}_t = (y_{1,t}, ..., y_{N,t})' \quad ; \quad \mathbf{x}_t = (x_{1,t}, ..., x_{N,t})' \tag{3}$$

$$\mathbf{w}_i = (w_{i,1}, ..., w_{i,N}) \tag{4}$$

with

$$w_{i,i} = 0, \text{ for } i = 1, ..., N \tag{5}$$

[4]Without loss of generality we assume only a single common factor. The model can be extended to multiple factors, see for example Pesaran (2006) and Chudik and Pesaran (2015).

and we assume that:

$$\sum_{j=1}^{N} w_{i,j} = 1 \tag{6}$$

The above assumption implies that the spatial weights matrix $\mathbf{W}$ is row-normalised. Together with fixed spatial weights[5] inherent in (1), this assumption ensures that the spatial weights matrix $\mathbf{W} = ((w_{i,j}))_{N\times N}$ has bounded row and column norms as $N \to \infty$. This is analogous to the spatial granularity condition and weak cross-sectional dependence in Pesaran (2006). Note that we allow for potentially negative spatial weights, which can be important in many applications (Bhattacharjee and Holly, 2013; Bailey et al., 2016). An analogous relationship holds for $\mathbf{x}$. Then the model in (1) can be written as:

$$y_{i,t} = \beta_{i,0} + \beta_{i,1}x_{i,t} + \beta_{i,2}x_{i,t-1} + \alpha_i y_{i,t-1} + \pi_i \mathbf{w}_i \mathbf{x}_t + \rho_i \mathbf{w}_i \mathbf{y}_t + e_{i,t}. \tag{7}$$

In the spatial equilibrium, cross-sectional units have the same values for each time period:

$$\begin{aligned} x_{i,t} &= x_{j,t} = x_t^* \\ y_{i,t} &= y_{j,t} = y_t^* \end{aligned} \tag{8}$$

and since $\sum_{j=1,i\neq j}^{N} w_{ij} = 1$, then

$$\begin{aligned} \sum_{j=1,i\neq j}^{N} w_{ij}x_t^* &= x_t^* \\ \sum_{j=1,i\neq j}^{N} w_{ij}y_t^* &= y_t^*. \end{aligned} \tag{9}$$

Two important points are in order. First, in the above, for simplicity of exposition, and without loss of generality, we made the assumption that the spatial long run relationships are:

$$\begin{aligned} \sum_{i=1}^{N} x_{i,t} &= c_{x,t} \\ \sum_{i=1}^{N} y_{i,t} &= c_{y,t}, \end{aligned}$$

which implies that the long run weights matrix is

$$\mathbf{W} = \frac{1}{N} 11' - \mathbf{I},$$

[5]By "fixed", we mean that the spatial weights are fixed numbers, for a given sample size n, that is, they are non-stochastic.

where $\mathbf{1}$ is a $T \times 1$ vector of ones. This is consistent with the use of cross-section averages to model strong dependence (Pesaran, 2006). This simplification is without loss of generality, because if the long run $\mathbf{W}$ has some other form, one can scale $\mathbf{y}$ and $\mathbf{x}$ to set the value at each location equal to its spatial lag. One can also model spatial strong dependence using other methods, for example statistical factors, principal components or interactive fixed effects (Bai and Ng, 2007; Bai, 2009). However, following Pesaran (2006), cross-section averages are adequate in large (N, T) settings.

Secondly, the discussion highlights also that there can be a distinction between the strongly dependent temporal and the weakly dependent spatial weights matrices. We use cross-section averages to represent the temporal relationships for simplicity of exposition, retaining the notation $\mathbf{W}$ for the weakly dependent (or granular) spatial dynamics. However, in our application, we fully exploit the flexibility of specifying weights matrices for the spatial and temporal dynamics.

Coming back to the derivation of the spatial error correction model, substituting the equilibrium values from (9) into equation (7) yields:

$$\begin{aligned} y_t^* =& \beta_{i,0} + \beta_{i,1} x_t^* + \beta_{i,2} x_{t-1}^* + \alpha_i y_{t-1}^* + \pi_i W_i x_t^* + \rho_i W_i y_{t-1}^* & (10) \\ =& \beta_{i,0} + (\beta_{i,1} + \pi_i) x_t^* + \beta_{i,2} x_{t-1}^* + \alpha_i y_{t-1}^* + \rho_i y_t^* \\ \Rightarrow y_t^* =& \frac{\beta_{i,0}}{1-\rho_i} + \frac{\beta_{i,1} + \pi_i}{1-\rho_i} x_t^* + \frac{\beta_{i,2}}{1-\rho_i} x_{t-1}^* + \frac{\alpha_i}{1-\rho_i} y_{t-1}^* & (11) \\ =& \frac{\beta_{i,0}}{\lambda_i} + \gamma_i x_t^* + \eta_i x_{t-1}^* + \delta_i y_{t-1}^* \end{aligned}$$

The final step defines the parameter values in the spatial equilibrium. Denoting by λ_i as the spatial equilibrium effect of $\mathbf{y}$, γ_i the spatial equilibrium effect of $\mathbf{x}$, δ_i the temporal equilibrium effect of $\mathbf{y}$ and η_i of $\mathbf{x}$:

$$\lambda_i = 1 - \rho_i \qquad \delta_i = \frac{\alpha_i}{\lambda_i} \qquad (12)$$

$$\gamma_i = \frac{\beta_{i,1} + \pi_i}{\lambda_i} \qquad \eta_i = \frac{\beta_{i,2}}{\lambda_i} \qquad (13)$$

Next, we define the spatial first difference as $\underline{\Delta} x_{i,t} = x_{i,t} - \mathbf{w}_i \mathbf{x}_t$. The spatial first difference is analogous to the temporal first difference for time series. However, whereas in the time dimension the (causal) ordering is evident, the ordering in a spatial context is less clearly defined. The spatial weight matrix specifies a partial ordering since it assigns non zero values only to those cross-sectional units which are related to each other. Analogous to the first time difference, spatial first differencing removes potential non-stationarity in the spatial dimension. We will refer to stationarity in the spatial dimension as granularity (Chudik et al., 2011). Granularity implies that the location of a unit within a space or a network is arbitrary and none of its connections to and from other units are dominating. This in turn implies weak cross-section dependence.

We use the long run coefficients from equation (12) and (13) to derive the spatial error correction model.

$$
\begin{aligned}
y_{i,t} =& \beta_{i,0} + \beta_{i,1} x_{i,t} + \eta_i \lambda_i x_{i,t-1} + (\gamma_i \lambda_i - \beta_{i,1}) \mathbf{w}_i \mathbf{x}_t + \delta_i \lambda_i y_{i,t-1} \quad (14) \\
& + (1 - \lambda_i) \mathbf{w}_i \mathbf{y}_t + e_{i,t} \\
y_{i,t} - \mathbf{w}_i \mathbf{y}_t =& \beta_{i,0} + \beta_{i,1} (x_{i,t} - \mathbf{w}_i \mathbf{x}_t) + \lambda_i \eta_i x_{i,t-1} + \gamma_i \lambda_i \mathbf{w}_i \mathbf{x}_t \quad (15) \\
& - \lambda_i \mathbf{w}_i \mathbf{y}_t + \delta_i \lambda_i y_{i,t-1} + e_{i,t} \\
\underline{\Delta} y_{i,t} =& \beta_{i,0} + \beta_{i,1} \underline{\Delta} x_{i,t} - \lambda_i (\mathbf{w}_i \mathbf{y}_t - \gamma_i \mathbf{w}_i \mathbf{x}_t) \quad (16) \\
& + \lambda_i \eta_i (x_{i,t-1} + \delta_i / \eta_i y_{i,t-1}) + e_{i,t}
\end{aligned}
$$

Equation (16) is an ECM in a combined spatial and temporal dimension. It has one cointegrating relationship between the spatial lags of $\mathbf{x}$ and $\mathbf{y}$. Analogously to a temporal ECM, λ_i defines the spatial equilibrium effect, or the spatial cointegration vector. The second term encompassed by $\lambda_i \eta_i$, that is $(x_{i,t-1} + \delta_i / \eta_i y_{i,t-1})$ refers to the temporal cointegration relationship. Next, we derive the conditions necessary for a pair of time and space equilibria to exist.

