MODELING US HOUSING PRICES BY SPATIAL DYNAMIC STRUCTURAL EQUATION MODELS

BY PASQUALE VALENTINI, LUIGI IPPOLITI AND LARA FONTANELLA

University of Chieti-Pescara

This article proposes a spatial dynamic structural equation model for the analysis of housing prices at the State level in the USA. The study contributes to the existing literature by extending the use of dynamic factor models to the econometric analysis of multivariate lattice data. One of the main advantages of our model formulation is that by modeling the spatial variation via spatially structured factor loadings, we entertain the possibility of identifying similarity "regions" that share common time series components. The factor loadings are modeled as conditionally independent multivariate Gaussian Markov Random Fields, while the common components are modeled by latent dynamic factors. The general model is proposed in a state-space formulation where both stationary and nonstationary autoregressive distributed-lag processes for the latent factors are considered. For the latent factors which exhibit a common trend, and hence are cointegrated, an error correction specification of the (vector) autoregressive distributed-lag process is proposed. Full probabilistic inference for the model parameters is facilitated by adapting standard Markov chain Monte Carlo (MCMC) algorithms for dynamic linear models to our model formulation. The fit of the model is discussed for a data set of 48 States for which we model the relationship between housing prices and the macroeconomy, using State level unemployment and per capita personal income.

1. Introduction. This paper is concerned with the modeling of housing prices at the State level in the US. Housing is a massive factor in people's consumption. For industrialized nations, for example, it is the biggest component in the basket of goods used for calculating the consumer price index. Also, the Bureau of Labor Statistics has estimated in 2010 that about 24 percent of the total consumption of American home owners goes toward housing. Hence, housing is big enough to leave a sizable footprint on the economy in general.

In the generic sense, housing is also an important social institution in our society. Not only does housing play a major role in any nation's economy, but it also provides people with the social values of shelter, security, independence, privacy and amenity. The state of the current economy and recent events in the housing sector have thus led to increased attention on the role of the housing sector in the economy as a whole.

Economists have studied the relationship between the housing sector and the macroeconomy since the 1970s. Several socio-economic variables and/or real estate characteristics are traditionally considered to have an impact on housing prices and several studies have thus been dedicated to the determination of fundamental factors explaining US housing price variations. Our primary purpose here is not to comprehensively examine all these variables. In fact, there is no single generally agreed upon set of variables used in testing models of housing prices in the literature. For a complete discussion on this point see, for example, Malpezzi (1999), Capozza et al. (2002) and Gallin (2008). It is thus beyond the scope of this paper to discuss the possible roles played by all fundamental factors in explaining the variation of housing prices. Hence, for simplicity, we only examine here the extent to which these prices are driven by the real per capita disposable income and the unemployment rate.

1.1. *The data*: *A brief description*.   The data analyzed in this paper are from the St. Louis Federal Reserve Bank database[1] and the Bureau of Labor Statistics[2] and consist of quarterly time series on 48 States (excluding Alaska and Hawaii) from 1984 (first quarter) to 2011 (fourth quarter). Figure 1 shows the time series of the real housing price index for the 48 United States grouped in the eight Bureau of Economic Analysis (BEA) regions. The time series are expressed in a logarithmic scale—see Section 8 for a complete description of the data set.

Figure 1 shows that there are interesting dynamic structures in the time series and that periodic patterns and common trend components are consistent features of the housing market. Specifically, it appears that housing prices have been rising rapidly. Since 1995 we have estimated that, on average, real housing prices have increased about 36 percent, roughly double the increase of previous housing price booms observed in the late 1980s. Moreover, we notice that housing prices continued to rise strongly during the 2001 recession and that the process of the housing price boom, which some have interpreted as a bubble, started in 1998, accelerated during the period 2003–2006 and burst in 2007. The prices have then been falling sharply over all the country.

The possibility of modeling all these dynamic features, as well as to obtain accurate housing price forecasts, is important for prospective homeowners, investors, appraisers and other real estate market participants, such as mortgage lenders and insurers.

The way in which housing prices spread out to surrounding locations over time are also of interest in the real estate literature. The co-movements shown by the time series within BEA regions suggest the presence of spatial correlation. As

---

[1]http://research.stlouisfed.org/fred2/.
[2]http://stats.bls.gov/cpi/home.htm#data.
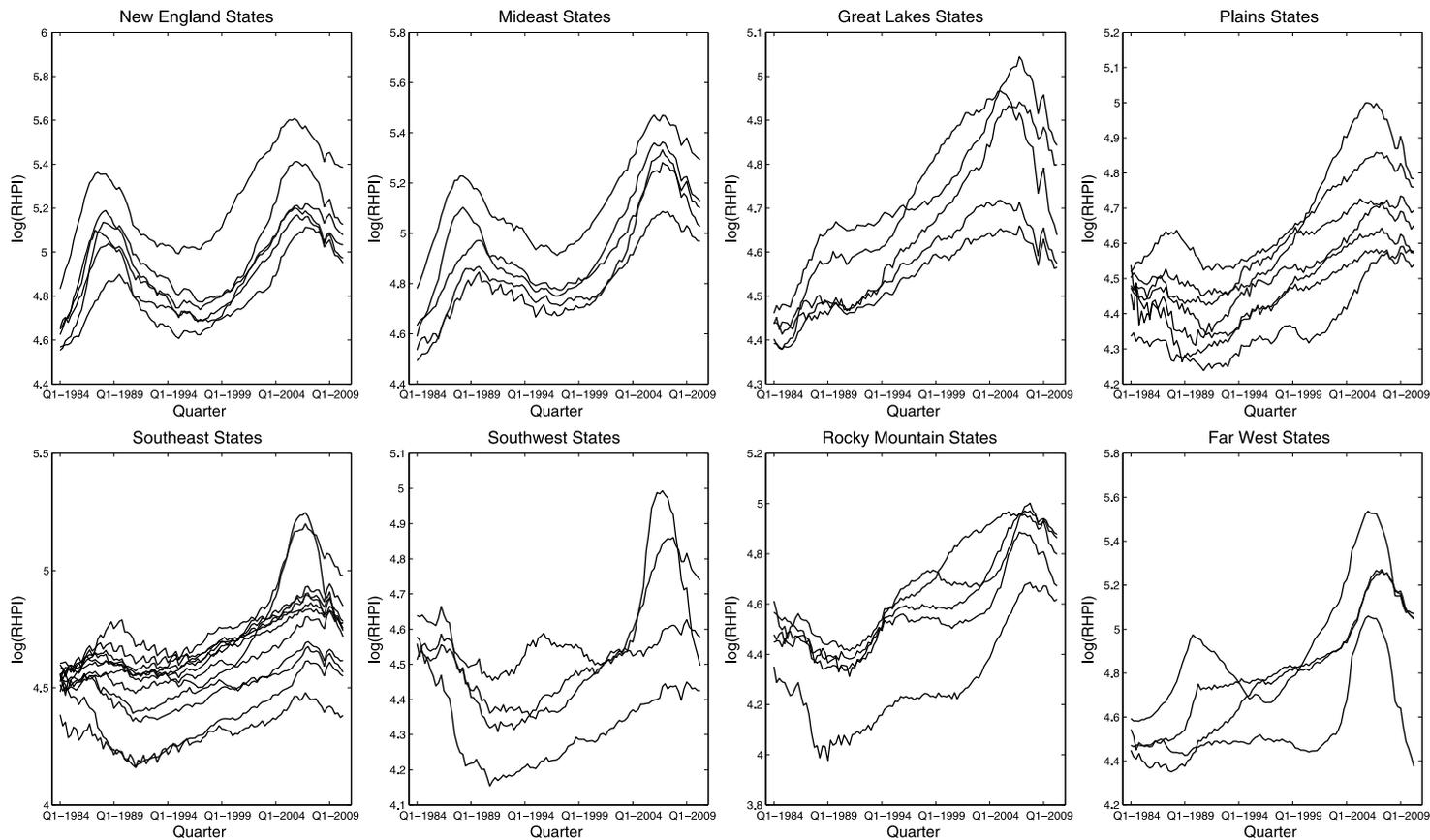
FIG. 1.   *Time series of the log-transformed real housing price index. The 48 United States are grouped in the eight Bureau of Economic Analysis (BEA) regions.*

