Spatial and Temporal Diffusion of House Prices in the UK

Sean Holly, M.Hashem Pesaran and Takashi Yamagata

December 2009

CWPE 0952

Spatial and Temporal Diffusion of House Prices in the UK*

Sean Holly*, M. Hashem Pesaran† and Takashi Yamagata‡

*Faculty of Economics and CIMF, University of Cambridge

†Faculty of Economics and CIMF, University of Cambridge and USC

‡Department of Economics and Related Studies, University of York

December 2009

Abstract

This paper provides a method for the analysis of the spatial and temporal diffusion of shocks in a dynamic system. We use changes in real house prices within the UK economy at the level of regions to illustrate its use. Adjustment to shocks involves both a region specific and a spatial effect. Shocks to a dominant region - London - are propagated contemporaneously and spatially to other regions. They in turn impact on other regions with a delay. We allow for lagged effects to echo back to the dominant region. London in turn is influenced by international developments through its link to New York and other financial centers. It is shown that New York house prices have a direct effect on London house prices. We analyse the effect of shocks using generalised spatio-temporal impulse responses. These highlight the diffusion of shocks both over time (as with the conventional impulse responses) and over space.

JEL Classification: C21, C23
Keywords: House Prices, Cross Sectional Dependence, Spatial Dependence.

*We are grateful to Alex Chudik, Chris Rogers, Ron Smith, Elisa Tosetti, and participants in the Cambridge Finance Workshop for helpful comments where a preliminary version of this paper was first presented.

1 Introduction

This paper provides a method for the analysis of the spatial and temporal diffusion of shocks in a dynamic system. We use changes in real house prices within the UK economy at the level of regions to illustrate its use. Adjustment to shocks involves both a region specific and a spatial effect. Shocks to a dominant region - London - are propagated contemporaneously and spatially to other regions. They in turn impact on other regions with a delay. We allow for lagged effects to echo back to the dominant region. London in turn is influenced by international developments through its link to New York and other financial centers. We analyse the effect of shocks using generalised spatio-temporal impulse responses. These highlight the diffusion of shocks both over time (as with the conventional impulse responses) and over space.

The present paper provides a relatively simple and consistent approach to modelling spatial and temporal adjustments quantitatively.[1] We approach the analysis from the perspective of recent developments in the literature on panel data models with a spatial dimension that manifests itself in the form of cross sectional dependence. One of the most important forms of cross section dependence arises from contemporaneous dependence across space and this is the primary focus of the spatial econometrics literature. This spatial dependence (Whittle, 1954) approach models correlations in the cross section by relating each cross section unit to its neighbour(s). Spatial autoregressive and spatial error component models are examples of such processes. (Cliff and Ord, 1973, Anselin, 1988, Kelejian and Robinson, 1995, Kelejian and Prucha, 1999, 2009, and Lee, 2004). Proximity, of course, does not have to be limited to proximity in space. Other measures of distance such as economic (Conley, 1999, Pesaran, Schuermann and Weiner, 2004), or social distance (Conley and Topa, 2002) could also be employed. In a regional context proximity of one region to another can depend on transport infrastructure. The ability to commute easily between two areas is likely to be a much better indication of economic inter-dependence than just physical closeness.[2]

Another approach to dealing with cross sectional dependence is to make use of multifactor error processes where the cross section dependence is characterized by a finite number of unobserved common factors, possibly attributable to economy-wide shocks that affect all units in the cross section, but with different intensities. With this approach the error term is a linear combination of a few common time-specific effects with heterogeneous factor loadings plus an idiosyncratic (individual-specific) error term. Pesaran (2006) has proposed an estimation method that consists of approximating the linear combinations of the unobserved factors by cross section averages of the dependent and explanatory variables and then running standard panel regressions augmented with the cross section averages. An advantage of this approach is that it yields consistent estimates even when the regressors are correlated with the factors, and the number of factors are unknown. A maximum likelihood procedure is also suggested by Bai (2009).

[1] The way in which space and time interact has long been a primary concern of epidemiologists and regional scientists. For a qualitative analysis of spatio-temporal processes using geographical information systems, see Peuquet (1994). However, there is also widespread interest in the environmental sciences as well. For a recent contribution see, for example, Kneib and Fahrmeir (2009). The study of the human brain through the use of magnetic resonance imaging also involves spatial-temporal modeling. See, for example, Gössl et al. (2001), and Fahrmeir and Gössl (2002).

[2] At the industry level interdependency between industries and firms are more likely to reflect patterns of intermediate input usage rather than physical proximity. See Horvath (1998, 2000) and Holly and Petrella (2008).

More recently Pesaran and Tosetti (2009) have sought to combine the insights of these two approaches and propose a panel model in which the errors are a combination of a multifactor structure and a spatial process. To achieve this a distinction is drawn between what is termed weak and strong cross section dependence. (Chudik, Pesaran and Tosetti, 2009). A process is said to be cross sectionally weakly dependent at a given point in time, if its weighted average at that time converges to its expectation in quadratic mean, as the cross section dimension is increased without bounds. If this condition does not hold, then the process is said to be cross sectionally strongly dependent. The distinctive feature of strong correlation is that it is pervasive, in the sense that it remains common to all units however large the number of cross sectional units. Significantly, spatial dependence typically entertained in the literature turns out to be weakly dependent in this framework. Holly, Pesaran and Yamagata (2009) model house prices at the level of US States where there is evidence of significant spatial dependence even when the strong form of cross sectional dependence has been swept up by the use of cross sectional averages. If we were to extend the sample by including regions or countries in Europe we would still expect that the spatial effects of New York State would be confined to its neighbouring states and not extend to Europe. By contrast common factors coming from the aggregate US economy could still have pervasive effects for regions of Europe.

As compared to purely spatial or purely factor models analysed in the literature, the spatio-temporal model estimated in this paper uses London house prices as the common factor and then models the remaining dependencies (contemporaneously or with a lag) conditional on London house prices. This allows us to consistently estimate separate conditional error correcting models for the different regions in the UK, which we then combine with a model for London to solve for a full set of spatio-temporal impulse response functions. Two alternative specifications are considered for London house prices, one specification that only depends on lagged London and neighbouring house price changes, and another which also depends on New York house prices.

While we are able to demonstrate that London is a dominant region for the rest of the UK, it is not immediately obvious why it should be uniquely so. One possibility we consider is that London is the largest city in Europe but more significantly is a major world financial centre. Developments in world financial markets can impact directly on the London housing market. Because London's traditional role as a financial and trading centre and the attraction that it has for economic migration of highly skilled workers, residential prices reflect both local factors in the UK but also movements abroad. In particular, there is a well established international market in residential property in which London along with New York plays a role. Our test results clearly show that New York house prices are significant drivers of house prices in the UK, but only through London. We also explored the possibility that Paris house prices could be one of the drivers of London house prices but found little evidence in its support.

It is important to note that the focus of our analysis differs from many others where the intention is to understand what determines regional house prices in terms of income, housing costs and other fixed factors to explain differences in regional house prices.[3] Although our approach does not preclude the inclusion of observable covariates such as incomes and interest rates we have focussed on the dispersion of house price shocks, conditioning on a dominant region (London) and neighbourhood effects, so the formulation is particularly parsimonious. It can be seen as a first step towards a more structural under-

[3] See, for example, Ashworth and Parker (1997), Cameron and Muellbauer (1998), Gallin (2006), and Holly et al. (2009).

standing of the inter-play of house price diffusion and the evolution of the real economy nationally and regionally.[4]

There have been a number of other studies that have considered the spatial diffusion of house prices. One of the first was Can (1990). He studied what he calls 'neighbourhood dynamics' by using a hedonic model of house prices where the price of a house depends on a series of characteristics, and incorporates both spatial spillover effects and spatial parametric drift. More recently Fingleton (2008) has developed a GMM estimator for a spatial house price model with spatial moving average errors. However, both of these studies confine themselves to the cross section dimension and do not consider the adjustment of prices over time. Studies of house prices that do consider both dimensions are van Dijk et al. (2007) and Holly et al. (2009). These studies develop a model that allows for stochastic trends, cointegration, cross-equation correlations and the latent-class clustering of regions. Dijk et al. apply their model to regional house prices in the Netherlands. They pick up a 'ripple' effect, by which shocks in one region are propagated to other regions. Holly et al. consider the evolution of real house prices and real disposable incomes across the 48 U.S. States and after allowing for unobserved common factors find statistically significant evidence of autoregressive spatial effects in the residuals of the cointegrating relations. Chudik and Pesaran (2009a) show that significant improvements in fit is achieved if Holly et al.'s regressions are augmented with spatially weighted cross sectional averages.

Conventional impulse response analysis traces out the effect of a shock over time. However, with a spatial dimension as well, dependence is both temporal and spatial (Whittle, 1954). Our results suggest that the effects of a shock decay more slowly along the geographical dimension as compared to the decay along the time dimension. For example, the effects of a shock to London on itself, die away and are largely dissipated after two years. By contrast the effects of the same shock on other regions takes much longer to dissipate, the further the region is from London. This finding is in line with other empirical evidence on the rate of spatial as compared to temporal decay discussed in Whittle (1956), giving the examples from variations of crop yields across agricultural plots, flood height and responses from population samples.

The rest of the paper is set out as follows: In Section 2 we propose a model of house price diffusion where we distinguish between the dominant and the non-dominant regions. In Section 3 we show how the individual models of regional house prices that have been treated separately for estimation purposes can be brought together and used for impulse response analysis along the time as well as the spatial dimensions. We also consider an extension of the basic model to allow for the effects of external shocks in the form of New York house prices. In Section 4 we report some empirical results using quarterly regional real house price data for the UK over the period 1974q1-2008q2. Finally, in Section 5 we

[4] Another branch of the literature explores the transmission of shocks in real estate markets including both residential and commercial property both within countries and across national borders. For example, Case et al. (2000) show that correlations between international real estate markets are high, given the degree to which they are segmented. They attribute a substantial amount of the correlation across world property markets to GDP which is correlated across countries. Herring and Wachter (1999) have pointed out that the 1997 Asian crisis was characterized by a collapse in real estate prices and a consequent weakening of the banking system before exchange rates came under attack. The essential link was that the real estate collapse impacted very negatively on the balance sheets of banks. Herring and Wachter point to a strong correlation between real estate cycles and banking crisis across a wide variety of countries. Bond, Dungey and Fry (2006) have also considered the transmission of real estate shocks during the East Asian crisis and the role they played in financial contagion.

draw some conclusions.

2 A Price Diffusion Model

Suppose we are interested in the diffusion of (log) prices, p_{it}, over time and regions indexed by $t = 1, 2, ..., T$ and $i = 0, 1, ..., N$ and we have *a priori* reason to believe that one of the regions, say region 0, is dominant in the sense that shocks to it propagate to other regions simultaneously and over time, whilst shocks to the remaining regions has little immediate impact on region 0, although we do not rule out lagged effects of shocks from regions $i = 1, 2, .., N$ to region 0. This is an example of a 'star' network within which each region can be viewed as a node. A first order linear error correction specification is given by[5]

$$\Delta p_{0t} = \phi_{0s}(p_{0,t-1} - \bar{p}^s_{0,t-1}) + a_0 + a_{01}\Delta p_{0,t-1} + b_{01}\Delta \bar{p}^s_{0,t-1} + \varepsilon_{0t} \tag{2.1}$$

and for the remaining regions

$$\begin{aligned}\Delta p_{it} &= \phi_{is}(p_{i,t-1} - \bar{p}^s_{i,t-1}) + \phi_{i0}(p_{i,t-1} - p_{0,t-1}) + a_i \\ &\quad + a_{i1}\Delta p_{i,t-1} + b_{i1}\Delta \bar{p}^s_{i,t-1} + c_{i0}\Delta p_{0t} + \varepsilon_{it},\end{aligned} \tag{2.2}$$

for $i = 1, 2, ..., N$. $\bar{p}^s_{it}$ denotes the spatial variable for region i defined by

$$\bar{p}^s_{it} = \sum_{j=0}^{N} s_{ij}p_{jt}, \text{ with } \sum_{j=0}^{N} s_{ij} = 1, \text{ for } i = 0, 1, ..., N. \tag{2.3}$$

The weights $s_{ij} \geq 0$ can be set *a priori*, either based on regional proximity or some economic measure of distance between regions i and j. In the empirical application we use a contiguity measure where s_{ij} is equal to $1/n_i$ if i and j share a border and zero otherwise, with n_i being the number of neighbors of i. $\bar{p}^s_{it}$ can be viewed as a local average price for region i. The weights s_{ij} can be arranged in the form of a spatial matrix, $\mathbf{S}$, which is row-standardized, namely $\mathbf{S}\boldsymbol{\tau}_{N+1}= \boldsymbol{\tau}_{N+1}$, where $\boldsymbol{\tau}_{N+1}$ is an $(N+1) \times 1$ vector of ones. Note also that when $\mathbf{S}$ captures contiguity measures its column norm will be bounded in N, namely that $\Sigma_{i=0}^{N}\ s_{ij} < K$, for all j, and as shown in Chudik and Pesaran (2009b), conditional on the dominant unit the remaining spatial dependence will be weak and conditional (on London) pair-wise dependence of the regions vanishes if N is sufficiently large.

The price equations are allowed to be error correcting, although whether they are is an empirical issue. In principle, the specification of the error correcting equations depends on the number of the cointegrating relations that might exist amongst the house prices across the N+1 regions. To avoid over-parametrization in the above specifications we have opted for relatively parsimonious specifications where London prices are assumed to be cointegrating with the average prices in the neighbourhood of London, $\bar{p}^s_{0t}$, whilst allowing for prices in other regions to cointegrate with London as well as with the neighbouring regions. In the case where prices cointegrate across all pairs of regions with the coefficients $(1, -1)$, it is easily seen that prices of each region must also cointegrate with prices of the neighbouring regions. It is interesting to note that the reverse is also true, namely if prices in each region cointegrate with prices of the neighbouring regions with the coefficients

[5] In the empirical application we allow for the possibility that the dynamics are of a higher order.

$(1,-1)$, then all price pairs would be cointegrating. Our error correcting specifications can therefore be justified as a parsimonious representation of pair-wise cointegration of prices across regions. The empirical validity of such an approximation for modelling of regional house prices in the UK will be investigated below.

Finally, note that the price change in the dominant region, Δp_{0t}, appears as a contemporaneous spatial effect for the i^{th} region. But there is no contemporaneous local average price included in the equation for Δp_{0t}. Implicit in the above specification is that conditional on the dominant region's price variable and lagged effects the shocks, ε_{it}, are approximately independently distributed across i. The assumption that $\Delta \bar{p}_{0t}$, is weakly exogenous in the equation Δp_{it}, $i = 1, 2, ..., N$ can be tested using the procedure advanced by Wu (1973), which can also be motivated using Hausman's (1978) type tests. Following Wu's approach denote the OLS residuals from the regression of the model for the dominant region by

$$\hat{\varepsilon}_{0t} = \Delta p_{0t} - \hat{\phi}_{0s}(p_{0,t-1} - \bar{p}^s_{0,t-1}) - \hat{a}_0 - \hat{a}_{01}\Delta p_{0,t-1} - \hat{b}_{01}\Delta \bar{p}^s_{0,t-1},$$

and run the auxiliary regression

$$\begin{aligned}\Delta p_{it} &= \phi_{is}(p_{i,t-1} - \bar{p}^s_{i,t-1}) + \phi_{i0}(p_{i,t-1} - p_{0,t-1}) \\ &\quad + a_i + a_{i1}\Delta p_{i,t-1} + b_{i1}\Delta \bar{p}^s_{i,t-1} + c_{i0}\Delta p_{0t} + \lambda_i \hat{\varepsilon}_{0t} + \eta_{it},\end{aligned} \tag{2.4}$$

and use a standard t-test to test the hypothesis that $\lambda_i = 0$ in this regression (for each i separately). This test is asymptotically equivalent to using Hausman's procedure which involves testing the statistical significance of the difference between the OLS and the IV estimates of $(a_i, a_{i1}, b_{i1}, c_{i0})$, using $\Delta p_{0,t-1}$ and $\Delta \bar{p}^s_{0,t-1}$ as instruments for Δp_{0t} in (2.2). It is clear that the test can only be computed if the instruments are not already included amongst the regressors of the model for the non-dominant regions. In our set up this is satisfied if $N > 1$. When $N = 1$, $\Delta \bar{p}^s_{0,t-1} = \Delta p_{1,t-1}$, and $\Delta \bar{p}^s_{1,t-1} = \Delta p_{0,t-1}$ and the model reduces to a bivariate VAR in Δp_{0t} and Δp_{1t}.