In the temporal equilibrium the value of the variables x and y is constant across the time dimension, such that:

$$
\begin{aligned}
y_{i,t} = y_{i,t-1} = y_i^* & \qquad x_{i,t} = x_{i,t-1} = x_i^* \quad (17) \\
\mathbf{w}_i \mathbf{y}_t = \mathbf{w}_i \mathbf{y}_{t-1} = \mathbf{w}_i \mathbf{y}^* & \qquad \mathbf{w}_i \mathbf{x}_t = \mathbf{w}_i \mathbf{x}_{t-1} = \mathbf{w}_i \mathbf{x}^* \quad (18)
\end{aligned}
$$

Using the temporal equilibrium conditions and equation (16) yields:

$$
\begin{aligned}
y_i^* =& \beta_{i,0} + \beta_{i,1} x_i^* - \beta_{i,1} \mathbf{w}_i \mathbf{y}^*, + \lambda_i \mathbf{w}_i \mathbf{y}^*, - \lambda_i \mathbf{w}_i \mathbf{y}^* \quad (19) \\
& + \lambda_i \eta_i x_i^* + \lambda_i \gamma y_i^* + \mathbf{w}_i \mathbf{y}^* \\
y_i^* =& \frac{\beta_{i,0}}{1 - \lambda_i \delta_i} + \frac{\beta_{i,1} + \lambda_i \eta_i}{1 - \lambda_i \delta_i} x_i^* + \frac{\lambda_i \gamma_i - \beta_{i,1}}{1 - \lambda_i \delta_i} \mathbf{w}_i \mathbf{x}^* + \frac{(1 - \lambda_i)}{1 - \lambda_i \delta_i} \mathbf{w}_i \mathbf{y}^* \quad (20) \\
=& \frac{\beta_{i,0}}{\phi_i} + \kappa_i x_i^* + \omega_i \mathbf{w}_i \mathbf{x}^* + \mu_i \mathbf{w}_i \mathbf{y}^* \quad (21)
\end{aligned}
$$

Analogous to the spatial equilibrium, we have the following coefficients under spatio-temporal equilibrium:

$$
\begin{aligned}
\phi_i = 1 - \lambda_i \delta_i & \qquad \kappa_i = \frac{\beta_{i,1} + \lambda_i \eta_i}{\phi_i} \quad (22) \\
\mu_i = \frac{1 - \lambda_i}{\phi_i} & \qquad \omega_i = \frac{\lambda_i \gamma_i - \beta_{i,1}}{\phi_i} \quad (23)
\end{aligned}
$$

Here, μ_i and ω_i capture the effect of the spatial lag and κ_i the effects of the explanatory variable in the spatio-temporal equilibrium. Equations (22) and

(23) imply:

$$\kappa_i \phi_i - \lambda_i \eta_i = \omega_i \phi_i - \lambda_i \delta_i \tag{24}$$

$$\phi_i \left(\kappa_i + \omega_i\right) = \lambda_i \left(\eta_i + \gamma_i\right) \tag{25}$$

Plugging the equilibrium coefficients into equation (16) gives us the spatio-temporal ECM:

$$\begin{aligned}
\underline{\Delta} y_{i,t} =& \beta_{i,0} + \beta_{i,1} \underline{\Delta} x_{i,t} + \lambda_i \left(\gamma_i \mathbf{w}_i \mathbf{x}_t - \mathbf{w}_i \mathbf{y}_t\right) && (26)\\
&+ \kappa_i \phi_i x_{i,t-1} - \beta_{i,1} y_{i,t-1} + \left(1 - \phi_i\right) y_{i,t-1} + e_{i,t} &&\\
=& \beta_{i,0} + \beta_{i,1} \underline{\Delta} x_{i,t} + \lambda_i \left(\gamma_i \mathbf{w}_i \mathbf{x}_t - \mathbf{w}_i \mathbf{y}_t\right) && (27)\\
&+ \phi_i \left(\kappa x_{i,t-1} - y_{i,t-1}\right) &&\\
&+ y_{i,t-1} - \beta_{i,1} x_{i,t-1} + e_{i,t} &&\\
=& \beta_{i,0} + \beta_{i,1} \underline{\Delta} x_{i,t} - \lambda_i \left(\mathbf{w}_i \mathbf{y}_t - \gamma_i \mathbf{w}_i \mathbf{x}_t\right) && (28)\\
&+ \beta_{i,1} \Delta x_{i,t} - \phi_i \left(y_{i,t-1} - \kappa_i x_{i,t-1}\right) &&\\
&+ y_{i,t-1} - \beta_{i,1} x_{i,t-1} + e_{i,t} &&\\
y_{i,t} - \mathbf{w}_i \mathbf{y}_t - y_{i,t-1} =& \beta_{i,0} + \beta_{i,1} \left(x_{i,t} - \mathbf{w}_i \mathbf{x}_t - x_{i,t-1}\right) && (29)\\
&- \lambda_i \left(\mathbf{w}_i \mathbf{y}_t - \gamma_i \mathbf{w}_i \mathbf{x}_t\right) &&\\
&- \phi_i \left(y_{i,t-1} - \kappa x_{i,t-1}\right) + e_{i,t} &&
\end{aligned}$$

Equation (29), the spatio-temporal ECM, is a central contribution of this paper. It is new to the literature and expresses precisely the nature of spatio-temporal short run dynamics and partial adjustment to the spatial and temporal long run equilibria. The short run effect is $\beta_{i,1}$, ϕ_i is the speed of error correction or the partial adjustment to the temporal long run equilibrium, and λ_i is the partial adjustment to the spatial long run equilibrium.

However, the term on the left hand side and the term capturing the spatial dynamics are not very informative. To provide better interpretation, we define joint spatio-temporal differencing as $y_{i,t} - y_{i,t-1} - \mathbf{w}_i \mathbf{y}_t + \mathbf{w}_i \mathbf{y}_{t-1} = \Delta \underline{\Delta} y_{i,t}$ and the equivalent for $\Delta \underline{\Delta} x_{i,t}$. The $\Delta \underline{\Delta}$ notation takes out first order nonstationarity across the two dimensions, space and time. It is equivalent in time series to transforming an I(1) process into a stationary I(0) by taking first differences across time. Here, the joint differencing is interpreted as temporal first difference of spatial difference, or vice versa.

Using this notation and adding on both sides $\mathbf{w}_i \mathbf{y}_t$ and $\mathbf{w}_i \mathbf{x}_t$ transforms equation (29) into:

$$\begin{aligned}
\Delta \underline{\Delta} y_{i,t} = & \beta_{i,0} + \beta_{i,1} \Delta \underline{\Delta} x_{i,t} + \left(\mathbf{w}_i \mathbf{y}_{t-1} - \beta_{i,1} \mathbf{w}_i \mathbf{x}_{t-1}\right) \\
& - \lambda_i \left(\mathbf{w}_i \mathbf{y}_t - \gamma_i \mathbf{w}_i \mathbf{x}_t\right) - \phi_i \left(y_{i,t-1} - \kappa_i x_{i,t-1}\right) + e_{i,t},
\end{aligned} \tag{30}$$

$\lambda_i \left(\mathbf{w}_i \mathbf{y}_t - \gamma_i \mathbf{w}_i \mathbf{x}_t\right)$ represents the spatial and $\phi_i \left(y_{i,t-1} - \kappa_i x_{i,t-1}\right)$ the temporal error correction term. However both terms are still potentially nonstationary with respect to the other dimension. We can rewrite the temporal long run relationship as

$$y_{i,t-1} - \kappa_i x_{i,t-1} = \underline{\Delta} y_{i,t-1} - \kappa_i \underline{\Delta} x_{i,t-1} + \mathbf{w}_i \mathbf{y}_{t-1} - \kappa_i \mathbf{w}_i \mathbf{x}_{t-1} \tag{31}$$

and the spatial long run relationship as

$$\mathbf{w}_i \mathbf{y}_t - \gamma_i \mathbf{w}_i \mathbf{x}_t = \mathbf{w}_i \Delta \mathbf{y}_t - \gamma_i \mathbf{w}_i \Delta \mathbf{x}_t + \mathbf{w}_i \mathbf{y}_{t-1} - \gamma_i \mathbf{w}_i \mathbf{x}_{t-1} \tag{32}$$

Then (30) is transformed as:

$$\begin{aligned}
\Delta\underline{\Delta} y_{i,t} =& \beta_{i,0} + \beta_{i,1} \Delta\underline{\Delta} x_{i,t} + \left(\mathbf{w}_i \mathbf{y}_{t-1} - \beta_{i,1} \mathbf{w}_i \mathbf{x}_{t-1}\right) & (33)\\
& - \lambda_i \left(\mathbf{w}_i \Delta \mathbf{y}_t - \gamma_i \mathbf{w}_i \Delta \mathbf{x}_t + \mathbf{w}_i \mathbf{y}_{t-1} - \gamma_i \mathbf{w}_i \mathbf{x}_{t-1}\right) \\
& - \phi_i \left(\underline{\Delta} y_{i,t-1} - \kappa_i \underline{\Delta} x_{i,t-1} + \mathbf{w}_i \mathbf{y}_{t-1} - \kappa_i \mathbf{w}_i \mathbf{x}_{t-1}\right) + e_{i,t} \\
=& \beta_{i,0} + \beta_{i,1} \Delta\underline{\Delta} x_{i,t} - \phi_i \left(\underline{\Delta} y_{i,t-1} - \kappa_i \underline{\Delta} x_{i,t-1}\right) \\
& - \lambda_i \left(\mathbf{w}_i \Delta \mathbf{y}_t - \gamma_i \mathbf{w}_i \Delta \mathbf{x}_t\right) \\
& + \mathbf{w}_i \mathbf{y}_{t-1} \left(1 - \lambda_i - \phi_i\right) - \mathbf{w}_i \mathbf{x}_{t-1} \left(\beta_{1,i} - \lambda_i \gamma_i - \phi_i \kappa_i\right) + e_{i,t} \\
=& \beta_{i,0} + \beta_{i,1} \Delta\underline{\Delta} x_{i,t} - \phi_i \left(\underline{\Delta} y_{i,t-1} - \kappa_i \underline{\Delta} x_{i,t-1}\right) \\
& - \lambda_i \left(\mathbf{w}_i \Delta \mathbf{y}_t - \gamma_i \mathbf{w}_i \Delta \mathbf{x}_t\right) & (34)\\
& + \alpha_i \rho_i \mathbf{w}_i \mathbf{y}_{t-1} + \left(\pi_i - \beta_{1,i} - \beta_{2,i}\right) \mathbf{w}_i \mathbf{x}_{t-1} + e_{i,t}
\end{aligned}$$