TABLE 1
Average of correlation coefficients within and between regions first difference log of real housing prices. BEA regions: New England (NE), Mideast (ME), Great Lakes (GL), Plains (PL), Southeast (SE), Southwest (SW), Rocky Mountain (RM), Far West (FW)

	NE	ME	GL	PL	SE	SW	RM	FW
NE	0.80	–	–	–	–	–	–	–
ME	0.72	0.74	–	–	–	–	–	–
GL	0.47	0.48	0.63	–	–	–	–	–
PL	0.23	0.25	0.35	0.50	–	–	–	–
SE	0.35	0.40	0.48	0.36	0.45	–	–	–
SW	0.24	0.29	0.35	0.42	0.42	0.47	–	–
RM	0.10	0.17	0.30	0.46	0.37	0.48	0.50	–
FW	0.33	0.46	0.42	0.34	0.37	0.40	0.41	0.50

stated in Holly, Pesaran and Yamagata (2010), it is possible that States that are contiguous may influence each other's housing prices. In fact, high prices in metropolitan areas may persuade people to commute from neighboring States. Labour mobility is quite high in the USA and lower housing prices may provide an incentive to migrate. Another possible source of cross-sectional dependence would be due to economy-wide common shocks that affect all cross section units. Changes in interest rates, oil prices and technology are examples of such common shocks that may affect housing prices, although with different degrees across States.

To explore the existence of spatial interactions, using data on the growth of real housing prices, Table 1 shows the simple correlation coefficients between each State, within and between correlations for the 8 BEA regions. The diagonal elements show the within region average correlation coefficients, while the off-diagonal elements give the between region correlation coefficients. Apart from the States of the Southeast, which are more correlated on average with the States of the Great Lakes than among themselves, the within region correlation is larger than the between region correlation. In general, on average, the correlations decline with distance, but it is interesting to note the quite high correlations between the East and West regions, that is, for States belonging to the Mideast and Far West regions. In general, there is more evidence in the raw data of a possible spatial pattern in real housing prices than in real incomes and unemployment rate.

1.2. *Related literature and the proposed model.* Modeling the spatio-temporal variability of housing prices has enjoyed widespread popularity in the last years. In order to obtain a high degree of accuracy in the results, the analysis of housing prices across US States requires the definition of a general and flexible econometric model where the temporal and cross-sectional dependencies must be accommodated. Several efforts have been made to develop spatio-temporal models but there is no single approach which can be considered uniformly as being the

most appropriate. For example, time series models have become increasingly sophisticated in their treatment of dynamics and trends over time, including the application of unit roots and cointegration techniques [Giussani and Hadjimatheou (1991), Meen (2001), Muellbauer and Murphy (1997)]. However, traditional approaches, such as those based on standard vector autoregression analysis (VAR), do not allow for a direct modeling of locational spillovers and are thus not consistent with the "ripple effect" theory [Meen (1999)]. A spatial adaptation of VARs, denoted as SpVAR models, explicitly considers the potential impacts of economic events in neighboring States and has been discussed in Kuethe and Pede (2011). The SpVAR is a specific version of the Spatio-Temporal Auto-Regressive Moving Average—(STARMA)—model introduced by Pfeifer and Deutsch (1980) where the linear dependencies are lagged in both space and time. Since STARMAs are an extension of the ARMA class of models [Box, Jenkins and Reinsel (1994)], they are particularly useful to produce temporal forecasts of the variable of interest. However, the STARMA specification also suffers from some disadvantages. First, because of the amount of computational effort required, STARMAs are in general only suitable for modeling data which are dense in time and sparse in space. For example, in Kuethe and Pede (2011) the analysis is only limited to 11 States (i.e., West Region). Secondly, the understanding of co-movements among US State housing prices (and other involved variables) is difficult when the number of the States is large. Knowledge of this covariation is required both to academics seeking to explain the economic nature and sources of variation and to practitioners involved in the development of trading strategies. Thirdly, as argued by Anselin [(1988), pages 11–14], the STARMA class does not offer a fully adequate modeling of the spatial dependence and heterogeneity of observations. The lack of an adequate treatment of a simultaneous (instantaneous) spatial dependence is also the main point of criticism raised by Cressie [(1993), page 450] to the STARMA methodology. In fact, in its standard specification, STARMA implicitly assumes that, conditional on past observations, the process is uncorrelated across space. This is undoubtedly a major shortcoming, since many observed series, as noted, for example, by Pfeifer and Deutsch (1981), show considerable contemporaneous correlation even after conditioning on the past history of the process. When the contemporaneous correlation is considered by the model, the observations become a nonlinear transformation of the innovations and, as a result, maximum likelihood estimation becomes much more difficult [Elhorst (2001), Di Giacinto et al. (2005)].