3 Spatio-temporal Impulse Response Functions

Although the regional price model can be de-coupled for estimation purposes, for simulation and forecasting the model represents a system of equations that needs to be solved simultaneously. We begin by writing the system of equations in (2.1) and (2.2) as

$$\Delta \mathbf{p}_t = \mathbf{a} + \mathbf{H}\mathbf{p}_{t-1} + (\mathbf{A}_1 + \mathbf{G}_1)\Delta \mathbf{p}_{t-1} + \mathbf{C}_0 \Delta \mathbf{p}_t + \boldsymbol{\varepsilon}_t, \tag{3.5}$$

where $\mathbf{p}_t = (p_{0t}, p_{1t}, ..., p_{Nt})'$, $\mathbf{a} = (a_0, a_1, ..., a_N)'$, $\boldsymbol{\varepsilon}_t = (\varepsilon_{0t}, \varepsilon_{1t}, ..., \varepsilon_{Nt})'$,

$$\mathbf{H} = \begin{pmatrix} \phi_{0s} & 0 & \cdots & 0 & 0 \\ -\phi_{10} & \phi_{1s} + \phi_{10} & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -\phi_{N-1,0} & 0 & \cdots & \phi_{N-1,s} + \phi_{N-1,0} & 0 \\ -\phi_{N0} & 0 & \cdots & 0 & \phi_{Ns} + \phi_{N0} \end{pmatrix} - \begin{pmatrix} \phi_{0s}\mathbf{s}'_0 \\ \phi_{1s}\mathbf{s}'_1 \\ \vdots \\ \phi_{N-1,s}\mathbf{s}'_{N-1} \\ \phi_{Ns}\mathbf{s}'_N \end{pmatrix}$$

$$\mathbf{A}_1 = \begin{pmatrix} a_{01} & 0 & \cdots & 0 & 0 \\ 0 & a_{11} & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{N-1,1} & 0 \\ 0 & 0 & \cdots & 0 & a_{N1} \end{pmatrix}, \quad \mathbf{G}_1 = \begin{pmatrix} b_{01}\mathbf{s}'_0 \\ b_{11}\mathbf{s}'_1 \\ \vdots \\ b_{N-1,1}\mathbf{s}'_{N-1} \\ b_{N1}\mathbf{s}'_N \end{pmatrix}, \quad \mathbf{C}_0 = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ c_{10} & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ c_{N-1,0} & 0 & \cdots & 0 & 0 \\ c_{N0} & 0 & \cdots & 0 & 0 \end{pmatrix},$$

where $\mathbf{s}_i' = (s_{i0}, s_{i1}, ..., s_{iN})$. Recall that $\mathbf{s}_i'\boldsymbol{\tau}_{N+1} = 1$. It is therefore, easily verified that $\mathbf{H}\boldsymbol{\tau}_{N+1} = \mathbf{0}$, and hence $\mathbf{H}$ is rank deficient. From this it also follows that one or more elements of the price vector, $\mathbf{p}_t$, must have a unit root.

Solving for price changes we have

$$\begin{aligned}\Delta\mathbf{p}_t &= (\mathbf{I}_{N+1} - \mathbf{C}_0)^{-1}\mathbf{a} + (\mathbf{I}_{N+1} - \mathbf{C}_0)^{-1}\boldsymbol{\Pi}\mathbf{p}_{t-1} \\ &\quad +(\mathbf{I}_{N+1} - \mathbf{C}_0)^{-1}(\mathbf{A}_1 + \mathbf{G}_1)\Delta\mathbf{p}_{t-1} + (\mathbf{I}_{N+1} - \mathbf{C}_0)^{-1}\boldsymbol{\varepsilon}_t,\end{aligned}$$

$$\Delta\mathbf{p}_t = \boldsymbol{\mu} + \boldsymbol{\Pi}\mathbf{p}_{t-1} + \boldsymbol{\Gamma}\Delta\mathbf{p}_{t-1} + \mathbf{R}\boldsymbol{\varepsilon}_t, \tag{3.6}$$

where

$$\begin{aligned}\boldsymbol{\mu} &= \mathbf{Ra}, \text{ with } \mathbf{R} = (\mathbf{I}_{N+1} - \mathbf{C}_0)^{-1}, \\ \boldsymbol{\Pi} &= \mathbf{RH}, \ \boldsymbol{\Gamma} = \mathbf{R}(\mathbf{A}_1 + \mathbf{G}_1).\end{aligned} \tag{3.7}$$

Since $\mathbf{R}$ is a non-singular matrix it follows that $\boldsymbol{\Pi}$ and $\mathbf{H}$ have the same rank, and $\boldsymbol{\Pi}$ will also be rank deficient ($\boldsymbol{\Pi}\boldsymbol{\tau}_{N+1} = \mathbf{RH}\boldsymbol{\tau}_{N+1} = \mathbf{0}$). Therefore, (3.6) represents a system of error correcting vector autoregressions in $\mathbf{p}_t$. The number of cointegrating or long run relations of the model depends on the rank of $\mathbf{H}$ (or $\boldsymbol{\Pi}$). The underlying price specifications in (2.1) and (2.2) imply the possibility of at most N cointegrating relations. For example, if all regional house prices were cointegrated with prices in London, then we will have N cointegrating relations. In what follows we denote the number of cointegrating relations by r $(0 \leq r \leq N)$, and write $\boldsymbol{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$ where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are $(N+1)\times r$ full column rank matrices. The choice of r is an empirical issue to which we shall return to below.

To examine the spatio-temporal nature of the dependencies implied by (3.6) we first write it as a vector autoregression (VAR)

$$\mathbf{p}_t = \boldsymbol{\mu} + \boldsymbol{\Phi}_1\mathbf{p}_{t-1} + \boldsymbol{\Phi}_2\mathbf{p}_{t-2} + \mathbf{R}\boldsymbol{\varepsilon}_t \tag{3.8}$$

where $\boldsymbol{\Phi}_1 = (\mathbf{I}_{N+1} + \boldsymbol{\alpha}\boldsymbol{\beta}' + \boldsymbol{\Gamma})$, and $\boldsymbol{\Phi}_2 = -\boldsymbol{\Gamma}$. The temporal dependence of house prices is captured by the coefficient matrices $\boldsymbol{\Phi}_1$ and $\boldsymbol{\Phi}_2$, and the spatial dependence by $\mathbf{R}$ and the error covariances, $Cov(\varepsilon_{it}, \varepsilon_{jt})$ for $i \neq j$. The temporal coefficients in $\boldsymbol{\Phi}_1$ and $\boldsymbol{\Phi}_2$ are also affected by the spatial patterns in the regional house prices as we have constrained the lagged effects and the error correction terms to match certain spatial patterns as characterized by the non zero values of s_{ij}. The above VAR model can now be used for forecasting or impulse response analysis.

For impulse response analysis it is important to distinguish between two types of counterfactuals. Assuming that the Wu test of the weak exogeneity of p_{0t} is not rejected, then it would be reasonable to assume that $Cov(\varepsilon_{0t}, \varepsilon_{it}) = 0$, for $i = 1, 2, ..., N$. In this case the impulse responses of a unit (one standard error) shock to house prices in the dominant region on the rest of regions at horizon h periods ahead will be given by

$$\begin{aligned}\mathfrak{g}_0(h) &= E(\mathbf{p}_{t+h} \,|\varepsilon_{0t} = \sqrt{\sigma_{00}}, \mathfrak{I}_{t-1}) - E(\mathbf{p}_{t+h} \,|\mathfrak{I}_{t-1}) \\ &= \sqrt{\sigma_{00}}\boldsymbol{\Psi}_h\mathbf{R}\ \mathbf{e}_0, \ \text{ for } h = 0, 1, ...,\end{aligned} \tag{3.9}$$

where $\mathfrak{I}_{t-1}$ is the information set at time $t-1$, $\sigma_{00} = Var(\varepsilon_{0t})$, $\mathbf{e}_0 = (1, 0, 0, ..., 0)'$, and

$$\boldsymbol{\Psi}_h = \boldsymbol{\Phi}_1\boldsymbol{\Psi}_{h-1} + \boldsymbol{\Phi}_2\boldsymbol{\Psi}_{h-2}, \text{for } h = 0, 1, ..., \tag{3.10}$$

with the initial values $\boldsymbol{\Psi}_0 = \mathbf{I}_{N+1}$, and $\boldsymbol{\Psi}_h = \mathbf{0}$, for $h < 0$.[6]

[6]See also Chapter 6 in Garratt et al. (2006).

To derive the impulse responses of a shock to non-dominant regions we need to allow for possible contemporaneous correlations across the regions i and j for $i, j = 1, 2, ..., N$. This can be achieved using the generalized impulse response function advanced in Pesaran and Shin (1998). In the present application we would have

$$\mathfrak{g}_i(h) = \frac{\mathbf{\Psi}_h \mathbf{R}\ \mathbf{\Sigma} \mathbf{e}_i}{\sqrt{\sigma_{ii}}}, \text{ for } h = 0, 1, ..., H \tag{3.11}$$

where $\mathbf{e}_i$ is an $(N+1) \times 1$ vector of zeros with the exception of its i^{th} element which is unity,

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma_{00} & 0 & 0 & \cdots & 0 & 0 \\ 0 & \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1,N-1} & \sigma_{1N} \\ 0 & \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2,N-1} & \sigma_{2N} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \sigma_{N-1,1} & \sigma_{N-1,2} & \cdots & \sigma_{N-1,N-1} & \sigma_{N,N-1} \\ 0 & \sigma_{N1} & \sigma_{N2} & \cdots & \sigma_{N-1,N} & \sigma_{NN} \end{pmatrix}, \tag{3.12}$$

where $\sigma_{ij} = E(\varepsilon_{it}\varepsilon_{jt})$. The elements of $\mathbf{\Sigma}$ can be estimated consistently from the OLS residuals, $\hat{\varepsilon}_{it}$ of the regressions for the individual regions, namely by, $\hat{\sigma}_{ij} = T^{-1}\Sigma_{t=1}^{T}\hat{\varepsilon}_{it}\hat{\varepsilon}_{jt}$, for $i, j = 1, 2, ..., N$, and $\hat{\sigma}_{00} = T^{-1}\Sigma_{t=1}^{T}\hat{\varepsilon}_{0t}\hat{\varepsilon}_{0t}$. With the above specification of $\mathbf{\Sigma}$, where its first row and column are restricted, $\mathfrak{g}_i(h) = \sigma_{ii}^{-1/2}\mathbf{\Psi}_h \mathbf{R}\ \mathbf{\Sigma}\mathbf{e}_i = \mathfrak{g}_0(h)$ for $i = 0$. So the generalized impulse response function (GIRF) defined by (3.11) with $\mathbf{\Sigma}$ given by (3.12) is applicable to the analysis of shocks to the dominant and non-dominant regions alike. The distinction between the two types of regions is captured by the zero-bordered form of $\mathbf{\Sigma}$. The recursive nature of the model with London as the dominant unit identifies the effects of shocks to London house prices. But due to the non-zero correlation of shocks across the remaining regions the effects of shocks to region $i \neq 0$ is not identified.

It is worth noting that

$$\mathbf{R} = (\mathbf{I}_{N+1} - \mathbf{C}_0)^{-1} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ c_{10} & 1 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ c_{N-1,0} & 0 & \cdots & 1 & 0 \\ c_{N0} & 0 & \cdots & 0 & 1 \end{pmatrix}. \tag{3.13}$$

Also since $s_{ii} = 0$, we have

$$\mathbf{A}_1 + \mathbf{G}_1 = \begin{pmatrix} a_{01} & b_{01}s_{01} & \cdots & b_{01}s_{0,N-1} & b_{01}s_{0N} \\ b_{11}s_{10} & a_{11} & \cdots & b_{11}s_{1,N-1} & b_{11}s_{1N} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ b_{N-1,1}s_{N-1,0} & b_{N-1,1}s_{N-1,1} & \cdots & a_{N-1,1} & b_{N-1,1}s_{N-1,N} \\ b_{N1}s_{N0} & b_{N1}s_{N1} & \cdots & b_{N1}s_{N,N-1} & a_{N1} \end{pmatrix}, \tag{3.14}$$

and recalling that $\Sigma_{j=0}^{N}s_{ij} = 1$, then $\left\|\mathbf{A}_1 + \mathbf{G}_1\right\|_r = \max_i(|a_{i1}|+|b_{i1}|) < K$, and $\left\|\mathbf{A}_1 + \mathbf{G}_1\right\|_c = \max_j\left(|a_{j1}| + \Sigma_{i=0}^{N}|b_{i1}|\, s_{ij}\right)$ which is also bounded in N since $\Sigma_{i=0}^{N}s_{ij} < K$, and $|a_{i1}|$ and $|b_{i1}|$ are bounded in N, by assumption.[7]

[7] $\left\|\mathbf{A}\right\|_r$ and $\left\|\mathbf{A}\right\|_c$ denote row and column matrix norm of $\mathbf{A}$.

In the context of the IVAR model discussed in Pesaran and Chudik (2009a,b), the spatio-temporal aspects of the underlying diffusion process are characterized by the coefficient matrices $\mathbf{R}$, $\mathbf{\Sigma}$, $\mathbf{\Phi}_1$ and $\mathbf{\Phi}_2$. In the present application, due to the dominance of region 0, we have

$$\|\mathbf{R}\|_c = 1 + \Sigma_{i=1}^{N} |c_{i0}| = O(N), \tag{3.15}$$

and the column norm of $\mathbf{R}$ is unbounded in N. As shown in Pesaran and Chudik the dominant region can be viewed as a common factor for the other regions, and conditional on p_{0t}, the cross dependence of p_{it} across $i = 1, 2, ..., N$ will be weak and not allowing for it at the estimation stage can only affect efficiency which will become asymptotically negligible when N is sufficiently large.

In cases where N is relatively small (as in the empirical application that follows), the estimation of the conditional models will be less efficient as compared to the full maximum likelihood estimation, treating all the equations of the model simultaneously. But such an estimation strategy is likely to be worthwhile only if N is really small, say 5 or 6. Even for moderate values of N the number of parameters to be estimated can be quite large compared to T (the time dimension) which could adversely affect the small sample properties of the maximum likelihood estimates.

Also, one can clearly generalize the model by introducing other (observed) national or international factors into the equation for the dominant region such as real disposable income (which need not be region specific), interest rates, or house prices in the US or Europe. In this paper we shall only consider the effects of New York house prices, and leave other extensions of the model to future research.

3.1 The Price Diffusion Model with Higher Order Lags

For higher lag order processes, the model of the dominant region is given by

$$\Delta p_{0t} = \phi_{0s}(p_{0,t-1} - \bar{p}^s_{0,t-1}) + a_0 + \sum_{\ell=1}^{k} a_{0\ell}\Delta p_{0,t-\ell} + \sum_{\ell=1}^{k} b_{0\ell}\Delta \bar{p}^s_{0,t-\ell} + \varepsilon_{0t} \tag{3.16}$$

and for the remaining regions by

$$\begin{aligned} \Delta p_{it} &= \phi_{is}(p_{i,t-1} - \bar{p}^s_{i,t-1}) + \phi_{i0}(p_{i,t-1} - p_{0,t-1}) \\ &\quad + a_i + \sum_{\ell=1}^{k} a_{i\ell}\Delta p_{i,t-\ell} + \sum_{\ell=1}^{k} b_{i\ell}\Delta \bar{p}^s_{i,t-\ell} + \sum_{\ell=0}^{k} c_{i\ell}\Delta p_{0,t-\ell} + \varepsilon_{it}, \end{aligned} \tag{3.17}$$

for $i = 1, 2, ..., N$. Combining (3.16) and (3.17), the full system of equations can be written as

$$\Delta \mathbf{p}_t = \mathbf{a} + \mathbf{H}\mathbf{p}_{t-1} + \sum_{\ell=1}^{k} (\mathbf{A}_\ell + \mathbf{G}_\ell)\Delta \mathbf{p}_{t-\ell} + \sum_{\ell=0}^{k} \mathbf{C}_\ell \Delta \mathbf{p}_{t-\ell} + \boldsymbol{\varepsilon}_t. \tag{3.18}$$

Solving for price changes we have

$$\Delta \mathbf{p}_t = \boldsymbol{\mu} + \mathbf{\Pi}\mathbf{p}_{t-1} + \sum_{\ell=1}^{k} \mathbf{\Gamma}_\ell \Delta \mathbf{p}_{t-\ell} + \mathbf{R}\boldsymbol{\varepsilon}_t, \tag{3.19}$$

where $\boldsymbol{\mu} = \mathbf{Ra}$ with $\mathbf{R} = (\mathbf{I}_{N+1} - \mathbf{C}_0)^{-1}, \mathbf{\Pi} = \mathbf{RH}$, as before, and

$$\mathbf{\Gamma}_\ell = \mathbf{R}(\mathbf{A}_\ell + \mathbf{G}_\ell + \mathbf{C}_\ell).$$

Writing (3.19) as a VAR in price levels

$$\mathbf{p}_t = \boldsymbol{\mu} + \sum_{\ell=1}^{k+1} \boldsymbol{\Phi}_\ell \mathbf{p}_{t-\ell} + \mathbf{R}\boldsymbol{\varepsilon}_t, \tag{3.20}$$

where $\boldsymbol{\Phi}_1 = \mathbf{I}_{N+1} + \boldsymbol{\Pi} + \boldsymbol{\Gamma}_1$, $\boldsymbol{\Phi}_\ell = \boldsymbol{\Gamma}_\ell - \boldsymbol{\Gamma}_{\ell-1}$ for $\ell = 2, ..., k$, and $\boldsymbol{\Phi}_{k+1} = -\boldsymbol{\Gamma}_k$.

The generalised impulse response function for the i^{th} region at horizon h is the same as before and is given by (3.11), except that $\boldsymbol{\Psi}_h$ must now be derived recursively using the following generalization of (3.10):

$$\boldsymbol{\Psi}_h = \sum_{\ell=1}^{k+1} \boldsymbol{\Phi}_\ell \boldsymbol{\Psi}_{h-\ell}, \text{ for } h = 0, 1, ..., \tag{3.21}$$

with the initial values $\boldsymbol{\Psi}_0 = \mathbf{I}_{N+1}$, and $\boldsymbol{\Psi}_h = \mathbf{0}$, for $h < 0$.

Note that the equations (3.16) and (3.17) could have different lag-orders for each terms for each region, k_{ia}, k_{ib}, k_{ic}, say. In the empirical section, we consider the case in which these lag-orders are selected by an information criteria (IC), such as Akaike IC (AIC) and Schwarz Bayesian Criterion (SBC). The heterogeneity of the lag orders across regions and variable type can be accommodated by defining $k = \max_i\{k_{ia}, k_{ib}, k_{ic}\}$, and then setting to zero all the lag coefficients that are shorter than k.