The advantage of equation (34) is that there are two distinct error correction terms: one capturing the temporal error correction $\phi_i \left(\underline{\Delta} y_{i,t-1} - \kappa_i \underline{\Delta} x_{i,t-1}\right)$ and the other capturing the spatial error correction $\lambda_i \left(\mathbf{w}_i \Delta \mathbf{y}_t - \gamma_i \mathbf{w}_i \Delta \mathbf{x}_t\right)$. Note that both these terms are stationary across one dimension and lagged (or spatially lagged) along the other. Then, if there is cointegration, both error correction terms are stationary across the two dimensions.

The final two terms $\alpha_i \rho_i \mathbf{w}_i \mathbf{y}_{t-1} + \left(\pi_i - \beta_{1,i} - \beta_{2,i}\right) \mathbf{w}_i \mathbf{x}_{t-1}$ have a strong dependence interpretation. Specifically as we show in the next section, these can be encompassed by common correlated effects of $\mathbf{y_{t-1}}$ and $\mathbf{x_{t-1}}$. This motivates the use of cross-section averages to capture spatial strong dependence, which in turn is justified in large (N, T) samples by common correlated effects (Pesaran, 2006). Hence, in some contexts, it may be appropriate to replace $\left(\mathbf{w}_i \Delta \mathbf{y}_t - \gamma_i \mathbf{w}_i \Delta \mathbf{x}_t\right)$ with $\left(\Delta \bar{\mathbf{y}}_t - \gamma_i \Delta \bar{\mathbf{x}}_t\right)$. Then, these terms can be interpreted as common correlated effects adjusted for strong spatial dependence. However, choice of the appropriate long run weights is typically context specific. We will discuss this issue in our application to UK house prices.

3 Spatio-Temporal ECM, Common Correlated Effects and Weak Dependence

The spatio-temporal error correction model in equation (34) highlights the importance of three dynamic processes: (a) spatio-temporal, short run dynamics $\beta_{i,1} \Delta\underline{\Delta} x_{i,t}$; and partial adjustment to two equilibrium relationships (b) temporal error correction $\phi_i\left(\underline{\Delta} y_{i,t-1}-\kappa_i \underline{\Delta} x_{i,t-1}\right)$ and (c) spatial error correction $\lambda_i\left(\mathbf{w}_i \Delta \mathbf{y}_t-\gamma_i \mathbf{w}_i \Delta \mathbf{x}_t\right)$. If both spatial and temporal equilibria exist, all the three corresponding regressors $\Delta \underline{\Delta} x_{i,t}$, $\left(\underline{\Delta} y_{i,t-1}-\kappa_i \underline{\Delta} x_{i,t-1}\right)$ and $\left(\mathbf{w}_i \Delta \mathbf{y}_t-\gamma_i \mathbf{w}_i \Delta \mathbf{x}_t\right)$, as well as the dependent variable $\Delta \underline{\Delta} y_{i,t}$, must all be stationary in the temporal domain and granular (weakly dependent) in the cross-sectional (spatial) domain. Naturally, achieving this requires adequate choice of spatial weights for spatial differencing ($\underline{\Delta}$) and spatial lagging/averaging ($\mathbf{w}_i$). The recent literatures on spatial strong and weak dependence and on estimation of weak dependence spatial weights matrices provide valuable insights in making these choices.

Pesaran (2006) showed that the common correlated effects (CCE) estimator achieves weak dependence by approximating factors underlying strong cross-sectional dependence using cross-sectional averages, $\bar{y}_t=\frac{1}{N} \sum_{i=1}^N y_{i,t}$ and $\bar{x}_t=\frac{1}{N} \sum_{i=1}^N x_{i,t}$. In a spatial equilibrium each cross-sectional unit has the same influence on all other units, so the spatial weights become $w_{i,j}=\frac{1}{N}\ \ \forall i, j \in N$ and $i \neq j$. Therefore the spatial lags can be rewritten as $\mathbf{w}_i \mathbf{y}_t+1/N y_{i,t}=1/N \sum_{i=1}^N y_{i,t}=\bar{y}_t$, respectively in their first difference as $\Delta \bar{y}_t=1/N \sum_{i=1}^N \Delta y_{i,t}$ where $\bar{y}_t$ is a scalar and the same for all cross-sectional units. We are going to incorporate the cross-sectional averages into Equation (34) in two steps. In the first step we replace the spatial lags $\mathbf{w}_i \mathbf{y}_{t-1}$ with $\bar{y}_{t-1}-1/N y_{i,t-1}$ and $\mathbf{w}_i \mathbf{x}_{t-1}=\bar{x}_{t-1}-1/N x_{i,t-1}$ which then gives:

$$
\begin{aligned}
\Delta \underline{\Delta} y_{i,t}= & \beta_{i,0}+\beta_{i,1} \Delta \underline{\Delta} x_{i,t}-\phi_i\left(\underline{\Delta} y_{i,t-1}-\kappa_i \underline{\Delta} x_{i,t-1}\right) \\
& -\lambda_i\left(\mathbf{w}_i \Delta \mathbf{y}_t-\gamma_i \mathbf{w}_i \Delta \mathbf{x}_t\right) \\
& +\alpha_i \rho_i\left(\bar{y}_{t-1}-\frac{1}{N} y_{i,t-1}\right)+\left(\pi_i-\beta_{1,i}-\beta_{2,i}\right)\left(\bar{x}_{t-1}-\frac{1}{N} x_{i,t-1}\right)+e_{i,t} \\
\approx & \beta_{i,0}+\beta_{i,1} \Delta \underline{\Delta} x_{i,t}-\phi_i\left(\underline{\Delta} y_{i,t-1}-\kappa_i \underline{\Delta} x_{i,t-1}\right) \qquad (35) \\
& -\lambda_i\left(\mathbf{w}_i \Delta \mathbf{y}_t-\gamma_i \mathbf{w}_i \Delta \mathbf{x}_t\right) \\
& +\alpha_i \rho_i \bar{y}_{t-1}+\left(\pi_i-\beta_{1,i}-\beta_{2,i}\right) \bar{x}_{t-1}+e_{i,t}
\end{aligned}
$$

The final step follows for a sufficiently large number of cross sectional units because $\bar{y}_t \pm \frac{1}{N} y_{i,t} \approx \bar{y}_t$ and $\mathbf{w}_i \mathbf{y}_t \approx \bar{y}_t$ as $\lim_{N \rightarrow \infty} \frac{1}{N} y_{i,t}=0$. Implicitly equation (35) relies on several (potentially different) spatial weight matrices. The last term $\alpha_i \rho_i \bar{y}_{t-1}+\left(\pi_i-\beta_{1,i}-\beta_{2,i}\right) \bar{x}_{t-1}$ uses cross-sectional averages to take out strong cross-sectional dependence. Given the asymptotic validity of the Pesaran (2006) common correlated effects (CCE) estimator, this is a natural choice. This still leaves the choice of spatial weights to model the short run and long run relationships. Now, we demonstrate that cross-sectional averages are also a natural choice for modeling the spatial error correction term. Then, in a second step we can further replace the spatial lag of the first differences in the spatial error correction term:

$$
\begin{aligned}
\Delta\underline{\Delta}y_{i,t} = {} & \beta_{i,0} + \beta_{i,1}\Delta\underline{\Delta}x_{i,t} - \phi_i\left(\underline{\Delta}y_{i,t-1} - \kappa_i\underline{\Delta}x_{i,t-1}\right) \\
& - \lambda_i\left(\Delta\bar{y}_t - \frac{1}{N}\Delta y_{i,t} - \gamma_i\left[\Delta\bar{x}_t - \frac{1}{N}\Delta x_{i,t}\right]\right) \\
& + \alpha_i\rho_i(\overline{y}_{t-1} - \frac{1}{N}y_{i,t-1}) + \left(\pi_i - \beta_{1,i} - \beta_{2,i}\right)(\overline{x}_{t-1} - \frac{1}{N}x_{i,t-1}) + e_{i,t}
\end{aligned}
\tag{36}
$$