Seemingly Unrelated Regression (SUR) and error correction panel data models [see, e.g., Meen (2001), Cameron, Muellbauer and Murphy (2006)] have also been largely used with spatial and time effects to investigate the evolution of housing prices. Apart from their rather complex structure, as STARMAs, these models are not suitable when the number of regions is relatively large. In fact, the application of an unrestricted SURE-GLS approach to large N (cross section dimension) and T (time series dimension) panels involves nuisance parameters that increase

at a quadratic rate as the cross section dimension of the panel is allowed to rise [Pesaran (2006)].

Recent research has found that in a data rich environment, dimension reduction in the form of factors is useful for exploratory analysis, prediction and policy analysis. Factor analysis assumes that the cross dependence can be characterized by a finite number of unobserved common factors, possibly due to economy-wide shocks that affect all States, albeit with different intensities. Thus, strong co-movement and high correlation among the series suggest that both observable and unobservable factors must be at place. The effects of common shocks on housing prices have been taken in consideration in van Dijk et al. (2011) and Holly, Pesaran and Yamagata (2010) by making use of the common correlated effects estimator [CCE, Pesaran (2006)] which controls for heterogeneity and spatial dependence. In these studies, the authors develop a panel data model where fixed mean effects, cointegration, cross-equation correlations and latent factors are considered. Furthermore, they show that by approximating the linear combinations of the unobserved factors by cross section averages of the dependent and explanatory variables, and by running standard panel regressions augmented with these cross section averages, spatial dependency can be eliminated.

Differently from these authors, we approach the analysis from the perspective of recent developments of dynamic factor models in the literature of spatio-temporal processes. We assume that the observed process can be modeled by a temporally dynamic and spatially descriptive model, hereafter referred to as the *spatial dynamic structural equation* model—SD-SEM. There are some important differences between our approach and the one discussed by Holly, Pesaran and Yamagata (2010) and van Dijk et al. (2011). Firstly, differently from these authors, we do not use cross section averages to eliminate cross-sectional dependencies. Instead, our model formulation exploits the spatio-temporal nature of the data and explicitly defines a nonseparable spatio-temporal covariance structure of the multivariate process. Secondly, because of the high dimensionality of the data, dimension reduction is important and we suggest modeling the temporal relationship between dependent and regressor variables in a latent space. The observed processes are thus described by a potentially small set of common dynamic latent factors. For all possible model candidates which may be specified, we use a multivariate autoregressive distributed-lag specification for these latent processes and, to account for situations in which two or more latent factors appear to exhibit a common trend, their cointegrating relationship is considered. Thirdly, by modeling the spatial variation via spatially structured factor loadings, we entertain the possibility of identifying clusters of States that share common time series components. This is one of the main advantages of our model formulation. Lastly, the model naturally allows for producing temporal and spatial predictions of the variables of interest. Note that although spatial interpolation is not a main task in lattice data applications, it may be an important issue in terms of missing data reconstruction

(i.e., partial or total reconstruction of the housing price time series). This problem would not be easily addressed by the other model formulations discussed above.

The SD-SEM represents a multivariate extension of the model recently proposed by Ippoliti, Gamerman and Valentini (2012) for modeling environmental coupled (correlated) spatio-temporal processes. Our spatio-temporal data are thus multivariate, in that more than one variable is typically measured at specific spatial sites (States) and different temporal instants. Furthermore, as in Lopes, Salazar and Gamerman (2008) and Ippoliti, Valentini and Gamerman (2012), we assume that the spatial dependence can be modeled through the columns of the factor loading matrices. However, differently from these authors, who refer to applications with spatially continuous (i.e., geostatistical) processes, we consider here applications with lattice data such that the factor loadings can be modeled as conditionally independent multivariate Gaussian Markov Random Fields—GMRFs. While models for multivariate geostatistical data have been extensively explored, models for lattice data have received less attention in the literature. For recent methodological developments the reader is referred to Sain and Cressie (2007), Sain, Furrer and Cressie (2011) and the references therein.

The SD-SEM is developed within a state-space framework and full probabilistic inference for the parameters is facilitated by Markov chain Monte Carlo (MCMC).

The remainder of the paper is organized as follows. In Section 2 we describe the general dynamic latent model, while in Section 3 specific attention is given to models which incorporate general forms of the spatial correlations and cross-correlations between variables at different locations. In Section 4 we describe the state-space formulation and in Section 5 discuss the nonstationary cases for the temporal dynamics of the latent factors. In Section 6 we consider Bayesian inferential issues and in Section 7 we describe forecasting strategies. In Section 8 we discuss fits of the model to the data set of US real housing prices, while Section 9 concludes the paper.

2. The spatial dynamic structural equation model.

Often observations are multivariate in nature, that is, we obtain vector responses at locations across space. For such data, we need to model both association between measurements at a location as well as association between measurements across locations. With increased collection of such multivariate spatial data, there arises the need for flexible explanatory stochastic models in order to improve estimation precision [see, e.g., Kim, Sun and Tsutakawa (2001)] and to provide simple descriptions of the complex relationships existing among the variables. In the following, a model formulation which describes the structural relations among the variables in a lower dimensional space is presented.

Assume that Y and X are two multivariate (multidimensional) spatio-temporal processes, that is, assume that several variables are measured at the node or interior (State), $\mathbf{s}$, of a lattice $\mathcal{L}$ and temporal instant $t \in \{1, 2, \ldots, T\}$. Hence, for n_y variables, we write $\mathbf{Y}(\mathbf{s}, t) = [Y_1(\mathbf{s}, t), \ldots, Y_{n_y}(\mathbf{s}, t)]'$, and the same holds for X,

for n_x variables. It is explicitly assumed that X is a predictor of Y, which is the process of interest.

Also, assume that N is the number of locations in $\mathcal{L}$ and let $\tilde{n}_y = n_y N$ and $\tilde{n}_x = n_x N$. Then, at a specific time t, the $(\tilde{n}_y \times 1)$ and $(\tilde{n}_x \times 1)$ dimensional spatial processes, Y and X, are denoted as $\mathbf{Y}(t) = [\mathbf{Y}(\mathbf{s}_1, t)', \ldots, \mathbf{Y}(\mathbf{s}_N, t)']'$ and $\mathbf{X}(t) = [\mathbf{X}(\mathbf{s}_1, t)', \ldots, \mathbf{X}(\mathbf{s}_N, t)']'$.