3.2 Bootstrap GIRF Confidence Intervals

As is well known the asymptotic standard errors of the estimated impulse responses are likely to be biased in small samples. There is also the possibility of non-Gaussian shocks that ought to be taken into account in small samples. For both of these reasons we compute bootstrapped confidence intervals for the estimates of $\mathfrak{g}_i(h)$, over h and i, to evaluate their statistical significance.

To this end denote the estimated model of (3.20) as $\mathbf{p}_t = \hat{\boldsymbol{\mu}} + \boldsymbol{\Sigma}_{\ell=1}^{k+1} \hat{\boldsymbol{\Phi}}_\ell \mathbf{p}_{t-\ell} + \hat{\mathbf{R}}\hat{\boldsymbol{\varepsilon}}_t$, and the estimated GIRF as $\widehat{\mathfrak{g}}_i(h) = \hat{\boldsymbol{\Psi}}_h \hat{\mathbf{R}}\ \hat{\boldsymbol{\Sigma}}\mathbf{e}_i / \sqrt{\hat{\sigma}_{ii}}$. We generate B bootstrap samples denoted by $\mathbf{p}_t^{(b)}$, $b = 1, 2, ..., B$, then compute bootstrap GIRF $\widehat{\mathfrak{g}}_i^{(b)}(h)$ for each $\mathbf{p}_t^{(b)}$. Firstly, the b^{th} bootstrap samples are obtained recursively based on the DGP

$$\mathbf{p}_t^{(b)} = \hat{\boldsymbol{\mu}} + \sum_{\ell=1}^{k+1} \hat{\boldsymbol{\Phi}}_\ell \mathbf{p}_{t-\ell}^{(b)} + \hat{\mathbf{R}}\hat{\boldsymbol{\varepsilon}}_t^{(b)}, \tag{3.22}$$

where $\boldsymbol{\varepsilon}_t^{(b)} = \hat{\boldsymbol{\Sigma}}^{1/2} \boldsymbol{\upsilon}_t^{*(b)}$, where the elements of $\boldsymbol{\upsilon}_t^{*(b)}$ are random draws from the transformed residual matrix, $\hat{\boldsymbol{\Sigma}}^{-1/2}(\hat{\boldsymbol{\varepsilon}}_1, \hat{\boldsymbol{\varepsilon}}_2, ..., \hat{\boldsymbol{\varepsilon}}_T)$, with replacement. The $k+1$ initial observations are equated to the original data.

Secondly using the obtained bootstrap samples $\mathbf{p}_t^{(b)}$, estimate the model (3.20), so that the b^{th} bootstrap GIRF is computed as

$$\widehat{\mathfrak{g}}_i^{(b)}(h) = \frac{\hat{\boldsymbol{\Psi}}_h^{(b)} \hat{\mathbf{R}}^{(b)}\ \hat{\boldsymbol{\Sigma}}^{(b)} \mathbf{e}_i}{\sqrt{\hat{\sigma}_{ii}^{(b)}}}, \ h = 0, 1, ..., H;\ i = 0, 1, 2, ..., N. \tag{3.23}$$

The $100(1-\alpha)\%$ confidence interval is obtained as $\alpha/2$ and $1-\alpha/2$ quantiles of $\widehat{\mathfrak{g}}_i^{(b)}(h)$ for each h and i.

The bootstrap samples are generated from the price equations selected at the estimation stage, namely with the same lag orders and restricted error correction terms.

3.3 The Model with New York House Prices

The baseline model can also be easily extended to include other observable effects such as international and domestic factors. Here we focus on the former and extend the model for London, (2.1), to include the New York house price changes, denoted by $\Delta p_{NY,t}$:

$$\begin{aligned}\Delta p_{0t} &= \phi_{0s}(p_{0,t-1} - \bar{p}^s_{0,t-1}) + \phi_{0NY}(p_{0,t-1} - p_{NY,t-1}) + a_0 \\ &\quad + a_{01}\Delta p_{0,t-1} + b_{01}\Delta \bar{p}^s_{0,t-1} + c_{NY,0}\Delta p_{NY,t} + \varepsilon_{0t}\end{aligned} \tag{3.1}$$

and the $\Delta p_{NY,t}$ is modelled as

$$\Delta p_{NY,t} = a_{NY} + a_{NY,1}\Delta p_{NY,t-1} + \varepsilon_{NY,t}. \tag{3.2}$$

It is assumed that ε_{0t} and $\varepsilon_{NY,t}$ are uncorrelated. This is not a restrictive assumption since $\Delta p_{NY,t}$ is already included in the London house price equation. As shown in Pesaran and Shin (1999), by including a sufficient number of lagged changes of $\Delta p_{NY,t}$ in (3.1) one can ensure that the errors from the two equations are uncorrelated.

The extended model, when combined with the price equations for other regions in the UK yields:

$$\Delta \mathbf{p}_t = \mathbf{a} + \mathbf{H}\mathbf{p}_{t-1} + (\mathbf{A}_1 + \mathbf{G}_1)\Delta \mathbf{p}_{t-1} + \mathbf{C}_0 \Delta \mathbf{p}_t + \boldsymbol{\varepsilon}_t, \tag{3.3}$$

where $\mathbf{p}_t = (p_{NY,t}, p_{0t}, ..., p_{Nt})'$, $\mathbf{a} = (a_{NY}, a_0, ..., a_N)'$, $\boldsymbol{\varepsilon}_t = (\varepsilon_{NY,t}, \varepsilon_{0t}, ..., \varepsilon_{Nt})'$,

$$\mathbf{H} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ -\phi_{0NY} & \phi_{0s} + \phi_{0NY} & 0 & \cdots & 0 & 0 \\ 0 & -\phi_{10} & \phi_{1s} + \phi_{10} & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & -\phi_{N-1,0} & 0 & \cdots & \phi_{N-1,s} + \phi_{N-1,0} & 0 \\ 0 & -\phi_{N0} & 0 & \cdots & 0 & \phi_{Ns} + \phi_{N0} \end{pmatrix} - \begin{pmatrix} 0 & \mathbf{0}'_{N+1} \\ 0 & \phi_{0s}\mathbf{s}'_0 \\ 0 & \phi_{1s}\mathbf{s}'_1 \\ \vdots & \vdots \\ 0 & \phi_{N-1,s}\mathbf{s}'_{N-1} \\ 0 & \phi_{Ns}\mathbf{s}'_N \end{pmatrix}$$

$$\mathbf{A}_1 = \begin{pmatrix} a_{NY,1} & 0 & \cdots & 0 & 0 \\ 0 & a_{01} & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{N-1,1} & 0 \\ 0 & 0 & \cdots & 0 & a_{N1} \end{pmatrix}, \mathbf{C}_0 = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ c_{NY,0} & 0 & 0 & \cdots & 0 & 0 \\ 0 & c_{10} & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & c_{N-1,0} & 0 & \cdots & 0 & 0 \\ 0 & c_{N0} & 0 & \cdots & 0 & 0 \end{pmatrix},$$

$$\mathbf{G}_1 = \begin{pmatrix} 0 & \mathbf{0}'_{N+1} \\ 0 & b_{01}\mathbf{s}'_0 \\ 0 & b_{11}\mathbf{s}'_1 \\ \vdots & \vdots \\ 0 & b_{N-1,1}\mathbf{s}'_{N-1} \\ 0 & b_{N1}\mathbf{s}'_N \end{pmatrix}.$$

Therefore, we have

$$\Delta \mathbf{p}_t = \boldsymbol{\mu} + \boldsymbol{\Pi}\mathbf{p}_{t-1} + \boldsymbol{\Gamma}\Delta \mathbf{p}_{t-1} + \mathbf{R}\boldsymbol{\varepsilon}_t, \tag{3.4}$$

where

$$\begin{aligned}\boldsymbol{\mu} &= \mathbf{Ra} \text{ with } \mathbf{R} = (\mathbf{I}_{N+2} - \mathbf{C}_0)^{-1}, \\ \boldsymbol{\Pi} &= \mathbf{RH}, \boldsymbol{\Gamma} = \mathbf{R}(\mathbf{A}_1 + \mathbf{G}_1).\end{aligned}$$

which as before leads to

$$\mathbf{p}_t = \boldsymbol{\mu} + \boldsymbol{\Phi}_1\mathbf{p}_{t-1} + \boldsymbol{\Phi}_2\mathbf{p}_{t-2} + \mathbf{R}\boldsymbol{\varepsilon}_t \tag{3.5}$$

As before the spatio-temporal properties of the model critically depend on the column norms of the coefficient matrices $\mathbf{R}$, $\boldsymbol{\Phi}_1$ and $\boldsymbol{\Phi}_2$. We saw that with London being the only dominant region only the first column of $\mathbf{R}$ was unbounded in N. For the present extended model we have

$$\mathbf{R} = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ c_{NY,0} & 1 & 0 & 0 & \cdots & 0 & 0 \\ c_{NY,0}c_{10} & c_{10} & 1 & 0 & \cdots & 0 & 0 \\ c_{NY,0}c_{20} & c_{20} & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ c_{NY,0}c_{N-1,0} & c_{N-1,0} & 0 & 0 & \cdots & 1 & 0 \\ c_{NY,0}c_{N0} & c_{N0} & 0 & 0 & \cdots & 0 & 1 \end{pmatrix},$$

which has two dominant columns. But since New York affects UK house prices only through London, there is only one dominant region as far as UK house prices are concerned, namely London. This is reflected in the columns of $\mathbf{R}$ since the non-zero elements of the second column are all proportional to the corresponding elements of the first column. Therefore, there is only one dominant unit driving UK house prices outside of London. Similar considerations also applies to the dynamic transmission of shocks.

The transmission of shocks through the idiosyncratic errors, ε_{it}, $i = NY, 0, 1, ..., N$, is characterised by the covariance matrix, $\boldsymbol{\Sigma}$:

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{NY} & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \sigma_{00} & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1,N-1} & \sigma_{1N} \\ 0 & 0 & \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2,N-1} & \sigma_{2N} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \sigma_{N-1,1} & \sigma_{N-1,2} & \cdots & \sigma_{N-1,N-1} & \sigma_{N,N-1} \\ 0 & 0 & \sigma_{N1} & \sigma_{N2} & \cdots & \sigma_{N-1,N} & \sigma_{NN} \end{pmatrix}.$$

where it is assumed that $\boldsymbol{\Sigma}$ has bounded matrix column (row) norm.

Finally, weak exogeneity of Δp_{0t} and the significance of the error correction terms in the Δp_{it} equations for $i = 1, 2, ..., N$ need to be tested as before.

4 Empirical Results

4.1 Regions and Their Connections

We apply the methodology described above to regional house prices (deflated by the general price level) in the UK using the quarterly mix-adjusted house price series collected by the Nationwide Building Society.[8] The panel data set covers quarterly real house price changes over the period 1973q4 to 2008q2 for London and 11 regions.[9]

[8] The mix adjustment of the house price index is intended to correct for price variations due to location and physical characteristics of the housing stock.

[9] See the Appendix A for data sources and other details.

The definition of regions used by the Nationwide (Table 1) differs in significant ways from the regional definitions used by the Office of National Statistics which are based on the Nomenclature of Territorial Units for Statistics (NUTS) of the European Union. The main differences arise with the definition of the North, the North West, East Anglia and the South East. The construction of the neighbourhood variables is described in Table 2. The general principle in construction was to use physically contiguous regions. However, that is not appropriate for London because the London Region is encircled first by the Outer Metropolitan Region which in turn is encircled by the Outer South East. In this case it may be inappropriate to rely solely on contiguity.

In general, contiguity can be a useful guide to determining the neighbours for each region. However, the relationships between house prices in different regions also interact with decisions to migrate and to commute. There are many ways in which information about house prices in different towns and regions are disseminated.[10] In this regard a major factor in this process is migration and commuting. However, as pointed out by Cameron and Muellbauer (1998) there are many barriers to mobility, especially because of large differences in the level of house prices in different regions (see also Barker, 2004). Commuting provides a substitute such that at the margin the relative price of a property of equivalent type (including an adjustment for different access to schools, countryside, etc.) at two different locations will depend on the time and cost of commuting. The extent of commuting within the UK can be seen from Table 3. This provides travel to work areas for 2001 obtained from the 2001 Census.[11] In terms of net flows the table reports the largest 10 areas by inflow and the largest 10 areas by outflow. London receives by far the largest inflow of workers each day with the large metropolitan cities of Manchester, Leeds, Glasgow and Birmingham next. In terms of outflows the three largest areas are Chelmsford and Braintree, Maidstone and North Kent, and Southend and Brentwood. Although some proportion of this could be to other areas, it is likely that the main destination is London. In Table 4 we identify a number of towns around London that have net outflows. All are connected to London via high speed rail[12], and all are in the Outer Metropolitan or Outer South East region. Of a total of 261,584 identified as commuting to outside of their town some 64% of commuters come from the Outer Metropolitan Region. By contrast commuting into the other large cities is almost exclusively from within the region in which the large cities are located. So because of the connections between areas and London provided by commuting, we used the OM and OSE regions as nearest neighbours to London.[13]

The neighbourhood connections in Table 2 with London connected to all regions forms a star network with London at its central hub. The extent of the interconnection of each region i can be measured by the degree of its centrality, c_i, defined as the number of regions with connections to region i divided by the total number of regions minus one.[14] The degree of centrality of the regional house price network is given in Table 5. Perhaps not surprisingly Scotland turns out to have the lowest degree of centrality (0.09) with

[10] For example, large chains of estate agents collate information on house prices across the country and allow the easy comparison of houses in different locations.

[11] These are provided by the Office of National Statistics for England and Wales and by the General Register Office for Scotland.

[12] For an empirical analysis of the effect on house prices of easy access to public transport (both rail and underground) see Gibbons and Machin (2005).

[13] The only other region that provides a significant number of commuters to London (approximately 16,000 according to the 2001 Census) is the South West. See Piggott (2007).

[14] See, for example, Chapter 2 in Goyal (2007).

London (by construction) the highest. Next most connected region is Outer South East (0.55), followed by East Midland, North West, and West Midland (each with $c_i = 0.45$). After Scotland the least connected region turns out to be East Anglia (0.18) which is nevertheless an important commuting region to London.

An alternative measure of connection for our purposes is commuting distances of the various regions from London. This measure is particularly relevant given the central role London house prices seem to play in the process of house price diffusion. Therefore, we consider ordering the regions (starting with London) according to their distance from London in the analysis of spatial impulse responses that follows. Regional distances from London are measured as geometrical average of distances from London to particular towns and cities in each region. For example, the distance to the North West is taken to be the geometrical average of the distance from London to Manchester, Liverpool and Lancaster.[15] Note that these distances are only used in our analysis to order the regions with respect to London. Local averages used in the regressions are measures of neighbourhood effects that are defined in terms of contiguity and not in terms of distances.

4.2 Cointegrating Properties of UK House Prices

The logarithm of real house prices and their quarterly rates of change across the 12 regions are displayed in Figure 1. There is a clear upward trend in real house prices over the 1974-2008 period, with prices in London and Outer Metropolitan areas rising faster than other regions. The bottom panel of the graph displays the considerable variations in house price changes that have taken place, both over time and across regions. It is also interesting to note that volatility of real house price changes (around 3.5% per quarter) are surprisingly similar across all the regions except for Scotland which is much lower at around 2.7% per quarter. The average rate of price increases in Scotland has been around 0.55% per quarter which is lower than the rate of increase of real house prices in London (at around 0.76% per quarter), but is in line with the rate of price rises in many other regions in the UK with much higher price volatilities.

The time series plots in Figure 1 suggest that the price series could be cointegrated across the regions, a topic which has attracted considerable attention in the literature. See, for example, Giussani and Hadjimatheou (1991), Alexander and Barrow (1994) and Ashworth and Parker (1997). Two related issues are also discussed in the literature. One concerns the possibility that house prices are convergent across regions, namely that shocks that move regional house prices apart are temporary (Holmes and Grimes, 2008). As we shall see below this requires that house price pairs are cotrending as well as cointegrating with coefficients that are equal but of opposite signs. The second issue relates to the so-called "ripple effect" hypothesis, under which shocks that originate in the London area and the South East fan out across the country, with the further away regions being the last to respond to the shock (Giussani and Hadjimatheou, 1991, Peterson et al., 2002).

Cointegration of regional house prices can be tested either jointly or pairwise. A joint test would involve setting up a VAR in all of the 12 regional house price series and then testing for cointegrating across all possible regions. This approach is likely to be statistically reliable only if the number of regions under consideration is relatively small, around 4-6, and the time series data available sufficiently long (say 120-150 quarters).

[15] Because of the concentric shape of the Outer Metropolitan and Outer South East regions we used a much larger number of towns and cities to compute the average distance.

The pairwise approach, developed in Pesaran (2007), can be used to test for cointegration either relative to a baseline (or an average) price level or for all possible pairs of prices. When applied to all possible pairs, the test outcome gives an estimate of the proportion of price pairs for which cointegration is not rejected, as well as providing evidence on possible clustering of cointegration outcomes. This approach is reliable when N and T are both sufficiently large. The full pairwise approach has been recently applied to UK house prices (over the period 1983q1-2008q4) by Abbott and de Vita (2009) who find no evidence of long-run convergence across regional house prices in the UK. Given the focus of our paper, namely the spatio-temporal nature of price diffusion, the issue of whether all regional house price pairs are cointegrating is of secondary importance. Clearly, our analysis can be carried out even if none of the house price pairs are cointegrated by simply setting the error correction coefficients (ϕ_{is} and ϕ_{i0}) in (2.1) and (2.2) to zero for all i. But it is important that cointegration is allowed for in cases where such evidence is found to be statistically significant.