For $\lim_{N\to\infty}\frac{1}{N}\Delta y_{i,t} = 0$ and Equation (36) can be then simplified to:

$$
\begin{aligned}
\Delta\underline{\Delta}y_{i,t} \approx {} & \beta_{i,0} + \beta_{i,1}\Delta\underline{\Delta}x_{i,t} - \lambda_i\left(\Delta\bar{y}_t - \gamma_i\Delta\bar{x}_t\right) \\
& - \phi_i\left(\underline{\Delta}y_{i,t-1} - \kappa_i\underline{\Delta}x_{i,t-1}\right) + \alpha_i\rho_i\overline{y}_{t-1} + \left(\pi_i - \beta_{1,i} - \beta_{2,i}\right)\overline{x}_{t-1} + e_{i,t}
\end{aligned}
\tag{37}
$$

Equation (37) clearly emphasises the connection between a heterogeneous panel data factor model (Pesaran, 2006) and our error correction model equation (34). Suppose we were not concerned with potential spatial cointegration and the corresponding error correction, but were mainly interested in modelling the temporal long run relationship along the lines of Pesaran and Smith (1995) and Pesaran et al. (1999). One would still be concerned with potential cross-sectional strong dependence and include cross-sectional averages as common correlated effects to address this issue. In fact, Pesaran (2006) showed that this empirical strategy is adequate in large samples. Now, since the model includes both short run dynamics and partial adjustment to a temporal long run equilibrium, one would need to include four cross-sectional averages in the model: $\Delta\bar{y}_t$, $\Delta\bar{x}_t$, $\overline{y}_{t-1}$ and $\overline{x}_{t-1}$. These are precisely the four cross-section averages included in equation (37). There are two implications of this observation. First, this equation can be interpreted as a common correlated effects estimator. In fact, even better, while coefficients on the cross-sectional averages typically have no economic interpretation, in our spatio-temporal error correction model, they have very precise interpretations. Second, this justifies cross-section average weights as a natural choice for two out of the four spatial weights matrices in our spatio-temporal error correction model. The choice for the remaining two matrices remain open, both denoted by $\underline{\Delta}$ in equation (37). In our empirical application, we model both using a spatial weight matrix from cross-correlations.

As discussed above, equation (37) contains the contemporaneous values of the difference of the cross-sectional averages in the spatial error correction term and first lag of the cross-sectional averages in a separate term. Next, we will collect the cross-sectional averages which leads to a representation in the fashion of Pesaran (2006):

$$
\begin{aligned}
\Delta\underline{\Delta}y_{i,t} &= \beta_{i,0} + \beta_{i,1}\Delta\underline{\Delta}x_{i,t} - \phi_i\left(\underline{\Delta}y_{i,t-1} - \kappa_i\underline{\Delta}x_{i,t-1}\right) \\
&\quad - \lambda_i\Delta\bar{y}_t + \lambda_i\gamma_i\bar{x}_t + \lambda_i\overline{y}_{t-1} - \lambda_i\gamma_i\overline{x}_{t-1} + e_{i,t} \\
&= \beta_{i,0} + \beta_{i,1}\Delta\underline{\Delta}x_{i,t} - \phi_i\left(\underline{\Delta}y_{i,t-1} - \kappa_i\underline{\Delta}x_{i,t-1}\right) \\
&\quad + \gamma_{i,y,0}\bar{y}_t + \gamma_{i,y,1}\overline{y}_{t-1} + \gamma_{i,x,0}\bar{x}_t + \gamma_{i,x,1}\overline{x}_{t-1} + e_{i,t}
\end{aligned}
\tag{38}
$$

where $\gamma_{i,y,0} = -\lambda_i$, $\gamma_{i,y,1} = \lambda_i + \alpha_i \rho_i$, $\gamma_{i,x,0} = \lambda_i \gamma_i$ and $\gamma_{i,x,1} = \pi_i - \beta_{1,i} - \beta_{2,i} - \gamma_i \lambda_i$. In the literature the nuisance coefficient estimates on the cross-sectional averages are not interpreted. However, as shown here, they can provide valuable insights into the spatial cointegration relationship and partial adjustment. Furthermore the last equation shows how the CCE estimator (Pesaran, 2006) accounts for spatially integrated processes and implicitly models cointegration in a fashion similar to a temporal ECM.

Thus, cross-section average weights capture, for large N and T, the spatial long run relationship and partial adjustment to it. Often interest also rests on spatial modelling of the weak dependence part, included in the short run dynamics, that is, in spatial weights for modelling $\underline{\Delta} y_{i,t}$ and $\underline{\Delta} x_{i,t}$. Here, we can draw upon the recent spatial econometrics literature, which shows that an unrestricted weak dependence $\mathbf{W}$ is not identified in general (Bhattacharjee and Jensen-Butler, 2013). Then, weak dependence can be modelled using one of several estimators under alternate identifying assumptions: (a) symmetry (Bhattacharjee and Jensen-Butler, 2013); (b) sparsity (Ahrens and Bhattacharjee, 2015; Lam and Souza, 2019); (c) symmetry and sparsity (Bailey et al., 2016); (d) asymmetric hub-and-spokes network (Bhattacharjee and Holly, 2013); and (e) recursive ordering (Basak et al., 2018).

For the remainder there are two reasons why we will focus on equation (35) rather than (38). Firstly, we are interested in estimating the spatial long run cointegrating relationship λ_i and the spatial long run coefficient γ_i. Secondly assuming a spatial equilibrium is a strong assumption and using cross-sectional averages in the spatial long run vector can take out too much of the spatial dependence.

In equation (35) the temporal long run relationship $\phi_i \left(\underline{\Delta} y_{i,t-1} - \kappa_i \underline{\Delta} x_{i,t-1} \right)$ depends on the cross-correlation spatial weight matrix. The term crucially requires the spatial first differences to be stationary and using the cross-correlation spatial weights might not be sufficient to do so. Therefore Equation (35) can be reformulated so it retains the cross-correlation spatial weights for the short run ($\Delta \underline{\Delta} x_{i,t}$) and for the long run spatial relationship, but uses cross-section averages for the long run temporal relationship. To do so, we redefine the first spatial difference in the temporal long run term ($\underline{\Delta} y_{i,t-1} - \kappa_i \underline{\Delta} x_{i,t-1}$) as $\underline{\underline{\Delta}} y_{i,t} = y_{i,t} - \bar{y}_t$ and $\underline{\underline{\Delta}} x_{i,t} = x_{i,t} - \bar{x}_t$, which then gives us a reformulation of Equation (35):

$$
\begin{aligned}
\Delta \underline{\Delta} y_{i,t} = \beta_{i,0} + \beta_{i,1} \Delta \underline{\Delta} x_{i,t} - \phi_i \left(\underline{\underline{\Delta}} y_{i,t-1} - \kappa_i \underline{\underline{\Delta}} x_{i,t-1} \right) \\
- \lambda_i \left(\mathbf{w}_i \Delta \mathbf{y}_t - \gamma_i \mathbf{w}_i \Delta \mathbf{x}_t \right) + \alpha_i \rho_i \overline{y}_{t-1} + \left(\pi_i - \beta_{1,i} - \beta_{2,i} \right) \overline{x}_{t-1} + e_{i,t}
\end{aligned} \tag{39}
$$

Equation (39) models spatial dependence using two different spatial weight matrices. The spatial short run relationship and the spatial long run relationship depend on the spatial first difference using a spatial weights matrix based on multiple testing of cross-correlations. The temporal long run relationship is modelled using cross-sectional averages. Cross-sectional averages are equivalent to a spatial lag in which all elements of the spatial weights matrix are $1/N$

and the averages represent a spatial equilibrium. The final term $\alpha_i \rho_i \overline{y}_{t-1} + (\pi_i - \beta_{1,i} - \beta_{2,i})\,\overline{x}_{t-1}$ further accounts for strong cross-sectional dependence. In order to remove remaining strong cross-sectional dependence further lags of $\bar{y}_t$ and $\bar{x}_t$ can be added following Chudik and Pesaran (2015).

In our application, we explore these two alternatives for modelling the long run relationship: common correlated effects (Pesaran, 2006) and cross-section correlations (Bailey et al., 2016). For weak dependence, we employ the estimator proposed in Bailey et al. (2016) based on multiple testing of cross-section correlations, under assumptions of sparsity and symmetry.

Hence, we propose estimation in two steps. In the first step, we estimate a simple model regressing y on x using standard panel models for potentially nonstationary data by including common correlated effects to account for strong cross-section dependence. After including sufficient temporal lags in this model to ensure weak dependence of the residuals, as evidenced using the Pesaran (2015) CD test, we estimate the weak dependence spatial weights by multiple testing of residual cross-correlations. We then construct $\Delta\underline{\Delta} y_{i,t}$ and $\Delta\underline{\Delta} x_{i,t}$. For the second step, we estimate the spatio-temporal ECM Equations (35) and (39) using the mean group estimator (Pesaran and Smith, 1995). Under homogeneity assumptions on the cointegrating relationships, one can also use the pooled mean group estimator of Pesaran et al. (1999).