Our model assumes that each multivariate spatial process, at a specific time t, has the following linear structure:

$$(1) \qquad \mathbf{X}(t) = \mathbf{m}_x(t) + \mathbf{H}_x \mathbf{f}(t) + \mathbf{u}_x(t),$$

$$(2) \qquad \mathbf{Y}(t) = \mathbf{m}_y(t) + \mathbf{H}_y \mathbf{g}(t) + \mathbf{u}_y(t),$$

where $\mathbf{m}_y(t)$ and $\mathbf{m}_x(t)$ are $(\tilde{n}_y \times 1)$ and $(\tilde{n}_x \times 1)$ mean components modeling the smooth large-scale temporal variability, $\mathbf{H}_y$ and $\mathbf{H}_x$ are measurement (factor loadings) matrices of dimensions $(\tilde{n}_y \times m)$ and $(\tilde{n}_x \times l)$, respectively, and $\mathbf{g}(t)$ and $\mathbf{f}(t)$ are m- and l-dimensional vectors of temporal common factors. Also, $\mathbf{u}_y(t)$ and $\mathbf{u}_x(t)$ are Gaussian error terms for which we assume $\mathbf{u}_y(t) \sim N(\mathbf{0}, \boldsymbol{\Sigma}_{u_y})$ and $\mathbf{u}_x(t) \sim N(\mathbf{0}, \boldsymbol{\Sigma}_{u_x})$. For simplicity, throughout the paper it is assumed that $\boldsymbol{\Sigma}_{u_y}$ and $\boldsymbol{\Sigma}_{u_x}$ are both diagonal matrices and that $m \ll \tilde{n}_y$ and $l \ll \tilde{n}_x$.

The temporal dynamic of the common factors is then modeled through the following state equations:

$$(3) \qquad \mathbf{g}(t) = \sum_{i=1}^{p} \mathbf{C}_i \mathbf{g}(t-i) + \sum_{j=1}^{q} \mathbf{D}_j \mathbf{f}(t-j) + \boldsymbol{\xi}(t),$$

$$(4) \qquad \mathbf{f}(t) = \sum_{k=1}^{s} \mathbf{R}_k \mathbf{f}(t-k) + \boldsymbol{\eta}(t),$$

where $\mathbf{C}_i$ $(m \times m)$, $\mathbf{D}_j$ $(m \times l)$, and $\mathbf{R}_k$ $(l \times l)$ are coefficient matrices modeling the temporal evolution of the latent vectors $\mathbf{g}(t) = [g_1(t), \ldots, g_m(t)]'$ and $\mathbf{f}(t) = [f_1(t), \ldots, f_l(t)]'$, respectively. Finally, $\boldsymbol{\xi}(t)$ and $\boldsymbol{\eta}(t)$ are independent Gaussian error terms for which we assume $\boldsymbol{\xi}(t) \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\xi)$ and $\boldsymbol{\eta}(t) \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\eta)$.

Equation (3) represents a Vector Autoregressive model with exogenous variables (VARX) where the variables in $\mathbf{g}(t)$, considered as endogenous (i.e., determined within the system), are controlled for the effects of other variables, $\mathbf{f}(t)$, considered as exogenous (i.e., determined outside the system and treated independently of the other variables)[3]. Equations (1)–(4) thus provide the basic formulation of the SD-SEM. One advantage of this model is that temporal forecasts of the variable of interest, Y, can be obtained by modeling the dynamics of a few

---

[3]The distinction between "exogenous" and "endogenous" variables in a model is subtle and is a subject of a long debate in the literature. See, for example, Engle, Hendry and Richard (1983), Osiewalski and Steel (1996). Gourieroux and Monfort [(1997), Chapter 10] also provide a clear distinction between the different exogeneity concepts.

common factors. Also, the model is spatially descriptive in that it can be used to identify possible clusters of locations whose temporal behavior is primarily described by a potentially small set of common dynamic latent factors. As it will be shown in the next section, flexible and spatially structured prior information regarding such clusters can be specified through the columns of the factor loading matrix.

3. Factor loadings and multivariate GMRFs.
A key property of much spatio-temporal data is that observations at nearby sites and times will tend to be similar to one another. This underlying smoothness characteristic of a space–time process can be captured by estimating the state process and filtering out the measurement noise. It is customary for dynamic latent models to refer to the unobserved (state) processes as the common factors and to refer to the coefficients that link the factors with the observed series as the factor loadings. It is assumed that these factor loadings have the nature of spatial processes and, extending results in Ippoliti, Valentini and Gamerman (2012), here the spatial dependence is modeled through a multivariate GMRF. Relevant papers useful for our purposes are Mardia (1988) and Sain and Cressie (2007), and we refer to them for known results on the model formulation.

Let $\mathbf{h}_{x_j} = [\mathbf{h}_{x_j}(\mathbf{s}_1)', \mathbf{h}_{x_j}(\mathbf{s}_2)', \ldots, \mathbf{h}_{x_j}(\mathbf{s}_N)']'$, that is, the jth column of $\mathbf{H}_x$, be a $\tilde{n}_x$-dimensional spatial process observed on $\mathcal{L}$—and similarly for $\mathbf{H}_y$. Also, let $[\mathbf{h}_{x_j}(\mathbf{s}_i)|R_{-i}]$ denote the conditional distribution of $\mathbf{h}_{x_j}(\mathbf{s}_i)$ given the rest (i.e., values at all other sites). Then, the GMRF is defined by the conditional mean

$$(5) \qquad \mathrm{E}\big(\mathbf{h}_{x_j}(\mathbf{s}_i)|R_{-i}\big) = \boldsymbol{\mu}_i^{(h_{x_j})} + \sum_{u \in \mathcal{S}_i} \mathbf{F}_{iu}^{(h_{x_j})}\big(\mathbf{h}_{x_j}(\mathbf{s}_u) - \boldsymbol{\mu}_u^{(h_{x_j})}\big)$$

and the conditional covariance matrix

$$(6) \qquad \mathrm{Var}\big(\mathbf{h}_{x_j}(\mathbf{s}_i)|R_{-i}\big) = \mathbf{T}_i^{(h_{x_j})},$$

where $\mathcal{S}_i$ is a finite subset of $\mathcal{L}$ containing neighbors of site $\mathbf{s}_i$, $\boldsymbol{\mu}_i^{(h_{x_j})}$ is a n_x-dimensional mean vector, and $\mathbf{F}_{iu}^{(h_{x_j})}$ is a $(n_x \times n_x)$ matrix of spatial regression parameters.