With this in mind, and given our *a priori* maintained hypothesis that London can be taken as the dominant region, in the left panel of Table 6 we present trace statistics for testing cointegration between London and region i house prices, computed based on a bivariate VAR(4) specification in p_{0t} and p_{it} for $i = 1, 2, ..., 11$. The null hypothesis that London house prices are not cointegrated with house prices in other regions is rejected at the 10% significance level or less in all cases, with the exception of the Outer Metropolitan, Wales, North and Scotland. The test results for Wales and to a lesser extent for Outer Metropolitan and North are marginal. Only in the case of Scotland do we find no evidence of cointegration with London house prices.

Cointegration whilst necessary for long-run convergence of house prices is not sufficient. We also need to establish that house prices are cotrending and that the cointegrating vector corresponding to (p_{it}, p_{0t}) is $(1, -1)$. The joint hypothesis that p_{it} and p_{0t} are cotrending and their cointegrating vector can be represented by $(1, -1)$ is tested using the log-likelihood ratio statistic which is asymptotically distributed as a χ^2_2. To carry out these tests we use bootstrapped critical values (given at the foot of Table 6) since it is well known that the use of asymptotic critical values can lead to misleading inferences in small samples.[16] The joint hypothesis under consideration is rejected at the 10% level only in the case of Outer South East and East Anglia. Once again, these rejections are rather marginal and none are rejected at the 5% level. Together these test results support the error correction formulations (2.1) and (2.2). This does not, however, mean that all the error correction terms must be included in all the price equations. The evidence of pairwise cointegration simply suggests that one or more of the error correction coefficients should be statistically significant in one or both of the price equations in a given pair.

The estimates of the error correction coefficients and their associated t-ratios for each price pair involving London are summarized in Table 7. The right hand panel gives the estimates for the London equation and the left panel for the other regions. As can be seen none of the error correction terms is significant in the London equation, whilst they are significant in the equations for all other regions, with the exception of Outer Metropolitan, which seems to have very similar features to London. These results are also compatible with London viewed as the dominant region. Prices in London are long run forcing for all other regions (with the exception of the Outer Metropolitan), whilst London is not long run forced by any other region. The concept of long-run forcing or long-run causality was first introduced in Granger and Lin (1995) and later applied to

[16] See, for example, Garratt, Lee, Pesaran and Shin (2006, Section 9.3.1).

cointegrating models in Pesaran, Shin and Smith (2000). It is to be distinguished from the more familiar notion of "Granger causality" which does not even allow for short term feedbacks from the non-causal to the causal regions.

4.3 Estimates of Regional House Price Equations

The regression results for a model in which London acts as the dominant region and there is the possibility of error correcting towards London and towards neighbouring regions are summarized in Table 8. Price equations for individual regions are estimated by ordinary least squares (OLS) which yield consistent estimates under the weak exogeneity of the London prices. For the London region there is no error correction term and London house price changes are regressed on their lagged values and lagged values of average house price changes in neighbouring regions. For the other 11 regions similar regressions are estimated but with contemporaneous and lagged changes in London house prices also included as additional regressors along with the error correction terms. Local averages are constructed as simple averages of house price changes of the neighbouring regions as set out in Table 2. The lag orders for each region is selected separately using the Schwarz Bayesian criterion using a maximum lag order of 4. We also used an unrestricted model of 4 lags in all variables, as well as using the Akaike criterion to select the lag orders and found very similar results.[17]

Estimates for the error correction coefficients, ϕ_{i0} and ϕ_{is} in equation (3.17), are provided in columns 2 and 3 of Table 8. The estimates, $\hat{\phi}_{i0}$, refer to the error correction term $(p_{i,t-1} - p_{0,t-1})$ which capture the deviations of region i house prices from London, and $\hat{\phi}_{is}$ is associated with $(p_{i,t-1} - \bar{p}^s_{i,t-1})$, which gives the deviations of region i^{th} house prices from its neighbours. In line with the literature the evidence on convergence of house prices across the regions is mixed. Considering the error correction terms measured relative to London we find that it is statistically significant in five regions (East Anglia, East Midlands, West Midlands, South West and North). By contrast, the error correction term measured relative to neighbouring regions is statistically significant only in the price equation for Scotland, with none of the error correction terms being statistically significant in the remaining six regions, which include, perhaps not surprisingly, the dominant region, London and its surrounding regions Outer Metropolitan and Outer South East. We are left with the three regions, Wales, Yorkshire and Humberside, and North West, for which the absence of any statistically significant error correcting mechanism in their price equations is difficult to explain. A number of factors could be responsible for this outcome. The sample period might not be sufficiently informative in this regard, or these regions might have different error correcting properties that our parsimonious specification can not take into account.

We now turn to the short-term dynamics and spatial effects. To somewhat simplify the reporting of the estimates, in columns 4-6 of Table 8 we report the sum of the lagged coefficients, with the associated t-ratios provided in brackets. Column 7 reports the contemporaneous impact of London on other regions. The estimates show a considerable degree of heterogeneity in lag lengths and short term dynamics. Surprisingly, the own lag effects are rather weak and are generally statistically insignificant. Own lag effects are statistically significant only in the regressions for the North. In contrast, the lagged price changes from neighbouring regions are generally strong and statistically highly significant, clearly showing the importance of dynamic spill-over effects from the neighbouring

[17] These are available from the authors on request.

regions.

The contemporaneous effect of London house prices are sizeable and statistically significant in all regions. The size of this effect seems to be closely related to the commuting distance of the region from London. Wales seems to be an exception although the rather high contemporaneous effects of London house price changes on Wales is partly offset by the large negative effect of lagged London price changes. It is also noticeable that the direct effect of London on the Outer South East and Wales is greater than the impact on the Outer Metropolitan Region, even though this region is physically contiguous with London.[18]

To ensure that the above results are not subject to simultaneity bias, we used the Wu-Hausman statistic to test the null hypothesis that changes in London house prices are exogenous to the evolution of house prices in other regions. The test statistics, reported in the 8th column of Table 8, clearly show that the null can not be rejected. This finding is consistent with earlier evidence provided by Giussani and Hadjimatheou (1991) who use cross-correlation coefficients and Granger causality tests to show the existence of a ripple effect in house price changes starting in Greater London and spreading to the North. But note that there are statistically significant short-run feedbacks to London house prices from its neighbouring regions. Therefore, London house prices are "Granger caused" by its neighbouring house prices, although as noted above house prices in London's neighbouring regions are not long- run causal for London. Evidently, a dominant region could be affected by its neighbours in the short-run but not in the long-run. Short-run feedbacks could be the result of forward looking behaviour on the part of the neighbours, for example.

4.4 On the Choice of the Dominant Region

Thus far we have carried out our empirical analysis on the maintained assumption that London is the dominant region. The results provided so far are in fact compatible with this view. First, we have shown in Table 7 that London prices are long-run causal for all regions with the exception of the Outer Metropolitan region. House prices in none of the other regions are long-run causal for London. Also the hypothesis that contemporaneous changes in London prices are weakly exogenous for house prices in all other regions can not be rejected (as can be seen from Wu-Hausman statistics in Table 8). We have also noted that the presence of short-term feedbacks to London from the neighbouring regions is not incompatible with our maintained hypothesis.

However, it is also possible that there may be other forms of pair-wise dominance. Scotland could be dominant for the North, for example. To shed light on such possibilities, in Table 9 we allow each region in turn to be 'dominant' and then use the Wu-Hausman statistics to test the hypothesis that the assumed 'dominant' region is weakly exogenous for the other 11 regions. The first column of this table confirms the earlier results that London can be regarded as weakly exogenous for the other 11 regions. But the null hypothesis that other regions are weakly exogenous is rejected, in the case of at least 2 regions. For some regions that are relatively remote from London the number of rejections is much higher. It is also worth noting from the first row of Table 9 that prices changes in none of the regions can be regarded as weakly exogenous in price equation for London.

[18] An inspection of the ratio of London house prices to prices elsewhere suggests that since the mid 1990s house prices in the Outer Metropolitan region have declined relative to London, though a similar pattern is not apparent for other regions.

The weak exogeneity test results in Table 9 and the earlier test results of long-run causality in Table 7 provide strong evidence in favour of London being the dominant region, with the Outer metropolitan region being the second best candidate.[19]

4.5 A Common Factor Representation

As shown in Chudik and Pesaran (2009b) the VAR model with a dominant unit can also be viewed as a VAR with a dynamic factor. Therefore, the price diffusion model proposed in this paper is also a dynamic factor model where the common factor is observed and identified as the London price level. But in reality unobserved common factors could still be present in addition to the dominant unit(s). However, to test for such a possibility requires N to be sufficiently large which is not the case in our application. An alternative specification would be to assume a common factor model without a dominant unit where all regions are treated symmetrically and all price changes are related to the same unobserved common factors. Under such an approach the common factors are typically estimated as the principle components of the regional price series which again require N to be relatively large. It is not clear to us if such a strategy would be statistically reliable in our application where $N = 12$. There is also the additional difficulty of how to interpret the results based on common factors, since a pure common factor representation would surely destroy the spatial features that we have managed to identify and estimate within our set up.

4.6 Spatial-temporal Impulse Responses

The individual house price equations summarized in Table 8 present a rather complicated set of dynamic and interconnected relations, with the parameter estimates only providing a partial picture of the spatio-temporal nature of these relationships. For a fuller understanding we need to trace the time profile of shocks both over time and across regions. Conventional impulse response analysis traces out the effect of a shock over time where the time series under examination is influenced by past values of itself and possibly other variables. However, when we have a spatial dimension as well, dependence extends in both directions, spatially and temporally (Whittle, 1954).

In Figure 2 we plot generalised impulse responses of the effects of a positive unit shock (one standard error) to London house prices on the level of house prices in London as well as in the other 11 regions. Part A of this figure shows the point estimates of the effects of the shock on the level of house prices across all the regions, whilst part B displays the 90% bootstrapped error bounds for each region separately.

The positive shock to London house prices spills over to other regions gradually raising prices across the country. Generally the closer is the region to London the more rapid the response to the shock. Scotland in particular, but also the North and the North West take considerably longer to adjust to the shock. But the effects eventually converge across all the regions, albeit rather slowly in the case of some regions. The bootstrapped error bounds also support this conclusion.

The spatio-temporal effects of the London shock is better captured in Figure 3 where the same information as in Figure 2 is plotted in terms of the change rather than the level of house prices. In this figure the responses to the London shock are plotted along the two

[19] The other main index of regional house prices, the Halifax, publishes a series for Greater London that incorporates the Outer Metropolitan with London.

dimensions with the regions ordered by distance from London. The rate of decay in the time dimension is captured by the downhill direction going from right to left. Movement initially along the ridge going from left to right captures the spatial pattern.

Another view of the results is provided in Figure 4 where the impulse response functions for the price changes are plotted across regions (again ordered by their distance to London) for different horizons, $h = 0, 1, ...12$. Figure 4 is a contour of the GIRF and clearly shows the leveling off of the effect of the shocks over time and across regions. But the decay along the geographical dimension seems to be slower as compared to the decay along the time dimension. This important feature of the impulse response is best seen in Figure 5 where the effect of a unit shock to London house prices on London over time are directly compared to the impact effects of the same shock on regions ordered by their distance from London. Broken lines are bootstrap 90% confidence band of the GIRFs for the regions. The decay of the impact effects across regions is noticeably slower than the time decay of the shock on London. It is also interesting that the lower bound of the regional decay curve is systematically above the time decay curve for London, which suggests that the difference in the two rates of decay could be statistically significant.

4.7 Effects of New York House Prices

We have found that London house prices are weakly exogenous for prices in other regions, and long run causal. It is now interesting to ask if there are any exogenous drivers for London house price. Why should exogenous shocks to UK house prices originate in London? There are a number of possible reasons for this. London and its surrounding regions account for the largest concentration of income and wealth in the country. Macroeconomic and financial shocks are likely to have their first effects in London, due to the role that London has played historically as one of the world's financial centres. London's close links to New York as the pre-eminent financial centre in the global economy could also be an important channel through which global financial shocks can travel to the rest of the UK through London. With this in mind we also examined the possibility that house prices in New York have an impact on London house prices.

The NY house price series we use is based on the Bureau of Labor Statistics' definition of Metropolitan and non-metropolitan areas deflated by the New York consumer price index.[20] The quarterly, time series of London and New York real house prices are shown in Figure 6 allowing for the different scaling of the two series. There is a clear relationship between the two house price series, with London prices tending to follow New York price relatively closely. But, due to the non-stationary nature of real house price series one needs to consider such visual relationships with care.[21]

Since quarterly data on New York house price series are available only from end of 1975 onwards, we re-estimated all the regional price equations over the sample period 1976q1-2008q2 so that the estimates for the extended model are all based on the same sample information. The regression results are reported in Table 10. We first note that the change in the estimation sample from 1973q4-2008q2 to 1976q1-2008q2 has not affected the estimates of the regional house prices and the results for these regions (reported under

[20] The areas included in the calculation of the NY house price series are the New York-White Plains-Wayne Metropolitan Division comprises the counties of Bergen, NJ, Bronx, NY, Hudson, NJ, Kings, NY, New York, NY, Passaic, NJ, Putnam County, NY, Queens, NY, Richmond, NY, Rockland, NY, and Westchester, NY. For data sources and other details see Appendix A.

[21] For example, it is not possible to reject the hypothesis that the two price series contain unit roots.

"Other UK Regions" in Table 10) are very similar to those reported in Tables 8.

Turning to the London price equation, we now focus on the extent to which London prices are influenced by New York house prices. We first estimated an error correcting regression for London with the term, $p_{0,t-1} - p_{NY,t-1}$, included as one of the regressors. But did not find the error correction term to be statistically significant. We then considered if there were short-term feedbacks from New York to London, and run a regression of London house price changes on lagged price changes in London and New York plus the contemporaneous price change in New York. In all cases there is a statistically significant lagged effect of New York house prices on London even when we condition on lagged neighbourhood effects for London. It is also interesting that lagged effects of New York house price changes on London prices are quantitatively more important than the contemporaneous effect of New York house price changes on London.[22]

We also explored whether Paris as a major part of the market for real estate internationally, might affect London house prices. However, using data from 1991 on apartment prices for Paris, with the bordering departments of Hauts-de-Seine, Seine-Saint-Denis and Val-de-Marne as contiguous regions, we did not find any role for Paris house prices in explaining London house prices.

So far we have assumed that New York house price changes affect UK regional house prices only through London. But it might be argued that these effects could be more pervasive, possibly influencing all regions directly. To test this hypothesis in Table 11 we give the F statistics for testing the joint significance of the contemporaneous and lagged effects of New York house prices changes in all the UK regional price equations. In panel A of the table the statistics are computed conditioning on the contemporaneous changes in London house prices, whilst in panel B only contemporaneous changes in New York house prices are included. When we condition on contemporaneous London prices the test results are highly significant only in the London equation. When contemporaneous London prices are excluded, New York house prices become significant in London's neighbouring regions and East Anglia, but are statistically insignificant in other regions. These test results taken together clearly show that New York house prices are significant drivers of house prices in the UK only through London.

To close the system (for the computation of impulse responses) we estimated a pure autoregression in New York house price changes. The lag order was selected by SBC which turned out to be 3. The estimates are given in Table 10 and show that New York house price changes are highly persistent which is important for the way shocks transmit from New York to London and then to the rest of UK. The impulse responses of a positive unit shock to New York house prices on New York and London house price changes is given in Figure 7 and show the highly persistent effects of New York house price changes on London. Initially the effect of the shock is much more pronounced on New York, but after one quarter the effects of the New York shock are very similar for London and New York, with the effects persisting more in London than in New York, although the differences are not statistically significant.

For completeness, we also computed the impulse responses of the effects of a unit shock to London house prices on the UK regions in the extended model with New York House prices included in the London equation. The results are summarized in Figure 8. Not surprisingly these impulse responses are very similar to those given in Figure 2 for the baseline model without New York house prices. Note that in our model a

[22]We also considered including changes in the GBP/US$ rate in the London house price equation, but did not find to be statistically significant, largely due to its excessive volatility.

shock to London house prices does not feedback to New York house prices. As a result the differences in impulse responses of a shock to London in the two models (with and without New York prices) are only due to the differences in the parameter estimates across the two models. In the present application these differences are rather small, thus explaining the similarities of the impulse responses in Figures 2 and 8.

What is of greater interest is the impulse responses of a New York shock on the level of regional house prices in the UK. The results of this shock scenario is given in Figure 9. The impact effects of the New York shock on UK house prices is very small, since changes in New York house prices only affect London prices directly and mainly with a lag. But once London prices start to change, as the result of the New York shock, the effects begin to travel to the rest of the UK directly through the contemporaneous effects of London house prices on the rest of the UK as well as indirectly through the spatial inter-linkages. The outcome is very similar to the regional impulse responses given in Table 8, with the regions further away from London being initially less affected, although all regional house prices eventual convergence due to the dominant role that London plays in the diffusion of house prices in the UK.

5 Conclusions

This paper suggests a novel way to model the spatial and temporal dispersion of shocks in non-stationary dynamic systems. Using UK regional house prices we establish that London is a dominant region in the sense of Chudik and Pesaran (2009b) and moreover that it is long run forcing in the sense of Granger and Lin (1995). House prices within each region respond directly to a shock to London and in turn the shock is amplified both by the internal dynamics of each region and by interactions with contiguous regions. Using this approach we can track the diffusion of shocks using spatial-temporal impulse responses. Furthermore, we identify an independent role for shocks to London coming from developments in house prices in New York. These proxy the effect of global financial developments on house prices in London.