4 An application to house prices in the UK

The interest that many social scientists have in housing reflects, among other things, the importance it has in household budgets, in the design of social policy and even in the behaviour of the macro economy. Big differences in the way in which housing and financial markets function around the world have profound effects on how output and inflation in the different countries respond to changes in short-term interest rates, as well as to external shocks to asset markets. An important aspect of the interaction between the housing market and the macro economy arises from the link to the labour market as, for example, differences in the level of house prices between regions within countries lowers labour mobility.

In this section, we develop an application to regional house prices in the UK based on our spatio-temporal ECM. In related work, Beenstock and Felsenstein (2010) estimated a spatial Durbin error model for house prices in Israel. While their focus lies in temporal cointegration, we explicitly model the spatial nonstationarity and thus a spatial long run equilibrium as well. In doing so, we allow for spatial short and long run processes using different spatial weight matrices.

4.1 Economic model

There is an extensive literature on the economics of housing and on the determination of house prices, yet many studies of house prices place more emphasis on demand compared to supply factors (Olsen, 1987). One reason for this is that fluctuations in house prices observed in many countries over time have the most immediate consequences for macroeconomic performance, reflecting factors on the demand side that trigger shifts along a very inelastic short run supply curve. However, if there is also an interest in the lower frequency movement of house prices, an analysis of how forces on the supply side impact upon house prices could be useful.

This is not to suggest that the theoretical literature has neglected supply side factors. The best known, and most elegant, models of the housing market derive the demand for housing from a well articulated utility maximising framework and allows the stock of housing to evolve in a similar manner to the practice in the modern literature on economic growth (Muth, 1976; Brueckner, 1981; Arnott et al., 1983, 1999; Glaeser et al., 2008; Glaeser and Gottlieb, 2009). Nevertheless, the housing stock is subject to a different process of construction and then refurbishment over the (extended) lifetime of the house. Here we do not focus on supply side factors.

On the demand side it is now standard to see the determination of house prices as the outcome of a market for the services of the housing stock and as an asset. A standard model of the demand for housing services includes permanent income, the real price of housing services and a set of other influences affecting changes in household formation such as demographic shifts. In equilibrium the real price of houses, p^h/p, is equal to the real price of household services, s, divided by the user cost of housing, c :

$$p^h/p = s/c.$$

Here, p is a general price index. Assume that alternative assets are taxed at the rate τ. c is then equal to the expected real after-tax rate of return on other assets with a similar degree of risk:

$$c = (r + \pi)(1 - \tau) - \pi^e,$$

where r is the risk-equivalent real interest rate on alternative assets and π^e is the expected rate of price inflation. Feldstein et al. (1978) assume that the alternative asset is some aggregate capital which can be financed by the issue of equity or the sale of bonds. The bonds are of an equivalent degree of risk to house ownership. Equity is riskier, so there is a market determined risk premium, ρ, on the holding of equity. In equilibrium the risk adjusted return on equity, ε, is equal to the return on bonds:

$$(1 - \tau)\varepsilon - \tau_c\pi - \rho = (r + \pi)(1 - \tau) - \pi^e$$

The return to equity is expressed as the dividend payout per unit of equity.

Another way of deriving the user cost of housing is to use the full intertemporal model of consumption in which in equilibrium the marginal rate of substitution between housing services and the flow of utility from consumption is:

$$\frac{u^h}{u^c} = \frac{p^h}{p}\{(1-\tau)(r+\pi) - \pi^e - \Delta(ph/p)^e$$

where $\Delta(ph/p)^e$ is the expected appreciation in the real price of houses.

The price of houses that satisfies the market for housing services and the asset market arbitrage condition is:

$$p^h/p = s/\{(1-\tau)(r+\pi) - \pi^e - \Delta(p/p)^e\} \tag{40}$$

The empirical model that can be derived from this form of analysis employs the device of proxying the unobservable real rental price of the flow of housing services, s, by the determinants of the demand for housing services, such as income and the housing stock. We take the above flexible model to data on UK house prices.

4.2 Data

We use quarterly panel data, from 1997q1 to 2016q4, across local authorities in England. House prices at the United Kingdom Land Registry which records all UK house transactions are available monthly at the local authority level for England and Wales (from January 1995), Scotland (from January 2004) and Northern Ireland (from January 2005). The average of the 3 months is used to construct the quarterly estimates (GOV.UK, 2020).

Data on gross disposable household income are from the Office of National Statistics Office of National Statistics (2020*c*). Quarterly estimates are obtained by quadratic interpolation from annual figures. Annual population figures are obtained from Office of National Statistics (2020*a*). Quarterly estimates are obtained by quadratic interpolation from annual figures. Real house prices are calculated by dividing by the CPI. CPI is the implicit deflator for consumer prices calculated as the ratio of current price consumer expenditure to constant price consumer expenditure (Office of National Statistics, 2020*b*).

We focus on a total of 326 English local authorities in our data set, made up of county councils, district councils, unitary authorities, metropolitan districts and London boroughs. We dropped the small local authority of Rutland which has a population of only about 30,000, so the sample for housing sales is very thin. The Isles of Scilly are included within Cornwall. The house price index for each local authority calculated from Land Registry data is an hedonic, mix adjusted index.

4.3 Discussion of Results

In this section we turn to a detailed discussion of the results. The steps we perform are outlined in more detail in the Appendix 6. First step estimation together with CCE cross-section averages produces residuals that satisfy the weak dependence condition based on the Pesaran (2015) CD test. We use these residuals to compute cross-sectional correlations, which are then used to conduct multiple testing to estimate the weak dependence spatial weights matrix, following the methodology in Bailey et al. (2016). This provides 238 non-zero elements in the spatial weights matrix, which is then row normalized using row sums of absolute values. The reported estimates in Table 1 are mean group panel estimates of a demand equation for real house prices across 324 local authorities. Estimation and inference is conducted in Stata using an extension of the `xtdcce2` command (Ditzen, 2018, 2019).

Column (1) of Table (1) is a standard panel data error correction model accounting for nonstationarity and possible cointegration in the temporal dimension; see, for example, Pesaran and Smith (1995). The evidence of cointegration is statistically significant but partial adjustment is weak and there is also substantial strong cross section dependence as evident from the CD test (Pesaran, 2015). Moreover, the long run relationship between real house prices and real incomes, γ, is rather high at 2.881. To correct for the strong cross sectional dependence we then used the common correlated effects mean group estimator in column (2) (Pesaran, 2006). This now eliminates the strong cross sectional dependence, and the exponent of cross sectional dependence, α falls to 0.565. There is significant cointegration and strong partial adjustment to a long run relationship between real house prices and real personal income of about 0.75. Thus far, these two columns represent what the current literature takes as best practice to estimating error correction models for panel data.

However, it is worthwhile examining the residuals of the model in more detail. At the moment the model ignores any possible overlap of house prices between local authority areas. An idiosyncratic shock to Manchester has no consequences for the behaviour of house prices in contiguous areas so it ignores all spatial effects, but most critically potential nonstationarity and cointegration across the spatial dimension (Holly et al., 2011). In Figures 1a - 2b we plot on a map of England[6] various features of the significant correlation coefficients after multiple testing (Bailey et al., 2016). Figure 1a plots the sum of the significant correlations for each local authority area. It does appear that there is a cluster of significant correlations around London and other large cities.[7] In Figures 1b and 2a we plot the negative and positive correlations and in Figure 2b the absolute sum of significant correlations. The plots suggest that there is a significant degree of spatial correlation between house prices in different local authorities that the results in columns (1) and (2) in Table 1 do not address.

We now use the results of sections 2 and 3 to integrate an explicit treatment of spatial as well as temporal effects into the model. This is the major contribution of this paper.

[6]Copyright for the shapefiles of England contains National Statistics data © Crown copyright and database right [2015]" and "Contains Ordnance Survey data © Crown copyright and database right [2015].

[7] These correlations may be picking up the commuting patterns around London and other major cities.

Column (3) shows estimates of a spatial error correction model which is an exact counterpart of the temporal error correction model in panel data settings shown in column (1). Here, the spatial difference of y, $\underline{\Delta}\mathbf{y}$, is a linear function of $\underline{\Delta}\mathbf{x}$ – the spatial short run dynamics – together with partial adjustment to a long run spatial equilibrium captured by the spatial weights of y and x, that is, $\mathbf{Wy}$ and $\mathbf{Wx}$. Remarkably, there is some indication of spatial cointegration as well, which provides some justification for the subject of this paper. Parallel to temporal cointegration, spatial cointegration here is interpreted as a (spatial) long run relationship between house prices and income whereby, for the *ith* local authority, if there is any disequilibrium between prices and income with its neighbours (as given by the spatial weights matrix), house prices in the *ith* local authority adjust to partially mitigate against this disequilibrium.