To take into account the effect of some explanatory variables, it is possible to parameterize the mean vector through the definition of a $(N \times q)$ design matrix, $\mathcal{D}^*$, such that $\boldsymbol{\mu}^{(h_{x_j})} = \mathcal{D}^*\boldsymbol{\beta}^{(h_{x_j})}$, with $\boldsymbol{\beta}^{(h_{x_j})}$ a $(q \times 1)$ vector of parameters. Assuming $\mathbf{c}_i$ is a vector of covariates for the ith location, we have $\boldsymbol{\mu}^{(h_{x_j})} = [\boldsymbol{\mu}_1^{(h_{x_j})'}, \ldots, \boldsymbol{\mu}_N^{(h_{x_j})'}]'$, with $\boldsymbol{\mu}_i^{(h_{x_j})} = \mathcal{D}_i^*\boldsymbol{\beta}_i^{(h_{x_j})}$, $\boldsymbol{\beta}^{(h_{x_j})} = [\boldsymbol{\beta}_1^{(h_{x_j})'}, \ldots, \boldsymbol{\beta}_{n_x}^{(h_{x_j})'}]'$, $\mathcal{D}_i^* = (\mathbf{I}_{n_x} \otimes \mathbf{c}_i')$, $i = 1, \ldots, n$, and $\otimes$ denoting the Kronecker product. For a discussion of different specifications of the matrix $\mathcal{D}^*$, see, for example, Ippoliti,

Valentini and Gamerman (2012) and Lopes, Salazar and Gamerman (2008). However, due to the static behavior of $\mathbf{h}_{x_j}$, only spatially-varying covariates will be considered in explaining the mean level of the GMRF.

With the definition of the conditional distributions, it follows [see Mardia (1988)] that the joint distribution of $\mathbf{h}_{x_j}$ is MVN($\boldsymbol{\mu}^{(h_{x_j})}, \boldsymbol{\Sigma}^{(h_{x_j})}$) with the covariance matrix specified as $\boldsymbol{\Sigma}^{(h_{x_j})} = \{\text{block}[-\mathbf{T}_i^{(h_{x_j})^{-1}} \mathbf{F}_{iu}^{(h_{x_j})}]\}^{-1}$, where $\mathbf{F}_{ii} = -\mathbf{I}$ and for a generic matrix $\mathbf{G}$, block$[\mathbf{G}_{iu}]$ denotes a block matrix with the (i, u)th block given by $\mathbf{G}_{iu}$ [see Sain and Cressie (2007)]. To guarantee that a proper probability density function is defined, the parametrization must ensure that $\boldsymbol{\Sigma}^{(h_{x_j})}$ is positive-definite and symmetric; hence, we require both $\mathbf{F}_{iu}^{(h_{x_j})} \mathbf{T}_u^{(h_{x_j})} = \mathbf{T}_i^{(h_{x_j})} \mathbf{F}_{ui}^{(h_{x_j})'}$ and block$[-\mathbf{T}_i^{(h_{x_j})^{-1}} \mathbf{F}_{iu}^{(h_{x_j})}]$ positive definite.

4. The state space formulation. As shown in Section 2, the temporal dynamic is modeled through the state equations (3) and (4). The specification of equation (4) is necessary to predict in time the latent process $\mathbf{f}(t)$ and thus to obtain k-step ahead forecasts of $\mathbf{g}(t)$ through equation (3). It is thus useful to specify the joint generation process for $\mathbf{g}(t)$ and $\mathbf{f}(t)$ as

$$
(7) \quad \begin{bmatrix} \mathbf{g}(t) \\ \mathbf{f}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{C}_1 & \mathbf{D}_1 \\ \mathbf{0} & \mathbf{R}_1 \end{bmatrix} \begin{bmatrix} \mathbf{g}(t-1) \\ \mathbf{f}(t-1) \end{bmatrix} + \cdots \\
+ \begin{bmatrix} \mathbf{C}_p & \mathbf{D}_p \\ \mathbf{0} & \mathbf{R}_p \end{bmatrix} \begin{bmatrix} \mathbf{g}(t-p) \\ \mathbf{f}(t-p) \end{bmatrix} + \begin{bmatrix} \boldsymbol{\xi}(t) \\ \boldsymbol{\eta}(t) \end{bmatrix},
$$

where it is assumed without loss of generality that $p \geq \max(s, q)$, $\mathbf{D}_i = \mathbf{0}$ for $i > q$ and $\mathbf{R}_j = \mathbf{0}$ for $j > s$. It follows that the joint generation process of $\mathbf{g}(t)$ and $\mathbf{f}(t)$ is a VAR(p) process of the type

$$
(8) \quad \mathbf{d}(t) = \boldsymbol{\Phi}_1 \mathbf{d}(t-1) + \cdots + \boldsymbol{\Phi}_p \mathbf{d}(t-p) + \boldsymbol{\varepsilon}(t),
$$

where

$$
\mathbf{d}(t) = \begin{bmatrix} \mathbf{g}(t) \\ \mathbf{f}(t) \end{bmatrix}, \qquad \boldsymbol{\Phi}_i = \begin{bmatrix} \mathbf{C}_i & \mathbf{D}_i \\ \mathbf{0} & \mathbf{R}_i \end{bmatrix}, \qquad \boldsymbol{\varepsilon}(t) = \begin{bmatrix} \boldsymbol{\xi}(t) \\ \boldsymbol{\eta}(t) \end{bmatrix}.
$$

The presence of the measurement and the state variables naturally leads to the state-space representation [Lutkepohl (2005)] of the SD-SEM model; given the data, this representation allows for a recursive estimate of the latent variables through the Kalman filter algorithm. The linear Gaussian state-space model is thus described by the following *state* and *measurement* equations:

$$
(9) \quad \boldsymbol{\alpha}(t) = \boldsymbol{\Phi}\boldsymbol{\alpha}(t-1) + \boldsymbol{\Xi}\boldsymbol{\zeta}(t),
$$

$$
(10) \quad \mathbf{z}(t) = \mathbf{H}\boldsymbol{\alpha}(t) + \mathbf{u}(t),
$$

where $\boldsymbol{\alpha}(t)$ is the state vector, $\boldsymbol{\Phi}$ is the nonsingular transition matrix, $\boldsymbol{\Xi}$ is a constant input matrix, $\mathbf{z}(t)$ is the measurement vector and $\mathbf{H}$ is the measurement