Modelling both the temporal (time series) and the spatial dimension at the same time means that we modify the conventional impulse response analysis. With a spatial dimension as well, dependence is both temporal and spatial (Whittle, 1954). The results then suggest that the effects of a shock decay more slowly along the geographical dimension as compared to the decay along the time dimension. When we shock London, the effects on London itself die away and are largely dissipated after two years. By contrast the effects of the shock to London on other regions takes much longer to dissipate, the further the region is from London. This finding is in line with other empirical evidence on the rate of spatial as compared to temporal decay discussed in Whittle (1956), giving the examples from variations of crop yields across agricultural plots, flood height and responses from population samples. A subject for further study is to see if this differential pattern of decay over time and space continues to prevail in other economic applications, not only in the case of house price diffusion, but also, for example, the diffusion of technological innovations using a suitable economic distance.

References

[1] Abbott, A. J. and G. DeVita (2009), Testing for long-run convergence across regional house prices in the UK: A pairwise approach. Mimeo, Oxford Brookes University.

[2] Anselin, L., (1988). *Spatial Econometrics: Methods and Models*, Boston: Kluwer Academic Publishers.

[3] Alexander, C. and Barrow, M. (1994). Seasonality and co-integration of regional house prices in the UK, *Urban Studies*, 31, 1667-1689.

[4] Ashworth, J. and S.C. Parker (1997). Modelling regional house prices in the UK. *Scottish Journal of Political Economy,* 44, 225-246.

[5] Bai, J., (2009). Panel data models with interactive fixed effects, *Econometrica*, 77, 1229-1279

[6] Barker, K. (2004). *Review of Housing Supply: Delivering Stability: Securing our future housing needs. Final Report - Recommendations.* www.barkerreview.org.uk. London: HM Treasury.

[7] van Dijk, B., Franses, P.H., Paap, R. and van Dijk, D (2007), Modeling Regional House Prices, *Tinbergen Institute,Erasmus University, Report* EI 2007-55.

[8] Bond, S.A., Dungey, M. and R. Fry (2006). A web of shocks: crises across Asian real estate markets. *Journal of Real Estate Finance and Economics,* 32, 253–274.

[9] Cameron, G. and J. Muellbauer, (1998). The housing market and regional commuting and migration choices. *Scottish Journal of Political Economy,* 45, 420-446.

[10] Can, A., (1990). The measurement of neighborhood dynamics in urban house prices. *Economic Geography* 66, 254-272

[11] Case, B., Goetzmann, W. and K.G. Rouwenhorst, (2000). Global real estate markets: Cycles and fundamentals. NBER Working Paper No. 7566.

[12] Chudik, A. and Pesaran, M. H., (2009a). Infinite dimensional VARs and factor models, *ECB working paper No. 998*, January 2009.

[13] Chudik, A. and Pesaran, M. H., (2009b). Econometric analysis of high dimensional VARs featuring a dominant unit, Cambridge University, under preparation.

[14] Chudik, A., Pesaran, M.H. and E. Tosetti, (2009). Weak and strong cross section dependence and estimation of large panels, CESifo Working Paper Series No. 2689.

[15] Cliff, A. and J.K. Ord, (1973). *Spatial Autocorrelation*, London: Pion.

[16] Conley, T.G., (1999), GMM Estimation with Cross Sectional Dependence, *Journal of Econometrics*, 92, 1-45.

[17] Conley, T.G. and G. Topa., (2002). Socio-economic Distance and Spatial Patterns in Unemployment, *Journal of Applied Econometrics*, 17, 303-327.

[18] Fahrmeir, L. and C. Gössl, (2002). Semiparametric Bayesian models for human brain mapping, *Statistical Modeling*, 2, 235-249.

[19] Fingleton B. (2008) Generalized Method of Moments estimator for a spatial model with Moving Average errors, with application to real estate prices, *Empirical Economics*, 34, 35-57.

[20] Gallin, J. (2006). The long-run relationship between house prices and income: Evidence from Local Housing Markets, *Real State Economics*, 34, 417–438.

[21] Garratt, A., K. Lee, M.H. Pesaran and Y. Shin, (2006). *Global and National Macroeconometric Modelling: A Long Run Structural Approach*, Oxford University Press, Oxford.

[22] Gibbons, S. and S. Machin, (2005). Valuing rail access using transport innovations. *Journal of Urban Economics,* 57, 148-169.

[23] Giussani, B. and G. Hadjimatheou, (1991). Modeling regional house prices in the U.K, *Journal of Financial Intermediation* 13, 414-435.

[24] Gössl, C., Auer, D. P. and L. Fahrmeir, (2001). Bayesian spatiotemporal inference in functional magnetic resonance imaging. *Biometrics*, 57, 554-562.

[25] Goyal, S., (2007), *Connections: An Introduction to the Economics of Networks*, Princeton University Press.

[26] Granger, C.W.J. and Lin, J.L. (1995). Causality in the Long Run, *Econometric Theory,* 11, 530-536.

[27] Hausman, J.A., (1978). Specification tests in econometrics. *Econometrica,* 46, 251-272.

[28] Herring, R. and S. Wachter, (1999). Real estate booms and banking busts: An international perspective. Wharton Working Paper Series 99-27.

[29] Holly, S., Pesaran, M.H. and T. Yamagata, (2009). A spatio-temporal model of house prices in the U.S.. *Journal of Econometrics,* forthcoming.

[30] Holly, S. and I. Petrella, (2008). Factor demand linkages and the business cycle: interpreting aggregate fluctuations as sectoral fluctuations. Cambridge Working Paper No. 0827.

[31] Holmes, M.J. and Grimes A. (2008). Is there long-run convergence among regional house prices in the UK? *Urban Studies*, 45, 1531-1544.

[32] Horvath, M., (1998). Cyclicality and sectoral linkages: Aggregate fluctuations from independent sectoral shocks. *Review of Economic Dynamics,* 1,, 781-808.

[33] Horvath, M., (2000). Sectoral shocks and aggregate fluctuations. *Journal of Monetary Economics,* 45, 69-106.

[34] Johansen, S., (1991). Estimation and hypothesis testing of cointegrating vectors in Gaussian vector autoregressive models. *Econometrica*, 59, 1551-1580.

[35] Kelejian, H.H., and Prucha, I. (1999), A generalized moments estimator for the autoregressive parameter in a spatial model, *International Economic Review*, 40, 509-533.

[36] Kelejian, H.H., and Prucha, I. (2009), Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances, *Journal of Econometrics*, forthcoming.

[37] Kelejian, H.H. and Robinson, D.P., (1995). Spatial autocorrelation: a suggested alternative to the autoregressive model, In: L. Anselin and R. Florax (Eds.), *New Directions in Spatial Econometrics*, Berlin: Springer-Verlag.

[38] Kneib, T. and Fahrmeir, L., (2009). A space-time study on forest health. In: R. Chandler and M. Scott (eds.), *Statistical Methods for Trend Detection and Analysis in the Environmental Sciences.* Wiley & Sons, Hoboken, forthcoming.

[39] Lee, L.F. (2004), Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models, *Econometrica*, 72, 1899-1925.

[40] Pesaran, B., and Pesaran, M.H., (2009). *Time Series Econometrics using Microfit 5*, Oxford University Press, Oxford.

[41] Pesaran, M.H., (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, 74, 967-1012.

[42] Pesaran, M.H., (2007). A pair-wise approach to testing for output and growth convergence. *Journal of Econometrics,* 138, 312–355.

[43] Pesaran, M.H., Schuermann, T. and S. Weiner, (2004). Modelling regional interdependencies using a global error-correcting macroeconometric model. *Journal of Business and Economics Statistics*, 22, 129-162.

[44] Pesaran, M. H., Shin, Y. and Smith, R.J. (2000). Structural analysis of vector error correction models with exogenous I(1) variables. *Journal of Econometrics*, 97, 293-343

[45] Pesaran, M.H., and Y. Shin, (1998). Generalized impulse response analysis in linear multivariate models. *Economics Letters,* 58, 17-29.

[46] Pesaran, M.H., and Y. Shin, (1999). An autoregressive distributed lag modelling approach to cointegration analysis, In: Steinar Strøm (ed.), *Econometrics and Economic Theory in the 20th Century*, Cambridge University Press, Cambridge.

[47] Pesaran, M.H. and E. Tosseti, (2009). Large panels with spatial correlations and common factors. http://www.econ.cam.ac.uk/faculty/pesaran/wp09/PesaranTosetti04Mar09.pdf.

[48] Peterson, W., Holly, S. and Gaudoin, P. (2002). *Further Work on an Economic Model of the Demand for Social Housing.* Report to the Department for the Environment, Transport and the Regions, London.

[49] Peuquet, D. J., (1994). It's about time: A conceptual rramework for the representation of temporal dynamics in geographic information systems. *Annals of the Association of American Geographers,* 84, 441-461.

[50] Piggott, G., (2007). *Commuting in London*, DMAG Briefing 2007-03, Greater London Authority, ISSN 1479-7879.

[51] Whitlle, P. (1954). On stationary processes in the plane, *Biometrika*, 41, 434-449.

[52] Whittle, P. (1956). On the variation of yield variance with plot size, *Biometrika*, 43, 337-343.

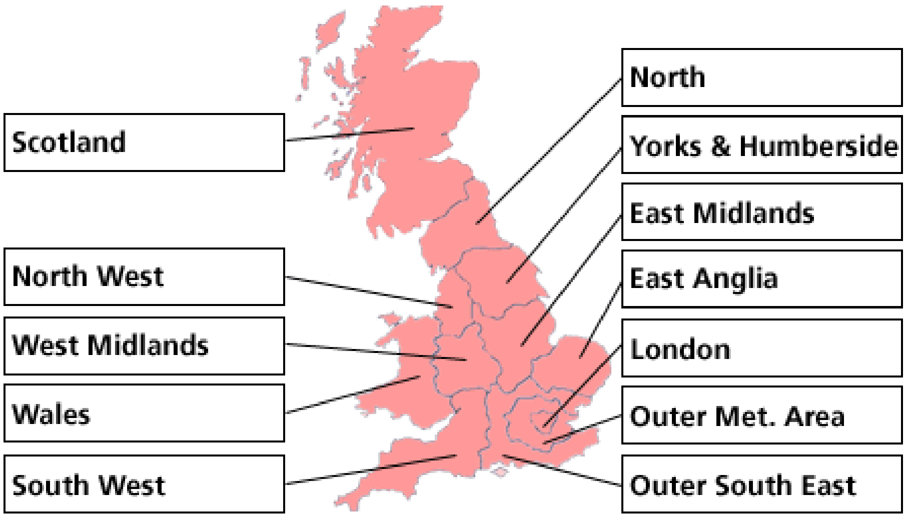
[53] Wu, D.M., (1973). Alternative tests of independence between stochastic regressors and disturbances. *Econometrica,* 41, 733-750.

Appendix A: Data Sources

Regional house prices (nominal, in British Pounds) in the UK are downloaded from the Homepage of the Nationwide Building Society and covers the period 1973q4 to 2008q2 (http://www.nationwide.co.uk/hpi/historical.htm). The regional price series are then deflated by the UK consumer price index (nominal consumers' expenditure divided by constant price consumers' expenditure) which is obtained from the Office of National Statistics. The New York house price index covers the shorter sample period of 1975q4 to 2008q2 and is constructed using data from the Federal Housing Finance Agency (FHFA). We use data for the New York-White Plains-Wayne Metropolitan Division. (http://www.fhfa.gov/Default.aspx?Page=87). The New York house price index is then deflated by the consumer price index of New York-Northern New Jersey-Long Island, NY-NJ-CT-PA, obtained from U.S. Bureau of Labor Statistics. The CPI data for New York is monthly (Series id: CUURA101SA0). Quarterly indices are constructed as simple averages of the monthly series.

Table 1: Regions, Abbreviations, and Data

Regions	Abbrev.	Regions	Abbrev.
East Anglia	EA	Outer South East	OSE
East Midlands	EM	Scotland	S
London	L	South West	SW
North	N	Wales	W
North West	NW	West Midlands	WM
Outer Metropolitan	OM	Yorkshire and Humberside	YH



Source: **Nationwide Homepage**

Notes: The Nationwide definition of regions differs from the ONS definition. The NW here excludes Cumbria which is added to the North East to comprise the North. The London region comprises the boroughs of London . The Outer Metropolitan region surrounds London and includes among others, South Essex, Reading, St Albans, and Medway. The Outer South East includes North Essex, Oxfordshire, Brighton, Southampton and Sussex. The map of definitions of regions is taken from http://www.nationwide.co.uk/hpi/regions.htm.

Table 2: Regions and Neighbours

Regions (Abbrev.)	Neighbours	Regions (Abbrev.)	Neighbours
East Anglia (EA)	EM, OSE	Outer South East (OSE)	EM, WM, EA, OM, L, SW
East Midlands (EM)	YH, NW, WM, EA, OSE	Scotland (S)	N
London (L)	OSE, OM.	South West (SW)	WM, OSE, W
North (N)	YH, NW, S	Wales (W)	NW, WM, SW
North West (NW)	N, YH, EM, WM, W	West Midlands (WM)	NW, EM, OSE, SW, W
Outer Metropolitan (OM)	EA, OSE, L	Yorkshire & Humberside (YH)	N, NW, EM

Table 3: Largest Ten Travel to Work Areas in terms of Net Inflows and Outflows

TTWA Name	Net Inflow(+)	in % of Total Employment
London	410108	9.70
Manchester	80029	9.47
Leeds	48643	11.10
Glasgow	47934	9.27
Birmingham	36353	5.29
Edinburgh	34521	10.45
Aberdeen	23878	11.72
Liverpool	20009	5.18
Bristol	18144	4.14
Newcastle & Durham	13448	2.93
TTWA Name	**Net Outflow(-)**	**in % of Total Employment**
Brighton	-15264	-9.10
Warrington & Wigan	-17331	-5.41
Luton & Watford	-18971	-6.27
Portsmouth	-20437	-7.99
Rochdale & Oldham	-21271	-13.18
Lanarkshire	-26901	-15.05
Wirral & Ellesmere Port	-27605	-20.44
Chelmsford & Braintree	-28300	-19.30
Maidstone & North Kent	-35055	-15.15
Southend & Brentwood	-41831	-19.26

Data Source: England & Wales: Office for National Statistics (ONS); Scotland: General Register Office for Scotland (GROS); Northern Ireland Statistics and Research Agency (NISRA)

Table 4: Net Outflows of Travel to Work Areas in Outer South East and Outer Metropolitan Regions

Regions	TTWA name	Net Outflow(-)	in % of Total Employment
OSE	Folkestone	-4124	-11.00
OSE	Dover	-4292	-13.07
OSE	Banbury	-5507	-9.87
OSE	Clacton	-5871	-24.42
OSE	Chichester & Bognor Regis	-6712	-8.03
OSE	Canterbury	-6834	-9.92
OM	Bedford	-7756	-9.64
OSE	Hastings	-8052	-13.99
OSE	Colchester	-8083	-8.95
OSE	Eastbourne	-8327	-13.17
OSE	Worthing	-10312	-14.31
OM	Stevenage	-10676	-6.78
OM	Harlow & Bishop's Stortford	-11304	-7.83
OM	Guildford & Aldershot	-14354	-4.08
OSE	Brighton	-15264	-9.10
OM	Luton & Watford	-18971	-6.27
OM	Chelmsford & Braintree	-28300	-19.30
OM	Maidstone & North Kent	-35055	-15.15
OM	Southend & Brentwood	-41831	-19.26

Data Source: England & Wales: Office for National Statistics (ONS); Scotland: General Register Office for Scotland (GROS); Northern Ireland Statistics and Research Agency (NISRA)

Table 5: Degree of Centrality of the Regional Network

Regions	Degree of Centrality of Network (c_i)
London	1.00
Outer South East	0.55
East Midland	0.45
North West	0.45
West Midland	0.45
North	0.27
Outer Metropolitan	0.27
South West	0.27
Wales	0.27
Yorkshire and Humberside	0.27
East Anglia	0.18
Scotland	0.09

Notes: The numbers reported are computed as $c_i = q_i/N$ with $N = 11$ where q_i is the number of regions with connections to region $i = 0, 1, ..., 11$. For example, London price changes are connected to all the remaining regions, giving $q_0 = N$. Similarly, price changes of the North region are connected to YH, NW, and S regions, yielding $q_{north} = 3$ (see Table 2)

Table 6: The Results of Trace Cointegration Tests with Unrestricted Intercepts and Restricted Trend Coefficients, and Tests of Over-identifying Restrictions in Bivariate VAR(4) Models of Log of Real House Prices of London and Other UK Regions (1974q4-2008q2)

	Trace Statistic (r is the number of cointegrating vectors)		H_0 : Cointegrating Vector is (1,-1)		
	$H_0 : r = 0$ vs $H_1 : r \geq 1$	$H_0 : r \leq 1$ vs $H_1 : r = 2$	LR Statistic	95% BCV	90%BCV
Outer Metropolitan	20.04	7.34	10.44	15.04	12.74
Outer South East	25.03*	7.83	9.54*	11.34	9.11
East Anglia	27.07**	9.62	9.86*	11.83	9.20
East Midlands	29.40**	9.79	5.44	10.22	7.75
West Midlands	24.00*	8.90	3.62	10.94	8.60
South West	28.80**	11.31*	2.13	9.96	7.99
Wales	22.67	6.68	5.03	11.14	8.86
Yorkshire & Humberside	23.56*	6.14	7.12	10.84	8.74
North West	23.09*	7.31	3.32	10.79	8.48
North	20.19	6.77	2.13	11.08	8.83
Scotland	16.92	4.04	1.83	11.37	9.49

Notes: The trace statistics reported are based on the bivariate VAR(4) specification of log of real house prices of London and other UK regions, with unrestricted intercepts and restricted trend coefficients. The trace statistic is the cointegration test statistic of Johansen (1991). The log-likelihood ratio (LR) statistic reported is for testing the cotrending restriction with the cointegrating vector given by (1,-1) for the log house prices in London and the other region. For the trace test, the 95% and 90% critical values of the test for $H_0 : r = 0$ are 25.77 and 23.08, and those for $H_0 : r \leq 1$ are 12.39 and 10.55, respectively. BCV stands for bootstrap critical values, which are based on 1000 bootstrap replications. ** signifies that the test rejects the null at the 5% level, and * at the 10% level. All statistics are computed using Microfit 5 (Pesaran and Pesaran, 2009).