Here, the role of the chosen spatial weights matrix is critical, because partial adjustment is with regard to disequilibrium amongst the neighbours of the *ith* local authority. Since our cross correlation weights matrix may be in part be capturing largely commuting patterns within local labour markets (Figures 1a to 2b), spatial cointegration here can be interpreted as housing market forces eliminating opportunities for local arbitrage.

In addition to spatial cointegration, and as expected, the model has strong spatial dependence. This strong spatial dependence is not fully addressed by including common correlated effects in Column (4). This is because nonstationary temporal dynamics have not yet been modelled. In both models (3) and (4), the spatial long run effect is modelled using the estimated spatial cross-correlation weights matrix (Bailey et al., 2016). Despite some residual strong spatial dependence, simultaneous evidence of spatial and temporal cointegration justifies our spatio-temporal ECM, to which we turn next.

Column (5) reports estimates of our basic spatio-temporal ECM model from Equation (35), which includes partial adjustment to a temporal equilibrium and a spatial equilibrium, but no common correlated effects. As expected from columns (1) through (4), we find strong evidence of cointegration in both dimensions. However, spatial strong dependence is still present, as is evident from the CD test (Pesaran, 2015). Also, the final two common correlated effects terms in Equation (35) are not included, so this model is not entirely consistent with our theory. Hence, in column (6), we also include in our model cross-section averages of y_t, x_t, y_{t-1} and x_{t-1}. The evidence of cointegration across both the temporal and spatial dimensions persists. The Pesaran CD test (Pesaran, 2015) statistic is much reduced, such that the null hypothesis can be rejected at levels just above 5%. Despite a much lowered CD test, a further reduction of the CD test statistic would be favourable. One reason why the statistic is not smaller may be that stationarity is not adequately achieved by taking spatial differences using the cross-correlation spatial weights matrix (Bailey et al., 2016). This line of reasoning is also supported by the evidence from columns (3) and (4) where a pure spatial ECM does not remove in itself strong cross section dependence. An alternative that can be considered here is the spatial weights implied by cross section averages, as discussed in Sections 2 and 3. The main place to apply these common correlated effects weights would be in the temporal partial adjustment term which relies critically on the spatial first differences being granular (weakly dependent). Hence, in the final column (7) we follow equation (38) and we apply common correlated effects weights only to the temporal long run. This retains the cross correlation spatial weights in the spatio-temporal short run dynamics and spatial error correction term.

Taking the above model to the data, in column (7), weak cross section dependence can no longer be rejected at the 5% level. This is supported by an estimated exponent of cross-sectional dependence $\hat{\alpha}$ of 0.505 with a lower 95% confidence bound of 0.497. To check on the remaining properties of the residuals for column (7) we used the multiple testing approach again to obtain any remaining significant cross-correlations. We find that only a single cross-correlation remains significant. This suggests that we may have adequately modelled strong and weak spatial dependence.

Evidence of significant short run dynamics and cointegration in the spatial and temporal dimensions is retained. Thus, column (7) represents our preferred model. In the temporal domain, there is about 20% partial adjustment, per quarter, to a long run relationship between house prices and income, γ, with a value of 1.304. The spatial long run house price elasticity of income, κ, is 1.418, and taking standard errors into consideration, this is very close to the estimate of γ, suggesting that house prices equilibrate spatially and temporally in a similar way. The strong partial adjustment is very notable, λ is very close to one suggesting that house prices in the i-th local authority adjust almost entirely to prices in other neighbourhoods, as captured by the spatial weighting matrix.

Finally, we also explored a couple of other possible model specifications. First, we attempted to model the spatial long run relationship using the cross section averages based on Equation (37). However, as would be expected from the above discussion, this model does not fit equally well, implying that the price-income relationship at the local authority level does not adjust to the national average. Second, we also explored a traditional and popular weights matrix based on geographic contiguity between local authorities.This model does not fit as well as our preferred model in column (7), with substantial spillovers beyond first order contiguous neighbours, implying that the spatial organisation is more nuanced. Local labour market dynamics and commuting for work may explain spatial dynamics better than simple geography.

5 Conclusion

We develop spatial and spatio-temporal Engle-Granger representations that provide corresponding error correction models (ECMs) that are new to the literature. The spatio-temporal ECM includes partial adjustment to two equilibria, one temporal and the other spatial, together with short run dynamics based on a spatio-temporal difference. The above ECM highlights the distinct role of spatial strong and weak dependence in nonstationary dynamic models. In addition, the role of strong dependence is important,and it can be modelled using the CCE estimator of Pesaran (2006). Weak dependence can be estimated using the various estimators of the spatial weights matrix available in the literature, and these estimated weights matrices can also be useful in understanding spatial cointegration.

Applied to data on house prices and personal incomes across local authorities in England, our model and estimation provides new evidence and interpretation of nonstationary spatio-temporal dynamics and partial adjustment to multiple equilibria. Importantly, there is evidence of spatial cointegration where there is a (spatial) long run relationship between house prices and income. The partial adjustment to this spatial equilibrium is very local, and may be explained in part by local labour markets and commuting for work.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Time	Time	Spatial	Spatial	Spatio-Temporal	Spatio-Temporal	Spatio-Temporal
Dep. Var.	$\Delta p_{i,t}$	$\Delta p_{i,t}$	$\underline{\Delta} p_{i,t}$	$\underline{\Delta} p_{i,t}$	$\Delta\underline{\Delta} p_{i,t}$	$\Delta\underline{\Delta} p_{i,t}$	$\Delta\underline{\underline{\Delta}} p_{i,t}$
$\hat{\beta}_1$	0.666***	0.401***	1.042***	0.147***	0.531***	0.288***	0.338***
	(0.020)	(0.037)	(0.096)	(0.048)	(0.027)	(0.038)	(0.039)
$\hat{\phi}$	-0.079***	-0.195***			-0.090***	-0.227***	-0.199***
	(0.002)	(0.007)			(0.005)	(0.008)	(0.008)
$\hat{\gamma}$	2.881***	0.751***			1.756***	0.138	1.304***
	(0.110)	(0.169)			(0.214)	(0.121)	(0.514)
$\hat{\lambda}$			-1.158***	-1.063***	-1.101***	-0.912***	-0.999***
			(0.114)	(0.055)	(0.043)	(0.040)	(0.040)
$\hat{\kappa}$			0.447***	0.647***	0.018	0.763***	1.418***
			(0.116)	(0.233)	(0.266)	(0.248)	(0.289)
$\hat{\beta}_0$	-7.278***	0.246	-0.406	2.522	-0.022	1.098***	0.271
	(0.564)	(0.695)	(0.445)	(2.285)	(0.056)	(0.372)	(0.293)
CSA time lags ($L_0 - L_1$)	none	0 - 1	none	0 - 4	none	0 - 1	0 - 2
R^2_{MG}	0.155	0.616	0.881	0.989	0.285	0.647	0.652
CD	1009.226	0.421	655.818	31.173	619.434	1.924	1.485
p-val	0.000	0.674	0.000	0.000	0.000	0.054	0.138
CI up	0.934	0.572	0.804	0.690	0.681	0.514	0.514
$\hat{\alpha}$	0.928	0.565	0.793	0.680	0.667	0.506	0.505
CI low	0.923	0.557	0.783	0.670	0.653	0.497	0.497
N	324	324	324	324	324	324	324
T	78	78	79	75	78	78	77

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

Table 1: Mean Group Estimation results with $\pi = \frac{1}{N}\sum_{i=1}^{N}\pi_i$ and $\pi_i = (\beta_{i,1}, \phi_i, \gamma_i, \lambda_i, \kappa_i, \beta_{i,0})$. The estimated equations are:

Column (1) & (2): $\Delta p_{i,t} = \beta_{0,i} + \beta_{1,i}\Delta y_{i,t} - \phi_i\left(p_{i,t-1} + \gamma_i y_{i,t-1}\right) + \sum_{l=L_0}^{L_1}\left(\gamma_{i,l,p}\bar{p}_{t-l} + \gamma_{i,l,y}\bar{y}_{t-l}\right) + \epsilon_{i,t}$