matrix. The sequences $\boldsymbol{\zeta}(t)$ and $\mathbf{u}(t)$ are assumed to be mutually independent zero mean Gaussian random variables with covariances $E\{\boldsymbol{\zeta}(t_i)\boldsymbol{\zeta}(t_j)'\} = \boldsymbol{\Psi}\delta_{ij}$ and $E\{\mathbf{u}(t_i)\mathbf{u}(t_j)'\} = \boldsymbol{\Sigma}_u\delta_{ij}$, where $E\{\cdot\}$ denotes the expectation and δ_{ij} the Kronecker delta function. In (9) and (10) we have the following specification:

$$\boldsymbol{\alpha}(t) = \begin{bmatrix} \mathbf{d}(t) \\ \mathbf{d}(t-1) \\ \vdots \\ \mathbf{d}(t-p+1) \end{bmatrix}, \qquad \boldsymbol{\Phi} = \begin{bmatrix} \boldsymbol{\Phi}_1 & \boldsymbol{\Phi}_2 & \cdots & \boldsymbol{\Phi}_p \\ \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \cdots & \mathbf{I} & \mathbf{0} \end{bmatrix},$$

$$\boldsymbol{\zeta}(t) = \begin{bmatrix} \boldsymbol{\varepsilon}(t) \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}, \qquad \mathbf{z}(t) = \begin{bmatrix} \mathbf{y}(t) \\ \mathbf{x}(t) \end{bmatrix}, \qquad \mathbf{H} = \begin{bmatrix} \mathbf{H}_y & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_x & \cdots & \mathbf{0} \end{bmatrix},$$

$$\boldsymbol{\Xi} = \begin{bmatrix} \mathbf{I} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}, \qquad \mathbf{u}(t) = \begin{bmatrix} \mathbf{u}_y(t) \\ \mathbf{u}_x(t) \end{bmatrix}.$$

5. Nonstationary latent factors. The dynamic specification for the state vector $\boldsymbol{\alpha}(t)$ is quite general. In fact, the family of time series processes that can be formulated as in equations (9) and (10) is wide and includes a broad range of nonstationary time series processes. Sometimes it may be advantageous to have a specification that decomposes the latent factors into stationary and nonstationary components, such as trend, periodic or cyclical components. The large scale dynamic components can in fact be directly specified through the common dynamic factors. In this case, for example, common seasonal factors can receive different weights for different columns of the factor loading matrix, so allowing different seasonal patterns for the spatial locations. For some specific examples, and for a wider discussion on this point, see Lopes, Salazar and Gamerman (2008) and Ippoliti, Valentini and Gamerman (2012).

5.1. *Cointegrated latent factors.* Nonstationarity can also occur when two or more latent factors appear to exhibit a common trend, and hence are cointegrated [Johansen (1988)]. In this case we have that one or more linear combinations of these factors are stationary even though individually they are not. If the factors are cointegrated, they cannot move too far away from each other and we should observe a stable long-run relationship among their levels. In contrast, a lack of cointegration suggests that such factors have no long-run link and, in principle, they can wander arbitrarily far away from each other.

In our model formulation we consider the case in which the exogenous factors are cointegrated among themselves as well as with the endogenous latent variables.

In this case the vector autoregressive process of equation (8) can be written in the error correction model (ECM) form as

$$(11) \qquad \Delta \mathbf{d}(t) = \tilde{\mathbf{A}} \mathbf{d}(t-1) + \sum_{i=1}^{p-1} \tilde{\boldsymbol{\Phi}}_i \Delta \mathbf{d}(t-i) + \boldsymbol{\varepsilon}(t),$$

where $\tilde{\mathbf{A}} = -\mathbf{I} + \sum_{i=1}^{p} \boldsymbol{\Phi}_i$, $\tilde{\boldsymbol{\Phi}}_i = -\sum_{j=i+1}^{p} \boldsymbol{\Phi}_j$ and Δ is the difference operator, that is, $\Delta \mathbf{d}(t) = \mathbf{d}(t) - \mathbf{d}(t-1)$. Full details of the vector error correction specification of equation (11) are provided in Appendix A where we also show that the matrix of long-run multipliers, $\tilde{\mathbf{A}}$, is an upper block triangular matrix. These single blocks, expressed as a product of parameter matrices, provide information about: (i) the cointegration structure *within* the exogenous and endogenous processes $\mathbf{f}(t)$ and $\mathbf{g}(t)$, and (ii) the cointegration *between* the two processes.

6. Inference and computations.

6.1. *Prior information.* Full probabilistic inference for the model parameters is carried out based on the following independent prior distributions. Throughout we shall use vec$(\cdot)$ to denote the vec operator and $G(a, b)$ to denote the Gamma distribution with mean a/b and variance a/b^2. Unless explicitly needed, full specifications of the priors are only given for X so that definitions for Y follow accordingly.

MEASUREMENT EQUATION. The precision matrix $\boldsymbol{\Sigma}_{u_x}^{-1}$ is assumed to be diagonal where each element has a Gamma prior distribution, $G(0.01, 0.01)$.

The prior distribution for $\boldsymbol{\beta}^{(h_{x_i})}$ $(i = 1, \ldots, l)$ is $N(\mathbf{0}, \sigma_\beta^2 \mathbf{I})$. Then, assuming a constant conditional covariance matrix, the prior on the inverse covariance matrix $\mathbf{T}^{(h_{x_i})^{-1}}$ is given by the Wishart distribution [Mardia, Kent and Bibby (1979)], that is, $\mathbf{T}^{(h_{x_i})^{-1}} \sim W(\varrho_x, (\varrho_x \mathbf{S}_x)^{-1})$, where $\varrho_x > l$ and $\mathbf{S}_x$ is a pre-specified symmetric positive definite matrix. To provide the prior specification for the joint distribution of the spatial regression parameters, we set $\mathbf{F}_{iu}^{(h_{x_i})} = \mathbf{F}^{(h_{x_i})}$ and, following Sain and Cressie (2007), we use the reparametrization $\tilde{\mathbf{F}}^{(h_{x_i})} = \mathbf{T}^{(h_{x_i})^{-1/2}} \mathbf{F}^{(h_{x_i})'} \mathbf{T}^{(h_{x_i})^{1/2}}$ and specify its prior to be proportional to $\exp\{-\boldsymbol{v}'\boldsymbol{v}/\varsigma^2\}$, where $\boldsymbol{v} = \text{vec}(\tilde{\mathbf{F}}^{(h_{x_i})'})$. The prior parameter ς is specified by choosing small values, since the prior for $\tilde{\mathbf{F}}^{(h_{x_i})}$ is concentrated around zero. Then, in both mean and variance of the GMRF processes we adopt priors centered around prefixed values, as defined in Section 3.