Table 7: Error Correction Coefficients in Cointegrating Bivariate VAR(4) of Log of Real House Prices of London and other UK regions (1974q4-2008q2)

	Error Correction Equation for London (p_{0t})				Error Correction Equation for other Regions (p_{it})			
Regions (i)	EC Coeff. $\hat{\phi}_{0i}$	t-ratio	$\overline{R}^2$	Serial Correlation	EC Coeff. $\hat{\phi}_{i0}$	t-ratio	$\overline{R}^2$	Serial Correlation
Outer Metropolitan	-0.02	-0.35	0.41	10.76**	-0.03	-0.62	0.49	15.16***
Outer South East	0.02	0.50	0.40	6.84	-0.09**	-2.03	0.45	8.12*
East Anglia	0.03	0.85	0.37	6.08	-0.08**	-2.54	0.35	4.36
East Midlands	0.02	0.67	0.36	4.33	-0.08***	-3.36	0.43	1.42
West Midlands	0.01	0.33	0.39	3.87	-0.07***	-3.05	0.35	6.79
South West	0.01	0.18	0.41	2.76	-0.11***	-3.03	0.40	11.61**
Wales	0.01	0.55	0.38	7.45	-0.05***	-2.88	0.33	6.48
Yorkshire & Humberside	0.01	0.44	0.40	4.59	-0.04***	-2.90	0.31	17.98***
North West	0.00	0.20	0.37	4.94	-0.04***	-3.04	0.42	2.96
North	-0.01	-0.42	0.37	9.88**	-0.05***	-3.05	0.21	1.26
Scotland	-0.01	-0.71	0.40	3.53	-0.03***	-2.84	0.14	4.79

Notes: For the London equations "EC Coeff." is the estimate of the coefficient of the error correction term, ϕ_{0i}, in $\Delta p_{0t} = \phi_{0i}(p_{0,t-1} - p_{i,t-1}) + \Sigma_{\ell=1}^{3} a_{0i,\ell}\Delta p_{0,t-\ell} + \Sigma_{\ell=1}^{3} b_{0i,\ell}\Delta p_{i,t-\ell} + \varepsilon_{0it}$. For other regions it is given by the estimate of ϕ_{i0} in the EC regressions, $\Delta p_{it} = \phi_{i0}(p_{i,t-1} - p_{0,t-1}) + \Sigma_{\ell=1}^{3} a_{i0,\ell}\Delta p_{0,t-\ell} + \Sigma_{\ell=1}^{3} b_{i0,\ell}\Delta p_{i,t-\ell} + \varepsilon_{i0t}$. Intercepts are included in all the regressions. The associated t-ratios for the error correction coefficients are given next to the coefficients. The $\bar{R}^2$ is the adjusted R^2. The column "Serial Correlation" reports the Breusch-Godfrey serial correlation test statistic which is distributed approximately as χ_4^2 under the null of no residual serial correlation. *** signifies that the test rejects the null at the 1% level , ** at the 5% level, and * at the 10% level.

Table 8: Estimation Results of Region Specific House Price Diffusion Equation with London as a Dominant Region (1974q1-2008q2)

Lag-orders $\{\hat{k}_{ia}, \hat{k}_{ib}, \hat{k}_{ic}\}$ selected by SBC										
Regions	EC1 $(\hat{\phi}_{i0})$	EC2 $(\hat{\phi}_{is})$	Own Lag Effects	Neighbour Lag Effects	London Lag Effects	London Contemporaneous Effects	Wu-Hausman Statistics	$\hat{k}_{ia}$	$\hat{k}_{ib}$	$\hat{k}_{ic}$
London	–	–	0.036 (0.246)	0.666* (4.314)	–	–		1	1	-
Outer Metropolitan	–	–	−0.103 (−1.095)	0.354* (4.107)	–	0.658*** (14.326)	1.018	1	1	0
Outer South East	–	–	−0.158 (−1.349)	0.423* (3.290)	–	0.746*** (15.017)	0.821	1	1	0
East Anglia	−0.045** (−2.002)	–	−0.033 (−0.320)	0.271* (2.158)	–	0.653*** (9.085)	-0.903	1	1	0
East Midlands	−0.057*** (−3.475)	–	−0.029 (−0.279)	0.808* (5.184)	−0.501*** (−4.459)	0.523*** (8.525)	-0.694	1	2	2
West Midlands	−0.061*** (−3.770)	–	−0.203* (−1.933)	0.791*** (4.952)	−0.442*** (−3.524)	0.498*** (7.043)	0.032	1	1	2
South West	−0.113*** (−4.557)	–	−0.026 (−0.249)	0.371*** (3.095)	−0.326*** (−2.744)	0.670*** (10.813)	-1.240	1	1	2
Wales	–	–	−0.137 (−1.414)	1.319*** (7.777)	−0.757*** (−6.645)	0.661*** (9.455)	-0.895	1	3	3
Yorkshire & Humberside	–	–	0.180 (1.338)	0.561*** (3.834)	−0.333*** (−3.047)	0.577*** (7.252)	-1.874*	2	1	2
North West	–	–	0.061 (−0.470)	0.918*** (6.399)	−0.452*** (−5.757)	0.423*** (7.751)	0.054	3	2	2
North	−0.039*** (−2.984)	–	−0.213** (−2.150)	0.750*** (5.074)	−0.235** (−2.248)	0.266*** (3.078)	0.610	1	1	1
Scotland	–	−0.098*** (−4.232)	0.019 (0.202)	0.050 (0.640)	–	0.326*** (5.266)	-1.174	1	1	0

Notes: This table reports estimates based on the price equations $\Delta p_{it} = \phi_{is}(p_{i,t-1} - \bar{p}^s_{i,t-1}) + \phi_{i0}(p_{i,t-1} - p_{0,t-1}) + \Sigma_{\ell=1}^{k_{ia}} a_{i\ell}\Delta p_{i,t-\ell} + \Sigma_{\ell=1}^{k_{ib}} b_{i\ell}\Delta\bar{p}^s_{i,t-\ell} + \Sigma_{\ell=1}^{k_{ic}} c_{i\ell}\Delta p_{0,t-\ell} + c_{i0}\Delta p_{0,t} + \varepsilon_{it}$, for $i = 1, 2, ..., N$. For $i = 0$, denoting the London equation, we have the additional *a priori* restrictions, $\phi_{00} = c_{00} = 0$. "EC1", "EC2", "Own lag effects", "Neighbour lag effects", "London lag effects", and "London contemporaneous effects" relate to the estimates of ϕ_{i0}, ϕ_{is}, $\Sigma_{\ell=1}^{k_{ia}} a_{i\ell}$, $\Sigma_{\ell=1}^{k_{ib}} b_{i\ell}$, $\Sigma_{\ell=1}^{k_{ic}} c_{i\ell}$, and c_{i0}, respectively. t-ratios are shown in the parenthesis. *** signifies that the test rejects the null at the 1% level , ** at the 5% level, and * at the 10% level. The error correction coefficients (ϕ_{is} and ϕ_{i0}) are restricted such that at least one of them are statistically significant at the 5% level. Wu-Hausman is the t-ratio for testing $H_0 : \lambda_i = 0$ in the augmented regression $\Delta p_{it} = \phi_{is}(p_{i,t-1} - \bar{p}^s_{i,t-1}) + \phi_{i0}(p_{i,t-1} - p_{0,t-1}) + \Sigma_{\ell=1}^{k_{ia}} a_{i\ell}\Delta p_{i,t-\ell} + \Sigma_{\ell=1}^{k_{ib}} b_{i\ell}\Delta\bar{p}^s_{i,t-\ell} + \Sigma_{\ell=0}^{k_{ic}} c_{i\ell}\Delta p_{0,t-\ell} + \lambda_i\hat{\varepsilon}_{0t} + \varepsilon_{it,}$, where $\hat{\varepsilon}_{0t}$ is the residual of the London house price equation, and the error correction coefficients are restricted as described above. In selecting the lag orders, k_{ia}, k_{ib}, and k_{ic} the maximum lag-order is set to 4. All the regressions include an intercept term.

Table 9 Wu-Hausman Statistics for Testing the Exogeneity of House Prices of the Assumed Dominant Region

		Assumed Dominant Region											
		L	OM	OSE	EA	EM	WM	SW	W	YH	NW	N	S
Price Equation	L	–	1.94*	3.46***	2.03**	3.60***	2.66***	2.22**	2.05**	2.28**	3.32***	3.50***	1.71*
	OM	1.02	–	2.35**	1.78*	4.64***	0.79	1.70*	1.45	2.55**	3.46***	3.95***	1.87*
	OSE	0.82	1.68*	–	1.17	4.26***	3.33***	2.45**	1.79*	2.87***	2.53**	5.46***	1.87*
	EA	-0.90	0.20	1.30	–	2.25**	4.07***	2.77***	0.33	3.62***	3.25***	4.00***	1.65*
	EM	-0.69	1.12	2.02**	1.46	–	-1.35	-2.14**	-4.39***	1.41	1.32	3.54***	-0.12
	WM	0.03	-0.25	1.10	0.98	1.00	–	-0.37	-1.22	0.40	1.41	2.34**	0.02
	SW	-1.24	-1.28	-0.37	-0.85	-1.01	2.65***	–	2.06**	0.38	1.15	2.89***	0.12
	W	-0.90	0.70	1.72*	-0.85	2.37**	0.50	0.79	–	-0.63	0.06	0.10	-2.44**
	YH	-1.87*	-0.01	1.28	-0.17	2.50**	-0.60	-1.61	-0.55	–	-1.92*	2.86***	1.29
	NW	0.05	1.91*	-0.38	1.53	4.30***	2.20**	1.70*	0.31	-0.69	–	1.30	1.09
	N	0.61	2.11**	1.13	0.60	4.04***	0.80	1.47	-0.62	1.26	0.76	–	-1.24
	S	-1.17	1.28	-0.10	-0.31	0.48	-1.57	-0.26	0.98	-1.04	1.73*	-0.59	–

Notes: The Wu-Hausman statistic is computed as the t-ratio for testing $H_0 : \lambda_i = 0$ in the augmented regression $\Delta p_{it} = \phi_{is}(p_{i,t-1} - \bar{p}^s_{i,t-1}) + \phi_{i0}(p_{i,t-1} - p_{0,t-1}) + \Sigma_{\ell=1}^{k_{ia}} a_{i\ell}\Delta p_{i,t-\ell} + \Sigma_{\ell=1}^{k_{ib}} b_{i\ell}\Delta\bar{p}^s_{i,t-\ell} + \Sigma_{\ell=0}^{k_{ic}} c_{i\ell}\Delta p_{0,t-\ell} + \lambda_i\hat{\varepsilon}_{0t} + \varepsilon_{it}$, where $\hat{\varepsilon}_{0t}$ is the residual of the house price equation for the assumed dominant region. The error correction coefficients (ϕ_{is} and ϕ_{i0}) are restricted so that at least one is statistically significant (see also the notes of Table 8). In order to avoid perfect multicollinearity, only one error correction term is included for the equation of Scotland when North is the assumed dominant region. *** signifies that the test is significant at the 1% level (absolute values larger than 2.576), ** at the 5% level (absolute values larger than 1.960), and * at the 10% level (absolute values larger than 1.645). In selecting the lag orders, SBC is used and the maximum lag-order is set to 4. All regressions include an intercept. See Table 2 for the regional mnemonics and a list of neighbours.

Table 10: Effects of New York House Price Changes on London, with London Treated as the Dominant Region in the UK (1976q1-2008q2)

								Lag-orders $\{\hat{k}_{ia}, \hat{k}_{ib}, \hat{k}_{ic}, \hat{k}_{i,NY}\}$ selected by SBC		
Regions (i)	EC1 ($\phi_{0,NY}$)	EC2 (ϕ_{0s})	Own Lag effects	Neighbour Lag Effects	New York Lag Effects	New York Contemporaneous Effects	Wu-Hausman Statistics	$\hat{k}_{ia}$	$\hat{k}_{ib}$	$\hat{k}_{i,NY}$
New York	–	–	0.764*** (8.560)	–	–	–	–	3	-	-
London	–	–	−0.094 (−0.641)	0.583*** (3.839)	0.316*** (2.745)	0.278** (2.445)	-0.139	1	1	1
Other UK Regions	EC1 ($\hat{\phi}_{i0}$)	EC2 ($\hat{\phi}_{is}$)	Own Lag Effects	Neighbour Lag Effects	London Lag Effects	London Contemporaneous Effects	Wu-Hausman Statistics	$\hat{k}_{ia}$	$\hat{k}_{ib}$	$\hat{k}_{ic}$
Outer Metropolitan	–	–	−0.118 (−1.217)	0.370*** (4.178)	–	0.656*** (13.944)	-0.481	1	1	-
Outer South East	–	–	−0.153 (−1.270)	0.410*** (3.103)	–	0.744*** (14.583)	-0.527	1	1	0
East Anglia	−0.044* (−1.904)	–	−0.041 (−0.384)	0.275** (2.083)	–	0.660*** (8.867)	-1.398	1	1	0
East Midlands	−0.057*** (−3.412)	–	−0.039 (−0.361)	0.824*** (5.099)	−0.501*** (−4.357)	0.513*** (8.101)	0.600	1	2	2
West Midlands	−0.056*** (−3.930)	–	0.000 (0.001)	0.583*** (3.866)	−0.425*** (−3.810)	0.492*** (7.793)	1.716*	1	1	2
South West	−0.113*** (−4.429)	–	−0.045 (−0.421)	0.400*** (3.167)	−0.339*** (−2.773)	0.681*** (10.601)	-0.467	1	1	2
Wales	–	–	−0.181* (−1.804)	1.429*** (8.001)	−0.769*** (−6.666)	0.672*** (9.368)	0.549	1	3	3
Yorkshire & Humberside	–	–	−0.065 (−0.605)	0.641*** (4.601)	–	0.458*** (6.438)	1.617	1	1	0
North West	–	–	−0.072 (−0.726)	1.034*** (8.048)	−0.472*** (−5.743)	0.428*** (7.446)	1.232	1	3	2
North	−0.040*** (−2.948)	–	−0.223** (−2.200)	0.751*** (4.954)	−0.240** (−2.228)	0.256*** (2.860)	2.207**	1	1	1
Scotland	–	−0.103*** (−4.360)	0.009 (0.087)	0.046 (0.587)	–	0.334*** (5.156)	0.828	1	1	0

Notes: The table report the estimates of the autoregressive model of New York real house price changes, $\Delta p_{NY,t} = \Sigma_{\ell=1}^{k_{NY}} a_{NY,\ell}\Delta p_{NY,t-\ell} + \varepsilon_{NY,t}$, the London house price changes, $\Delta p_{0t} = \phi_{0s}(p_{0,t-1} - \bar{p}^s_{0,t-1}) + \phi_{0,NY}(p_{0,t-1} - p_{NY,t-1}) + \Sigma_{\ell=1}^{k_{0a}} a_{0\ell}\Delta p_{0,t-\ell} + \Sigma_{\ell=1}^{k_{0b}} b_{0\ell}\Delta\bar{p}^s_{0,t-\ell} + \Sigma_{\ell=0}^{k_{0,NY}} c_{NY\ell}\Delta p_{NY,t-\ell} + \varepsilon_{0t}$, with the remaining regional price equations as specified at the foot of Table 8. The differences between the results reported for regional equations in this table and those reported in Table 8 are due only to the differences in the estimation sample. For the lag order selection the maximum lag-order of the autoregression of $\Delta p_{NY,t}$ and the regressions of Δp_{it}, $i = 0, 1, ..., N$, are set to 8 and 4, respectively.

Table 11: F Statistics for Joint Significance of Contemporaneous and Lagged Effects of New York House Price Changes in the UK Regional House Price Equations

A: With Contemporaneous London House Price Changes		
Regions	F Statistics	p-values
London	5.323	0.000
Outer Metropolitan	0.564	0.640
Outer South East	1.499	0.219
East Anglia	0.531	0.662
East Midlands	0.570	0.636
West Midlands	2.585	0.057
South West	0.798	0.498
Wales	0.784	0.505
Yorkshire & Humberside	0.685	0.563
North West	0.858	0.465
North	0.386	0.763
Scotland	1.522	0.213
B: Without Contemporaneous London House Price Changes		
Regions	F Statistics	p-values
London	5.323	0.000
Outer Metropolitan	2.479	0.065
Outer South East	3.456	0.019
East Anglia	2.844	0.041
East Midlands	0.597	0.618
West Midlands	0.036	0.991
South West	1.772	0.156
Wales	0.851	0.469
Yorkshire & Humberside	0.930	0.429
North West	0.316	0.814
North	0.196	0.899
Scotland	1.060	0.369

Notes: The F statistics are for testing the joint hypothesis $H_0 : c_{i,NY,\ell} = 0$, for $\ell = 0, 1, 2$, in the region-specific equations. The statistics in panel A are based on the regressions: $\Delta p_{it} = intercept + \phi_{is}(p_{i,t-1} - \bar{p}^s_{i,t-1}) + \phi_{i0}(p_{i,t-1} - p_{0,t-1}) + \Sigma^2_{\ell=1} a_{i\ell}\Delta p_{i,t-\ell} + \Sigma^2_{\ell=1} b_{i\ell}\Delta\bar{p}^s_{i,t-\ell} + \Sigma^2_{\ell=0} c_{i\ell}\Delta p_{0,t-\ell} + \Sigma^2_{\ell=0} c_{i,NY,\ell}\Delta p_{NY,t-\ell} + \varepsilon_{it}$, for $i = 1, 2...., 11$. In panel B the F statistics are computed using the same regressions except that contemporaneous change in London prices ($\Delta p_{0,t}$) is excluded from the regression. For London the regression equation used to compute the F statistic is given by $\Delta p_{0t} = intercept + \Sigma^2_{\ell=1} a_{0\ell}\Delta p_{0,t-\ell} + \Sigma^2_{\ell=1} b_{0\ell}\Delta\bar{p}^s_{0,t-\ell} + \Sigma^2_{\ell=0} c_{i,NY,\ell}\Delta p_{NY,t-\ell} + \varepsilon_{it}$, which is the same in both panels. The error correction coefficients (ϕ_{is}, and ϕ_{i0}) are restricted as before (see the notes to Table 8). The p-values for the tests are given in the last column.