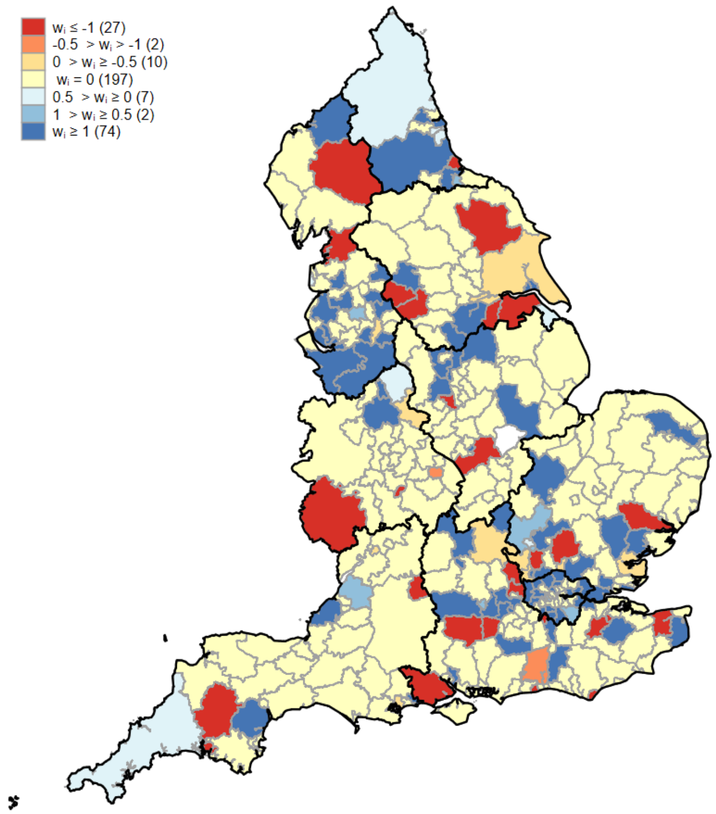
Column (3) & (4): $\underline{\Delta} p_{i,t} = \beta_{0,i} + \beta_{1,i}\underline{\Delta} y_{i,t} - \lambda_i\left(\mathbf{w}_i\Delta\mathbf{p}_t + \kappa_i\mathbf{w}_i\Delta\mathbf{y}_t\right) + \sum_{l=L_0}^{L_1}\left(\gamma_{i,l,p}\bar{p}_{t-l} + \gamma_{i,l,y}\bar{y}_{t-l}\right) + \epsilon_{i,t}$

Column (5) & (6): $\Delta\underline{\Delta} p_{i,t} = \beta_{0,i} + \beta_{1,i}\Delta\underline{\Delta} y_{i,t} - \phi\left(\underline{\Delta} p_{i,t-1} + \gamma_i\underline{\Delta} y_{i,t-1}\right) - \lambda_i\left(\mathbf{w}_i\Delta\mathbf{p}_t + \kappa_i\mathbf{w}_i\Delta\mathbf{y}_t\right) + \sum_{l=L_0}^{L_1}\left(\gamma_{i,l,p}\bar{p}_{t-l} + \gamma_{i,l,y}\bar{y}_{t-l}\right) + \epsilon_{i,t}$

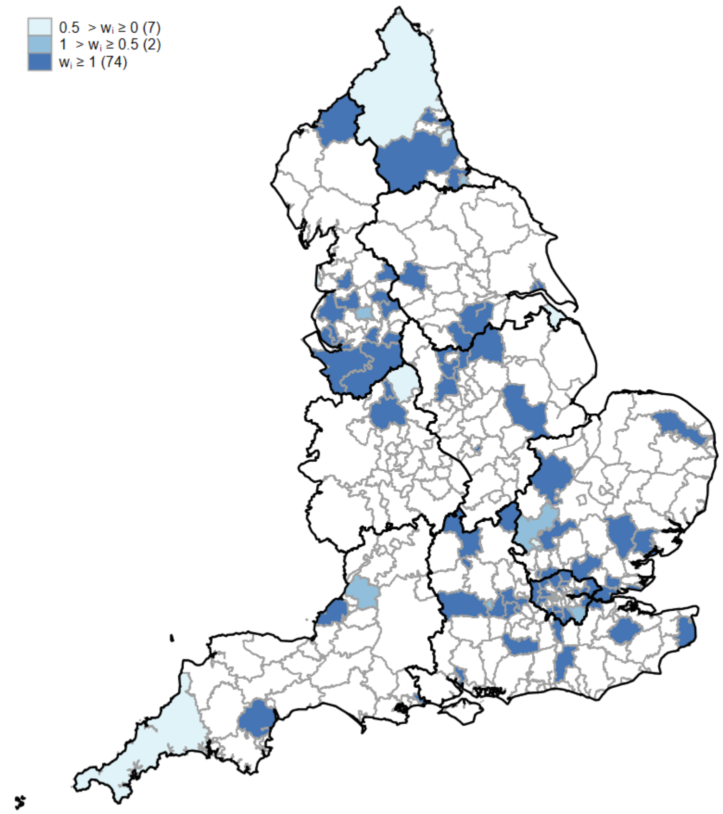
Column (7): $\Delta\underline{\underline{\Delta}} p_{i,t} = \beta_{0,i} + \beta_{1,i}\Delta\underline{\underline{\Delta}} y_{i,t} - \phi\left(\underline{\underline{\Delta}} p_{i,t-1} + \gamma_i\underline{\underline{\Delta}} y_{i,t-1}\right) - \lambda_i\left(\mathbf{w}_i\Delta\mathbf{p}_t + \kappa_i\mathbf{w}_i\Delta\mathbf{y}_t\right) + \sum_{l=L_0}^{L_1}\left(\gamma_{i,l,p}\bar{p}_{t-l} + \gamma_{i,l,y}\bar{y}_{t-l}\right) + \epsilon_{i,t}$

where $p_{i,t}$ are house prices and $y_{i,t}$ is income. Column (1) - (6) are based on Equation (35) and Column (7) on Equation (39).

CSA time lags is the structure of the lags of the cross-sectional averages, L_0 and L_1 defines the lower and upper bound of the lags of the cross-sectional averages. R^2_{MG} is the mean group adjusted R^2 from Holly et al. (2010). $\hat{\alpha}$ is an estimate of the exponent of cross-sectional dependence (Bailey et al., 2019). Confidence intervals for $\hat{\alpha}$ were calculated using 100 bootstrap repetitions. $\bar{y}_t$ and $\bar{p}_t$ are the cross-sectional averages, $\underline{\Delta}$ indicates the first difference in the spatial dimension using the cross-correlation weight matrix, $\underline{\underline{\Delta}}$ indicates the first difference in the spatial dimension using cross-sectional averages, Δ is the first lag in the time dimension. For the individual steps of the estimation see Section 3 and the Appendix.

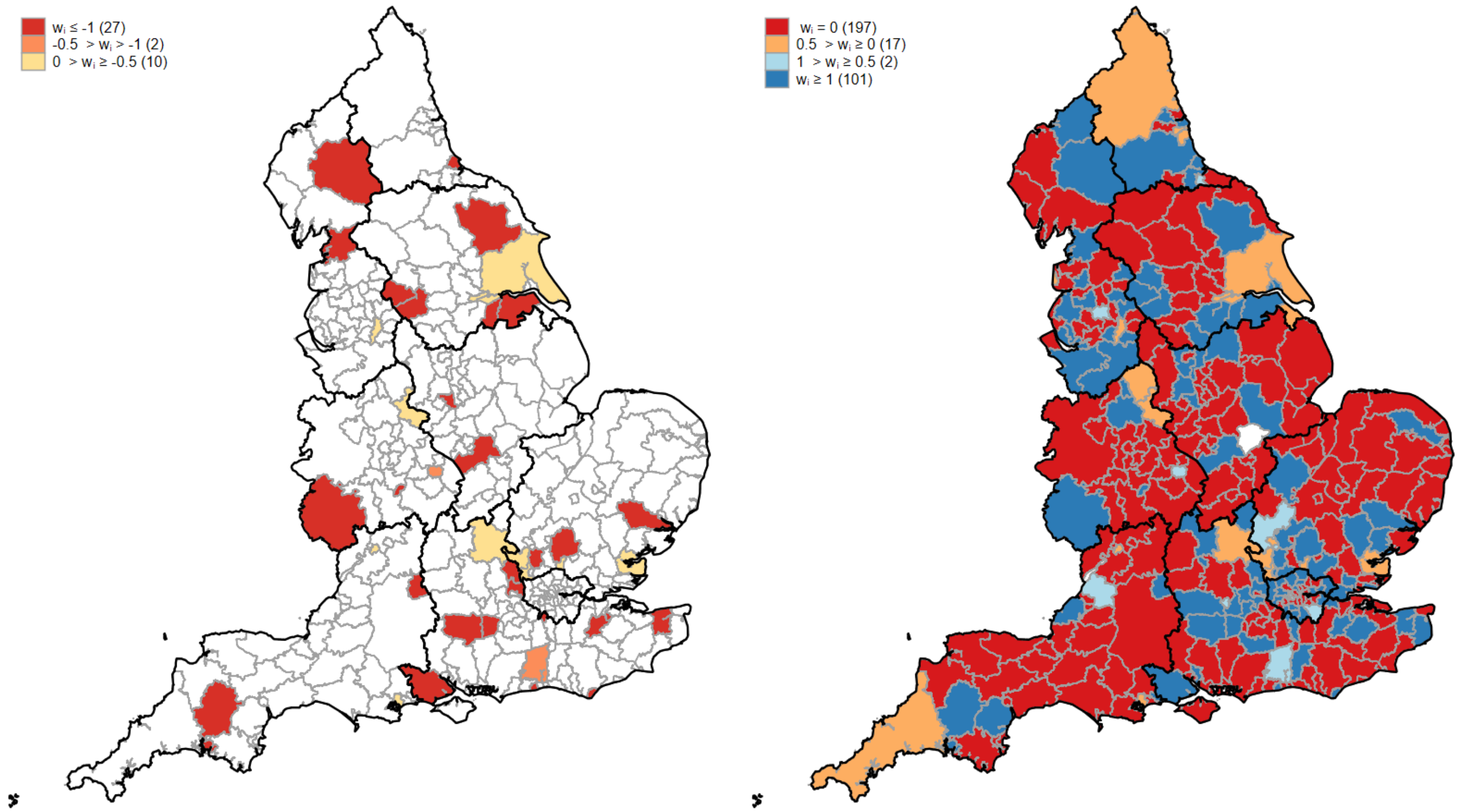


(a) All weights ($w_i = \sum_{j=1}^{N} w_{i,j}$)



(b) Weights larger than zero ($w_i = \sum_{j=1}^{N} w_{i,j}, w_{i,j} > 0$)

Figure 1: Sum of spatial weights of 324 local authorities in England. Spatial weights are generated from significant cross-correlations.