STATE EQUATION. When stationarity conditions are met for the latent processes the prior distributions for the state equation coefficients can be specified as proposed in Lopes, Salazar and Gamerman (2008). For the cointegration case, since the formulation given in equation (11) is quite general, and many plausible restricted models can be envisaged, Stochastic Search Variable Selection (SSVS) priors [see Jochmann et al. (2013)] are used for the parameters of the state equations. Note that these plausible models may differ in the choice of the restrictions

on the cointegration space, the number of exogenous and endogenous latent variables, and the lag length allowed for the autoregression.

The error covariance matrices are assumed to be decomposed as $\boldsymbol{\Sigma}_\xi^{-1} = \mathbf{V}_\xi \mathbf{V}_\xi'$ and $\boldsymbol{\Sigma}_\eta^{-1} = \mathbf{V}_\eta \mathbf{V}_\eta'$, where $\mathbf{V}_\xi$ and $\mathbf{V}_\eta$ are upper-triangular matrices. Then, the SSVS priors involve using a standard Gamma prior for the square of each of the diagonal elements of $\mathbf{V}_{(\cdot)}$ and the SSVS mixture of normals prior for each element above the diagonal [George, Sun and Ni (2008)]. Note that if the error covariance matrices are chosen to be diagonal, then the computation of the posterior simplifies considerably.

Since $\tilde{\mathbf{A}}$ is potentially of reduced rank and crucial issues of identification may arise in the ECM form, linear identifying restrictions are usually imposed. However, because of local identifiability problems and the restriction on the estimable region of the cointegrating space [Koop et al. (2006)], the so-called *linear normalization* approach also suffers from several drawbacks. To overcome these problems, we thus adopt the SSVS approach proposed by Jochmann et al. (2013) which, defining priors on the cointegration space, is facilitated by the computation of Gaussian posterior conditional distributions [Koop, Leon-Gonzalez and Strachan (2010)]. A brief summary of the SSVS priors used in this paper is provided in Appendix B. For a more complete description, the reader is referred to Jochmann et al. (2013) and Koop, León-González and Strachan (2010).

Finally, the prior for the latent process $\boldsymbol{\alpha}(t)$ is provided by the transition equation and is completed by $\boldsymbol{\alpha}(0) \sim N(\mathbf{a}_0, \boldsymbol{\Sigma}_{\alpha 0})$, for known hyperparameters $\mathbf{a}_0$ and $\boldsymbol{\Sigma}_{\alpha 0}$ [Durbin and Koopman (2001), Rosenberg (1973)].

6.2. *The likelihood function.* To specify the likelihood function, without loss of generality, it will be assumed that $\mathbf{m}_y(t) = \mathbf{0}$ and $\mathbf{m}_x(t) = \mathbf{0}$. Conditional on $\boldsymbol{\alpha}(t)$, for $t = 1, \ldots, T$, the SD-SEM model can be rewritten as $\mathbf{Z} = \boldsymbol{\alpha}\mathbf{H}' + \mathbf{U}$, where $\mathbf{Z} = [\mathbf{z}(1), \ldots, \mathbf{z}(T)]'$ and $\boldsymbol{\alpha} = [\boldsymbol{\alpha}(1), \ldots, \boldsymbol{\alpha}(T)]'$. The error matrix, $\mathbf{U}$, is of dimension $(T \times n)$, where $n = \tilde{n}_x + \tilde{n}_y$, and follows a matrix-variate normal distribution, that is, $\mathbf{U} \sim N(\mathbf{0}, \mathbf{I}_T, \boldsymbol{\Sigma}_u)$—see Dawid (1981) and Brown, Vannucci and Fearn (1998). Then the deviance, minus twice the log-likelihood is

$$\mathfrak{D}(\mathbf{z}|\boldsymbol{\Theta}, \boldsymbol{\Sigma}_u, \mathbf{H}, \boldsymbol{\alpha}, m, l)$$
$$= Tn \log(2\pi) + T \log|\boldsymbol{\Sigma}_u| + \text{trace}\{\boldsymbol{\Sigma}_u^{-1}(\mathbf{Z} - \boldsymbol{\alpha}\mathbf{H}')'(\mathbf{Z} - \boldsymbol{\alpha}\mathbf{H}')\},$$

where $\boldsymbol{\Theta}$ is the full set of model parameters.

6.3. *Posterior inference.* Posterior inference for the proposed class of spatial dynamic factor models is facilitated by MCMC algorithms. Standard MCMC for dynamic linear models are adapted to our model specification such that, conditional on l and m, posterior and predictive analysis are readily available. In the following, we provide some information on the relevant conditional distributions.

By denoting with "u" the suffix for the unobserved data, posterior inference is based on summarizing the joint posterior distribution $p(\mathbf{Z}^u, \boldsymbol{\Theta}, \boldsymbol{\alpha}(0), \boldsymbol{\alpha}|\mathbf{Z})$.

The common factors are jointly sampled by means of the well-known forward filtering backward sampling (FFBS) algorithm [Carter and Kohn (1994), Frühwirth-Schnatter (1994)]. All other full conditional distributions are "standard" multivariate Gaussian or Gamma distributions. An exception is for the spatial parameter matrices, $\tilde{\mathbf{F}}^{(h_{y_i})}$ and $\tilde{\mathbf{F}}^{(h_{x_i})}$, and the covariance matrices, $\mathbf{T}^{(h_{y_i})^{-1}}$ and $\mathbf{T}^{(h_{x_i})^{-1}}$, which are sampled using a Metropolis–Hastings step. Specific details for the implementation of the full conditional distributions can be found in Lopes, Salazar and Gamerman (2008), Sain and Cressie (2007) and Jochmann et al. (2013).

6.4. *Model identification.* Some restrictions on $\mathbf{H}_y$ and $\mathbf{H}_x$ are needed to define a unique model free from identification problems. Several possibilities can be considered and the solution adopted here is to constrain the measurement matrices so that they are lower triangular, assumed to be of full rank. We note here that we have proper but quantitatively vague priors which can lead to posteriors that are computationally indistinguishable from improper ones with the consequence of an MCMC convergence failure. Hence, to avoid relying so strongly on the prior specification, we prefer to focus on models which are identified in a frequentist sense. The approach is fully discussed in Ippoliti, Valentini and Gamerman (2012) and Strickland et al. (2011).

A critical comment to be borne in mind is that the chosen order of the univariate time series in the measurement vector influences interpretation of the factors and may impact on model fitting and assessment, the interpretation of factors if such is desired, and the choice of the number of factors. In such cases, the ordering becomes a modeling decision to be made on substantive grounds, rather than an empirical matter to be addressed on the basis of model fit. However, from the viewpoint of forecasting the ordering is irrelevant. For a detailed discussion on these points see, for example, Lopes and West (2004).