Figure 1: UK Real House Prices by Regions

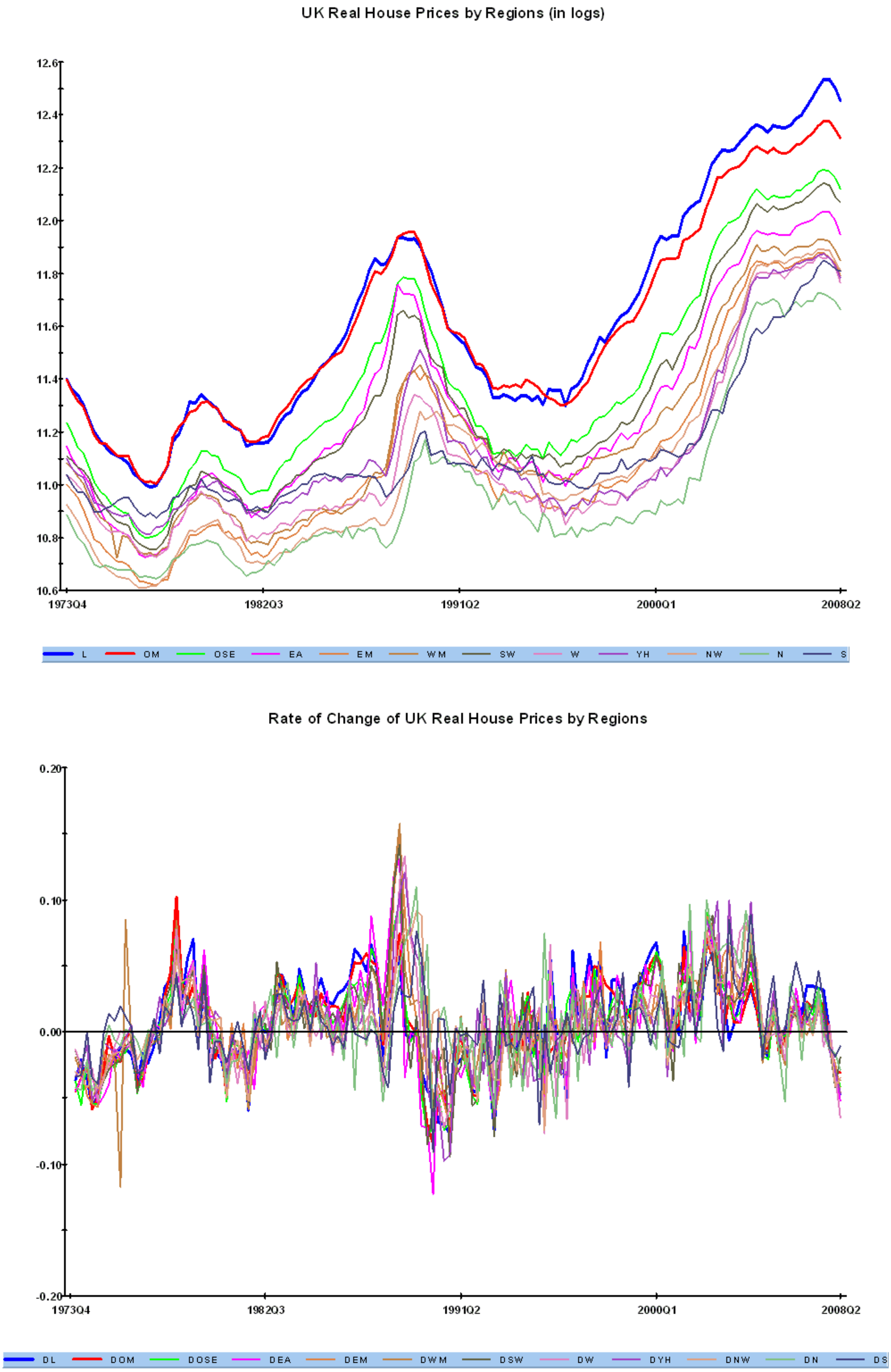
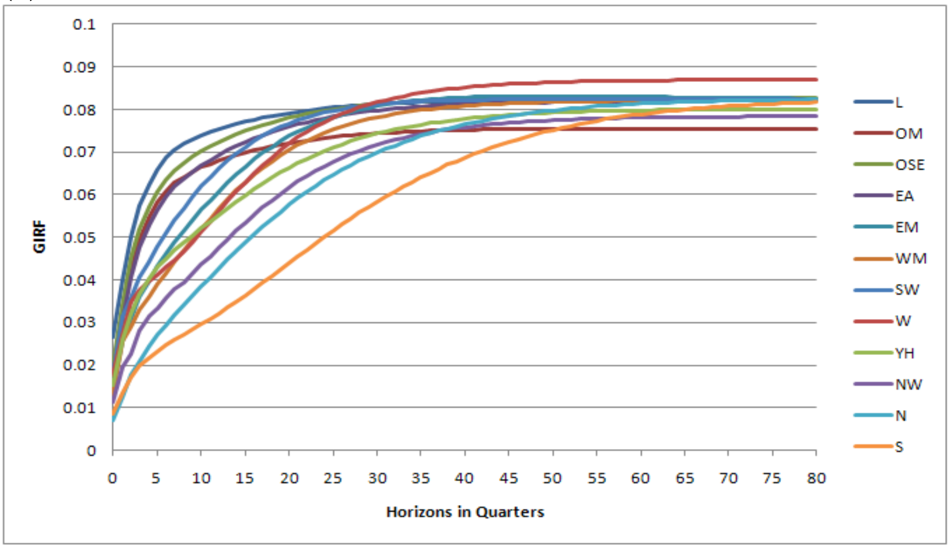
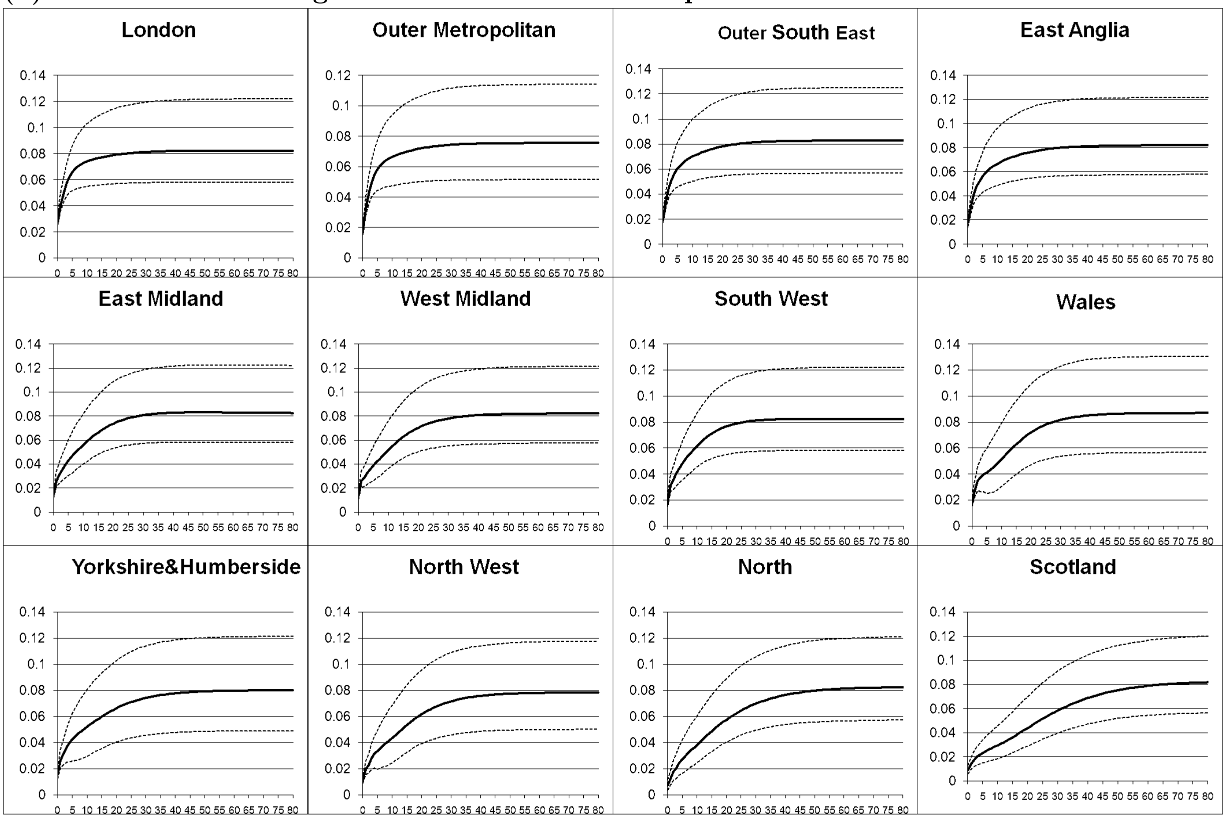


Figure 2: Generalised Impulse Responses of a Positive Unit (one s.e.) Shock to London House Prices

(a) GIRFs for All Regions

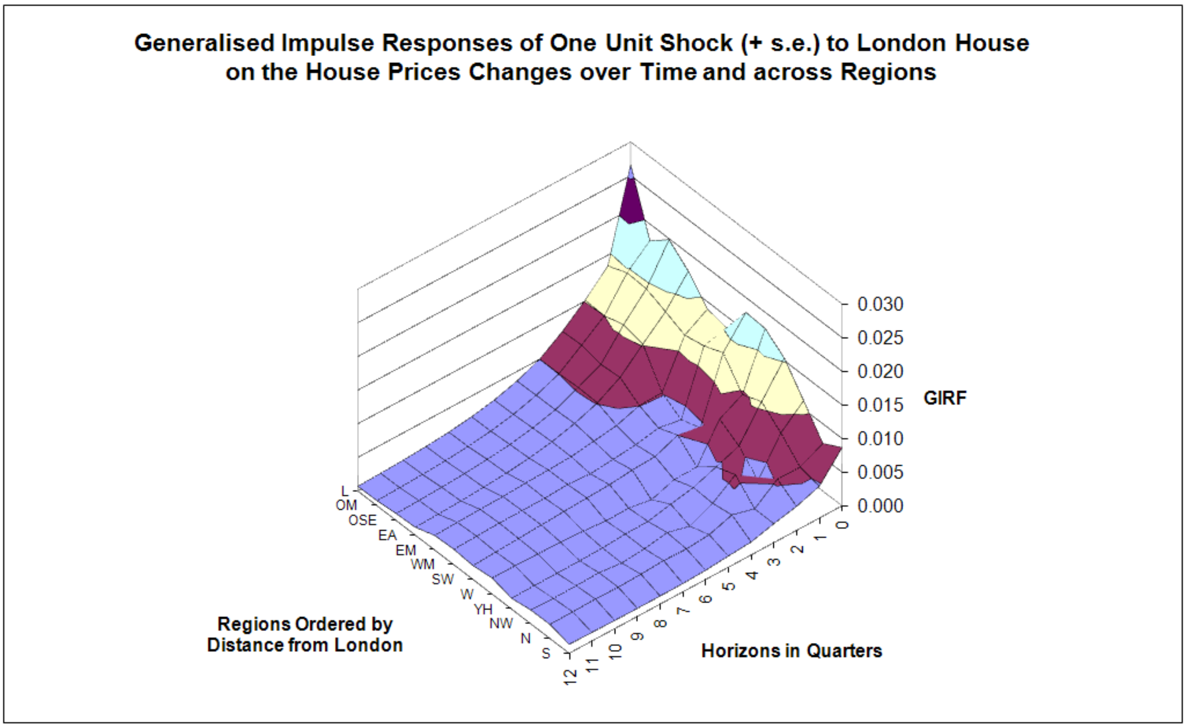


(b) GIRFs for Each Region with the 90% Bootstrap Error Bounds



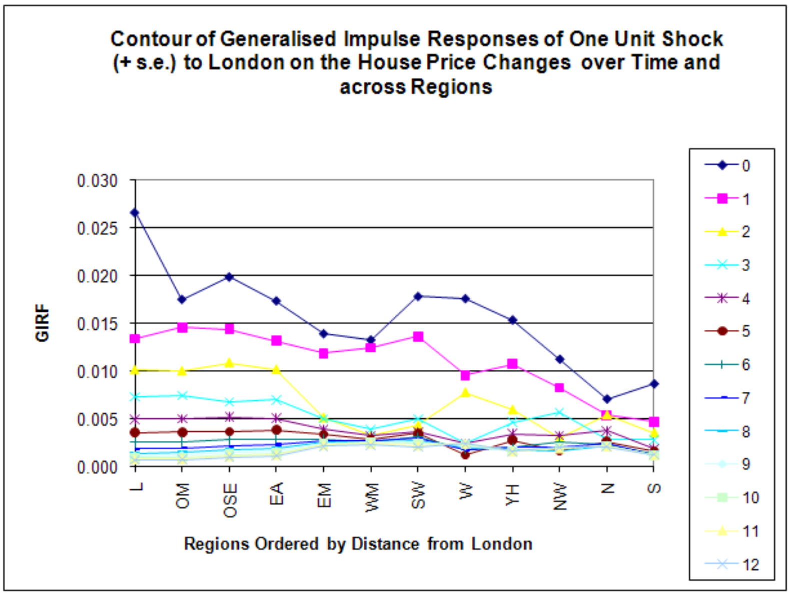
Notes: Figures show that the mean estimates (solid line) with 90% bootstrap error bounds (broken line, based on 10000 bootstrap samples) in the case of the model without New York price changes: $\Delta p_{it} = \phi_{is}(p_{i,t-1} - \bar{p}^s_{i,t-1}) + \phi_{i0}(p_{i,t-1} - p_{0,t-1}) + \Sigma_{\ell=1}^{k_{ia}} a_{i\ell}\Delta p_{i,t-\ell} + \Sigma_{\ell=1}^{k_{ib}} b_{i\ell}\Delta \bar{p}^s_{i,t-\ell} + \Sigma_{\ell=1}^{k_{ic}} c_{i\ell}\Delta p_{0,t-\ell} + \varepsilon_{it}$,for $i = 0, ..., N$, including an intercept. Lag-orders $\{\hat{k}_{ia}, \hat{k}_{ib}, \hat{k}_{ic}\}$ are selected by SBC. The restrictions of error correction terms, which are explained in the note to Table 8 (see the notes therein), as well as selected lag-orders by SBC are imposed in the estimation in bootstrap procedure.

Figure 3:



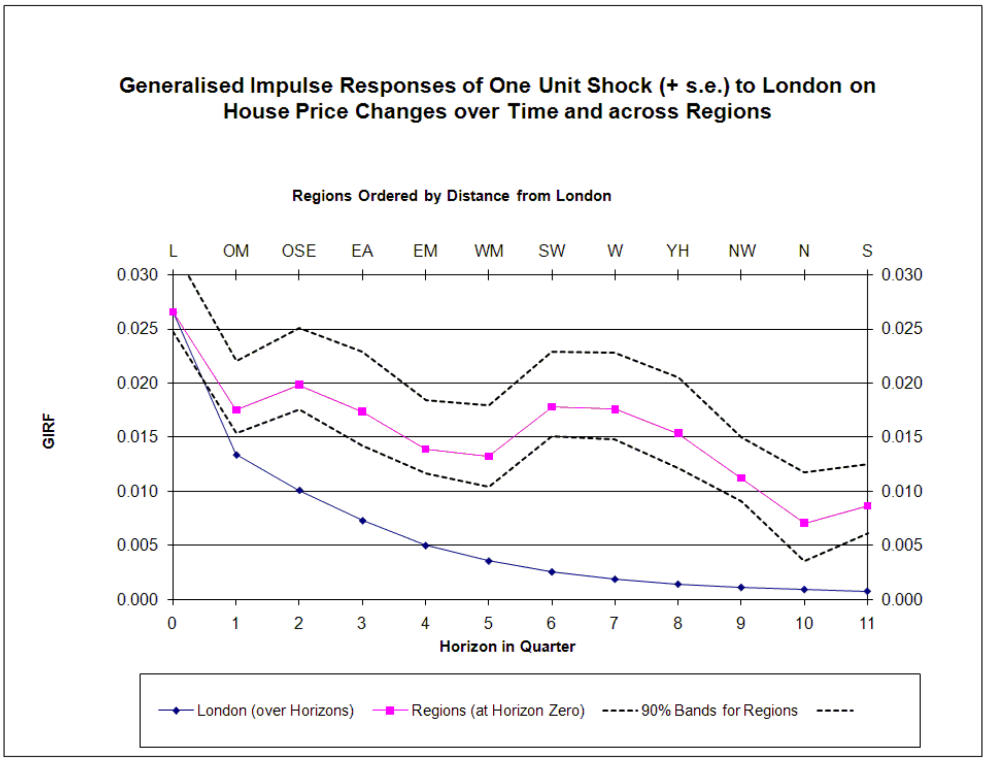
Notes:The regions are ordered by distance from London. See also notes to Figure 2. See Table 2 for the abbreviations of regions and their neighbours.