(a) Negative spatial weights ($w_i = \sum_{j=1}^{N} w_{i,j}, w_{i,j} < 0$)

(b) Absolute value of spatial weights ($w_i = \sum_{j=1}^{N} |w_{i,j}|$)

Figure 2: Sum of spatial weights of 324 local authorities in England. Spatial weights are generated from significant cross-correlations.

References

Ahrens, Achim and Arnab Bhattacharjee. 2015. "Two-step lasso estimation of the spatial weights matrix." *Econometrics* 3(1):128–155.

Arnott, Richard, Ralph Braid, Russell Davidson and David Pines. 1999. "A general equilibrium spatial model of housing quality and quantity." *Regional Science and Urban Economics* 29(3):283–316.

Arnott, Richard, Russell Davidson and David Pines. 1983. "Housing Quality, Maintenance and Rehabilitation." *Review of Economic Studies* 50(3):467–494.

Bai, Jushan. 2009. "Panel Data Models With Interactive Fixed Effects." *Econometrica* 77(4):1229–1279.

Bai, Jushan and Serena Ng. 2007. "Determining the number of primitive shocks in factor models." *Journal of Business and Economic Statistics* 25(1):52–60.

Bailey, Natalia, George Kapetanios and M. Hashem Pesaran. 2019. "Exponent of Cross-Sectional Dependence for Residuals." *Sankhya B. The Indian Journal of Statistics* forthcomin.

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

Basak, Gopal K., Arnab Bhattacharjee and Samarjit Das. 2018. "Causal ordering and inference on acyclic networks." *Empirical Economics* 55(1):213–232.

Beenstock, Michael and Daniel Felsenstein. 2010. "Spatial error correction and cointegration in nonstationary panel data: Regional house prices in Israel." *Journal of Geographical Systems* 12(2):189–206.

Bhattacharjee, Arnab and Chris Jensen-Butler. 2013. "Estimation of the spatial weights matrix under structural constraints." *Regional Science and Urban Economics* 43(4):617–634.

Bhattacharjee, Arnab and Sean Holly. 2013. "Understanding Interactions in Social Networks and Committees." *Spatial Economic Analysis* 8(1):23–53.

Brueckner, Jan K. 1981. "A dynamic model of housing production." *Journal of Urban Economics* 10(1):1–14.

Chudik, Alexander and M. Hashem Pesaran. 2015. "Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors." *Journal of Econometrics* 188(2):393–420.

Chudik, Alexander, M. Hashem Pesaran and Elisa Tosetti. 2011. "Weak and strong cross-section dependence and estimation of large panels." *The Econometrics Journal* 14(1):C45–C90.

Ditzen, Jan. 2018. "Estimating dynamic common-correlated effects in Stata." *The Stata Journal* 18(3):585 – 617.

Ditzen, Jan. 2019. "Estimating long run effects in models with cross-sectional dependence using xtdcce2." *CEERP Working Paper Series* No. 7.

Feldstein, Martin, Jerry Green and Eytan Sheshinski. 1978. "Inflation and Taxes in a Growing Economy with Debt and Equity Finance." *Journal of Political Economy* 86(2):S53–S70.

Glaeser, Edward L., Joseph Gyourko and Albert Saiz. 2008. "Housing supply and housing bubbles." *Journal of Urban Economics* 64(2):198 – 217.

Glaeser, Edward L. and Joshua D. Gottlieb. 2009. The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States. NBER Working Papers 14806 National Bureau of Economic Research, Inc.

GOV.UK. 2020. "UK House Price Index.".
URL: *Link*

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

Holly, Sean, M. Hashem Pesaran and Takashi Yamagata. 2011. "The spatial and temporal diffusion of house prices in the UK." *Journal of Urban Economics* 69(1):2–23.

Lam, Clifford and Pedro C.L. Souza. 2019. "Estimation and Selection of Spatial Weight Matrix in a Spatial Lag Model." *Journal of Business and Economic Statistics* 0(0):1–41.

Muth, R.J. 1976. A vintage model with housing production. In *Mathematical Land Use Theory*, ed. G.J. Papageorgiou. Lexingtom Books, D.C. Heath.

Office of National Statistics. 2020*a*. "Estimates of the population for the UK, England and Wales, Scotland and Northern Ireland.".
URL: *Link*

Office of National Statistics. 2020*b*. "GDP quarterly national accounts time series.".
URL: *Link*

Office of National Statistics. 2020*c*. "Regional gross disposable household income: all NUTS level regions.".
URL: *Link*

Olsen, E.O. 1987. The demand and supply of housing services: a critical survey of the empirical literature. In *Handbook of Regional and Urban Economics, Vol II*, ed. Edwin S. Mills. Amsterdam: North Holland pp. 989–1022.

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

Pesaran, M. Hashem. 2015. "Testing Weak Cross-Sectional Dependence in Large Panels." *Econometric Reviews* 34(6-10):1089–1117.

Pesaran, M. Hashem and Ron Smith. 1995. "Estimating long-run relationships from dynamic heterogeneous panels." *Journal of Econometrics* 68(1):79–113.

Pesaran, M. Hashem, Yongcheol Shin, and Ron P. Smith. 1999. "Pooled Mean Group Estimation of Dynamic Heterogeneous Panels." *Journal of the American Statistical Association* 94(446):621 –634.

6 Appendix: Estimation Steps

1. Estimate a simple 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 $\rho_{i,j} = \frac{1}{N} \sum_{t=1}^{T} \hat{\epsilon}_{i,t} \hat{\epsilon}_{j,t}$:

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

3. Use multiple testing to obtain significant cross-correlations with $\rho_{i,j} > c_p = \phi^{-1}\left(1 - \frac{p/2}{n^{\delta}}\right)$ which then gives W and row standardise W.

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

5. Calculate $\Delta\underline{\Delta} y_{i,t} = y_{i,t} - y_{i,t-1} - \mathbf{w}_i \mathbf{y}_t + \mathbf{w}_i \mathbf{y}_{t-1}$, $\underline{\Delta} y_{i,t} = y_{i,t} - \mathbf{w}_i \mathbf{y}_t$ and same for $\Delta\underline{\Delta} x_{i,t}$ and $\underline{\Delta} x_{i,t}$.

6. Estimate the following models:

$$\begin{aligned}\Delta\underline{\Delta} y_{i,t} &= \beta_{i,0} + \beta_{i,1} \Delta\underline{\Delta} x_{i,t} - \lambda_i \left(\mathbf{w}_i \Delta \mathbf{y}_t - \omega_i \mathbf{w}_i \Delta \mathbf{x}_t \right) \\ &\quad - \phi_i \left(\underline{\Delta} y_{i,t-1} - \kappa_i \underline{\Delta} x_{i,t-1} \right) \\ &\quad + \alpha_i \rho_i \overline{\mathbf{y}}_{t-1} + \left(\pi_i - \beta_{1,i} - \beta_{2,i} \right) \overline{\mathbf{x}}_{t-1} + e_{i,t}\end{aligned}$$

respectively for large N:

$$\begin{aligned}\Delta\underline{\Delta}y_{i,t} &= \beta_{i,0} + \beta_{i,1}\Delta\underline{\Delta}x_{i,t} - \lambda_i\left(\Delta\bar{y}_t - \tilde{\omega}_i\Delta\bar{x}_t\right) \\ &- \phi_i\left(\underline{\Delta}y_{i,t-1} - \kappa_i\underline{\Delta}x_{i,t-1}\right) \\ &+ \alpha_i\rho_i\overline{\mathbf{y}}_{t-1} + \left(\pi_i - \beta_{1,i} - \beta_{2,i}\right)\overline{\mathbf{x}}_{t-1} + e_{i,t}\end{aligned}$$

Further cross-sectional averages can be added to both regressions.