Figure 4:



Notes: The regions are ordered by distance from London. See also notes to Figure 2. See Table 2 for the abbreviations of regions and their neighbours.

Figure 5:



Notes: This figure shows the effect of a unit shock to London house prices on London over time, and the impact effects of the same shock on regions ordered by their distance from London. Broken lines are bootstrap 90% confidence band of the GIRFs for the regions, based on 10000 bootstrap samples. See Table 2 for the abbreviations of regions and their neighbours.

Figure 6:

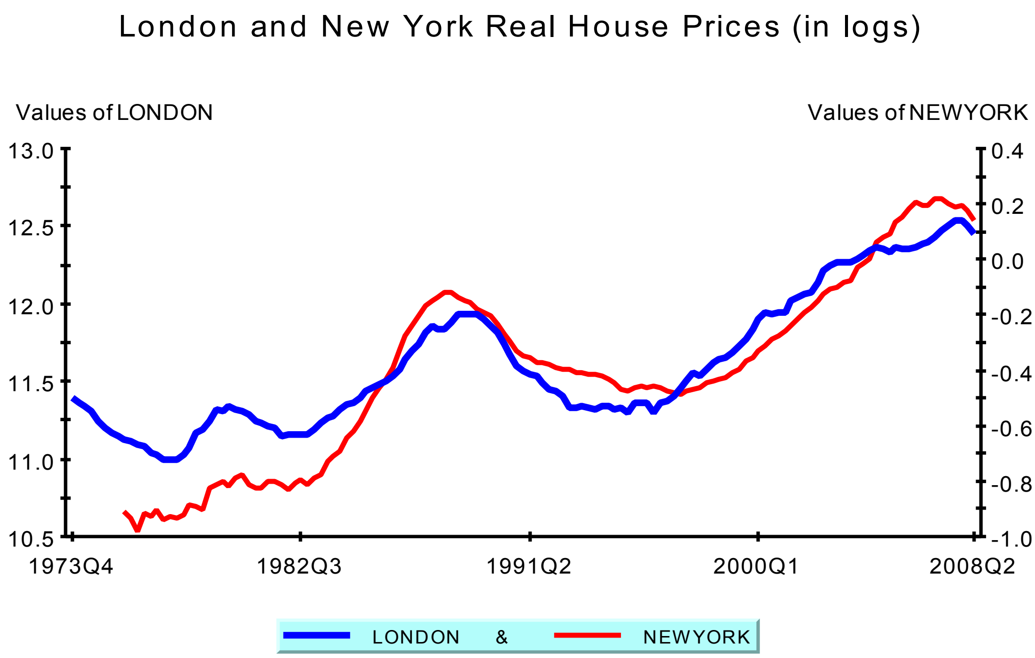
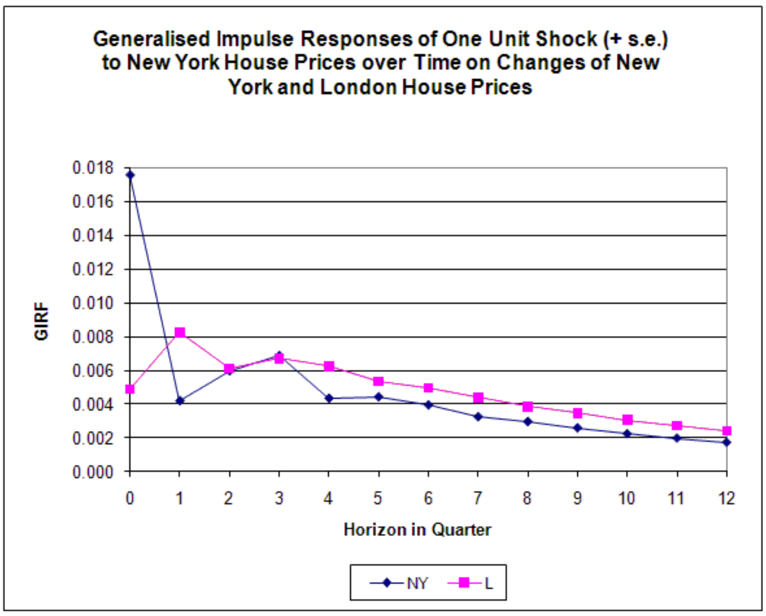


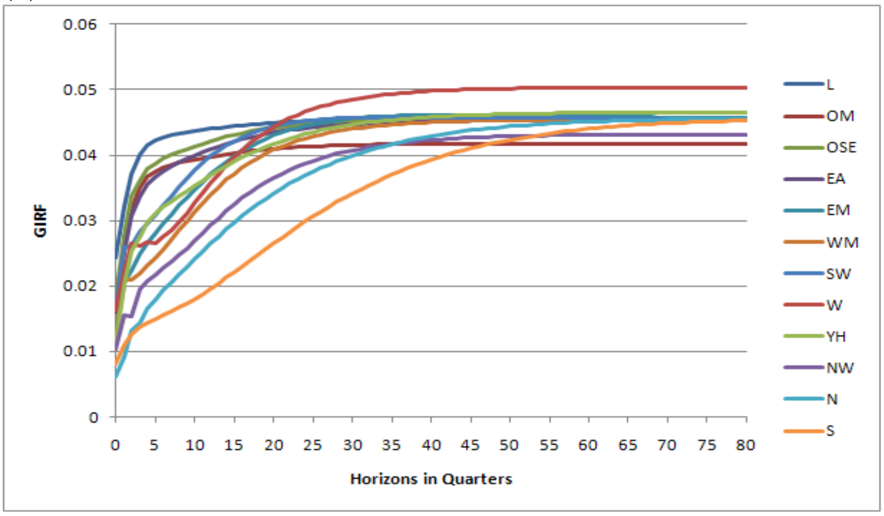
Figure 7:



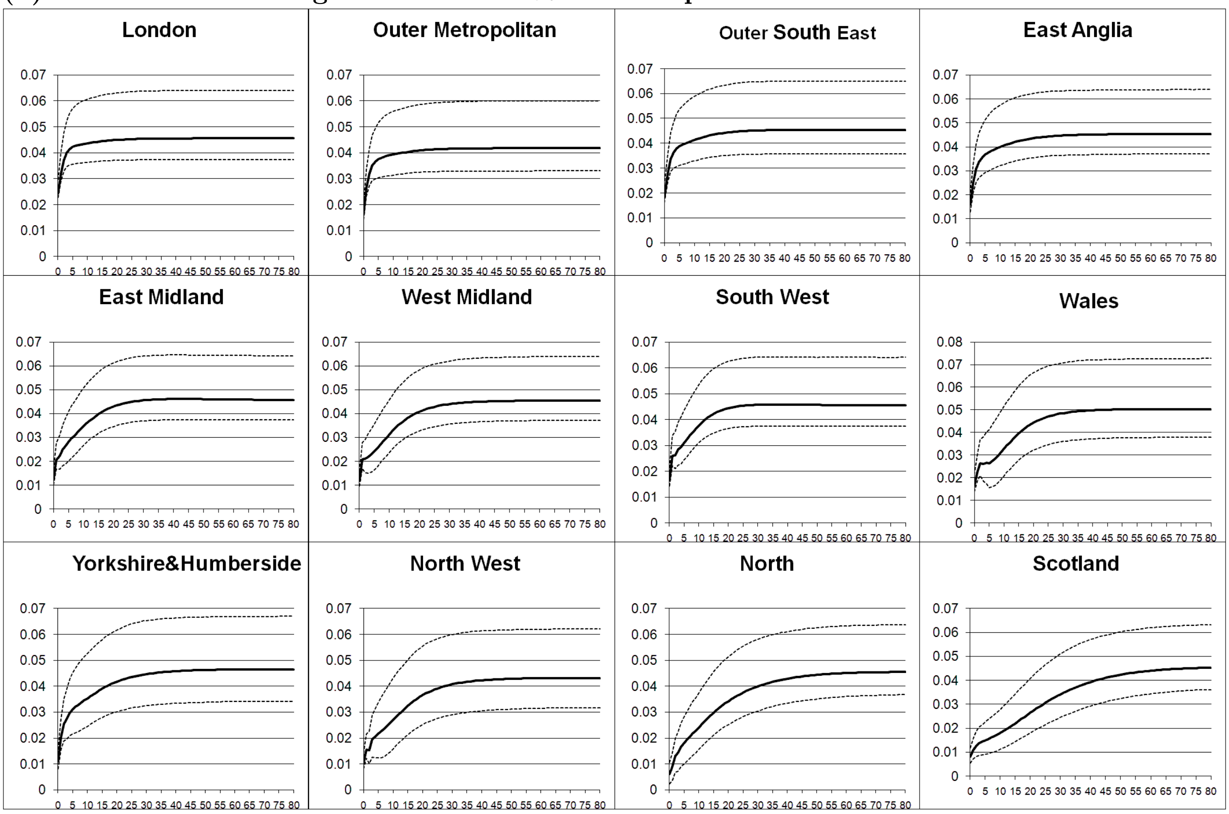
Note: See the notes to Table 10.

Figure 8: Generalised Impulse Responses of a Positive Unit (one s.e.) Shock to London House Prices in the Model with New York House Prices

(a) GIRFs for All Regions



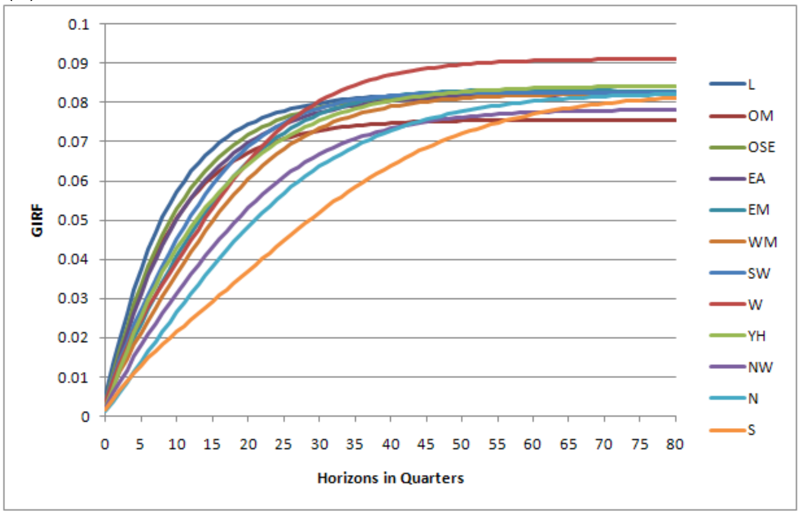
(b) GIRFs for Each Region with the 90% Bootstrap Error Bounds



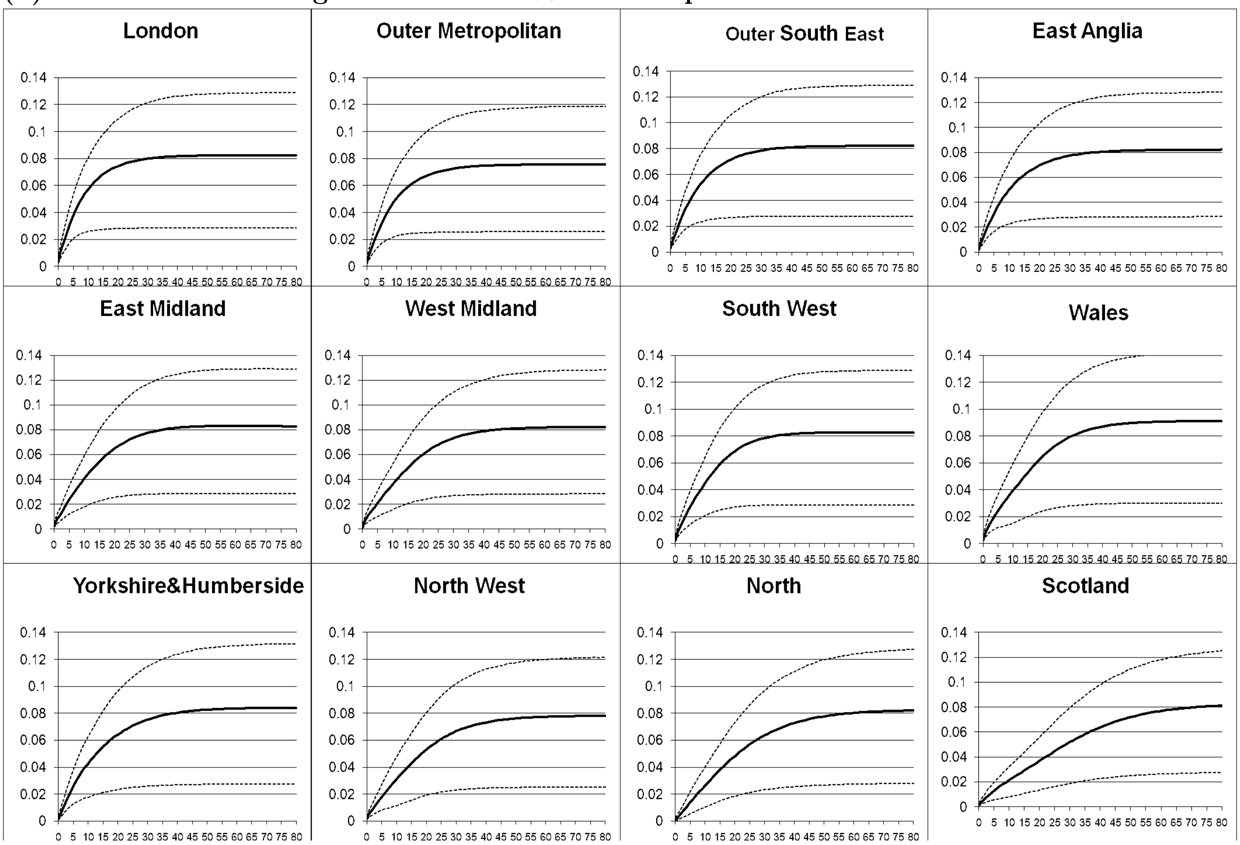
Notes: Figures show that the bootstrap mean estimates (solid line) with 90% bootstrap error bounds (broken line, based on 10000 bootstrap samples) in the case of the model with New York house price: which is a system of the autoregresive model of changes in New York house prices, $\Delta p_{NY,t}$, the London house prices, $\Delta p_{0t} = \phi_{0s}(p_{0,t-1} - \bar{p}^s_{0,t-1}) + \phi_{0,NY}(p_{0,t-1} - p_{NY,t-1}) + \sum_{\ell=1}^{k_{0a}} a_{0\ell}\Delta p_{0,t-\ell} + \sum_{\ell=1}^{k_{0b}} b_{0\ell}\Delta \bar{p}^s_{0,t-\ell} + \sum_{\ell=0}^{k_{0c}} c_{NY\ell}\Delta p_{NY,t-\ell} + \varepsilon_{0t}$, and the price equations for the remaining regions: $\Delta p_{it} = \phi_{is}(p_{i,t-1} - \bar{p}^s_{i,t-1}) + \phi_{i0}(p_{i,t-1} - p_{0,t-1}) + \sum_{\ell=1}^{k_{ia}} a_{i\ell}\Delta p_{i,t-\ell} + \sum_{\ell=1}^{k_{ib}} b_{i\ell}\Delta \bar{p}^s_{i,t-\ell} + \sum_{\ell=0}^{k_{ic}} c_{i\ell}\Delta p_{0,t-\ell} + \varepsilon_{it}$,for $i = 0, ..., N$, including an intercept. Lag-orders $\{\hat{k}_{ia}, \hat{k}_{ib}, \hat{k}_{ic}\}$ are selected by SBC. For the lag order selection the maximum lag-order of autoregression of $\Delta p_{NY,t}$ and the regression of Δp_{it}, $i = 0, 1, ..., N$, are set to 8 and 4, respectively. The restrictions of error correction terms, which are explained in the note to Table 7 (see the notes therein), as well as selected lag-orders by SBC are imposed in the estimation in bootstrap procedure.

Figure 9: Generalised Impulse Responses of a Positive Unit (one s.e.) Shock to New York House Prices in the Model with New York House Prices

(a) GIRFs for All Regions



(b) GIRFs for Each Region with the 90% Bootstrap Error Bounds



Notes: Figures show that the bootstrap mean estimates (solid line) with 90% bootstrap error bounds (broken line, based on 10000 bootstrap samples) of the New York model, which is a system of the autoregresive model of changes in New York house prices, $\Delta p_{NY,t}$, the London house prices, $\Delta p_{0t} = \sum_{\ell=1}^{k_{0a}} a_{0\ell}\Delta p_{0,t-\ell} + \sum_{\ell=1}^{k_{0b}} b_{0\ell}\Delta\bar{p}^{s}_{0,t-\ell} + \sum_{r=0}^{k_{0c}} c_{NYr}\Delta p_{NY,t-r} + \varepsilon_{0t}$, and the price equations for the remaining regions: $\Delta p_{it} = \sum_{\ell=1}^{k_{ia}} a_{i\ell}\Delta p_{i,t-\ell} + \sum_{\ell=1}^{k_{ib}} b_{i\ell}\Delta\bar{p}^{s}_{i,t-\ell} + \sum_{r=0}^{k_{ic}} c_{i0r}\Delta p_{0,t-r} + \varepsilon_{it}$, for $i = 1, 2, ..., N$. All regressions include an intercept. For the lag order selection the maximum lag-order of autoregression of $\Delta p_{NY,t}$ and the regression of Δp_{it}, $i = 0, 1, ..., N$, are set to 8 and 4, respectively. The restrictions of error correction terms, which are explained in the note to Table 7 (see the notes therein), as well as selected lag-orders by SBC are imposed in the estimation in bootstrap procedure.