6.5. *Model selection.* With this class of model, an important issue is the selection of m and l. Several Bayesian selection methods have been developed and for a discussion, see, for example, Section 4.1 in Lopes, Salazar and Gamerman (2008). Here, we consider a simple approach which only considers the variable of interest, Y, and that consists in the minimization of the following predictive model choice statistic [PMCC, Gelfand and Ghosh (1998)]:

$$\text{PMCC} = \frac{\zeta}{\zeta + 1} G + P,$$

where, for our proposed model, $G = \sum_{i,t} (\mathbf{Y}(\mathbf{s}_i, t) - E[\mathbf{Y}(\mathbf{s}_i, t)_{\text{rep}}])^2$ and $P = \sum_{i,t} \text{Var}[\mathbf{Y}(\mathbf{s}_i, t)_{\text{rep}}]$.

This statistic is based on replicates, $\mathbf{Y}(\mathbf{s}, t)_{\mathrm{rep}}$, of the observed data and the summation is taken over $i = 1, \ldots, N$, and $t = 1, \ldots, T$. Essentially, the PMCC quantifies the fit of the model by comparing features of the posterior predictive distribution, $p(\mathbf{Y}(\mathbf{s}, t)_{\mathrm{rep}}|\mathbf{Y}(\mathbf{s}, t))$, to equivalent features of the observed data. The quantity G is a measure of goodness of fit while P is a penalty term. As the models become increasingly complex the goodness-of-fit term will decrease but the penalty term will begin to increase. Overfitting of model results in large predictive variances and large values of the penalty function. The choice of ζ determines how much weight is placed on the goodness-of-fit term relative to the penalty term. As ζ goes to infinity, equal weight is placed on these two terms. Banerjee, Carlin and Gelfand (2004) mention that ordering of models is typically insensitive to the choice of ζ, therefore, we fix $\zeta = \infty$. Notice that at each iteration of the MCMC we can obtain replicates of the observations given the sampled values of the parameters.

7. Uses of the model. In this section we provide specific details on how to obtain temporal forecasts of the variable of interest Y.

7.1. *Unconditional forecasting.* Temporal forecasts of the variable Y are directly obtained through the state space formulation of the model. In fact, it is easy to show that since $\boldsymbol{\alpha}(t)|\boldsymbol{\alpha}(t-1) \sim N(\boldsymbol{\Phi}\boldsymbol{\alpha}(t-1), \boldsymbol{\Sigma}_\alpha)$, the k-step ahead forecast for the dynamic factors is given by $p(\boldsymbol{\alpha}(t+k)|\boldsymbol{\Theta}) \sim N(\boldsymbol{\Phi}^{(k)}\boldsymbol{\alpha}(t), \boldsymbol{\Omega}^{(k)})$, where $\boldsymbol{\Omega}^{(k)} = \sum_{j=1}^{k} \boldsymbol{\Phi}^{(k-j)} \boldsymbol{\Sigma}_\alpha \boldsymbol{\Phi}^{(k-j)'}$. Therefore, the k-step ahead predictive density, $p(\mathbf{z}(t+k)|\mathbf{Z})$, of the joint process $\mathbf{Z} = [Y\,X]$ is given by

$$p(\mathbf{z}(t+k)|\mathbf{Z}) = \int p(\mathbf{z}(t+k)|\boldsymbol{\alpha}(t+k), \mathbf{H}, \boldsymbol{\Theta}) p(\boldsymbol{\alpha}(t+k)|\boldsymbol{\alpha}(t), \mathbf{H}, \boldsymbol{\Theta})$$

$$\times \, p(\boldsymbol{\alpha}(t), \mathbf{H}, \boldsymbol{\Theta}|\mathbf{Z}) \, d\boldsymbol{\alpha}(t+k) \, d\boldsymbol{\alpha}(T) \, d\mathbf{H} \, d\boldsymbol{\Theta}.$$

Draws from $p(\mathbf{z}(t+k)|\mathbf{Z})$ can be obtained in three steps. Firstly, $\boldsymbol{\Theta}$ is sampled from its joint posterior distribution via MCMC. Secondly, conditionally on $\boldsymbol{\Theta}$, the common factors $\boldsymbol{\alpha}(t+k)$ are independent of $\mathbf{Z}$ and can be sampled from $p(\boldsymbol{\alpha}(t+k)|\boldsymbol{\Theta})$. Thirdly, $\mathbf{z}(t+k)$ is sampled from $p(\mathbf{z}(t+k)|\boldsymbol{\alpha}(t+k), \mathbf{H}, \boldsymbol{\Theta})$.

7.2. *Conditional forecasting.* The forecasting procedure described above is obtained under the hypothesis that the predictor X is unknown for the period of interest. However, quite flexible forecasts can also be obtained conditional on the potential future paths of specified variables in the model. In fact, it may happen that some of the future values of certain variables are known, because data on these variables are released earlier than data on the other variables. By incorporating the knowledge of the future path of the X variable, in principle, it should be possible to obtain more reliable forecasts of Y.

Another use of conditional forecasting is the generation of forecasts conditional on different "policy/exploratory" scenarios. These scenario-based conditional forecasts allow one to answer the question: *if something happens to X in the future, how will it affect forecasts of Y in the future?* Hence, a plurality of plausible alternative futures for X can be considered and temporal forecasts of $\mathbf{g}(t)$ can be produced conditional on a specific path of $\mathbf{f}(t)$. Under these assumptions, in the following, we propose a simple procedure to obtain $\mathbf{g}(T + k)$ given $\mathbf{f}(T + 1), \ldots, \mathbf{f}(T + k)$, and all present and past information, thus avoiding the use of equation (4) to obtain k-step ahead forecasts of $\mathbf{f}(t)$.

Suppose that for the period $T + 1, T + 2, \ldots, T + k$, X is known (or fixed a priori) and that $\mathbf{X}_k = [\mathbf{x}(T + 1), \mathbf{x}(T + 2), \mathbf{x}(T + k)]$. Then, k-step ahead forecasts of $\mathbf{g}(t)$ may be obtained conditional on $\mathbf{f}_k = [\mathbf{f}(T + 1), \mathbf{f}(T + 2), \ldots, \mathbf{f}(T + k)]$, where $\mathbf{f}_k = \mathbf{H}_x^\dagger \mathbf{X}_k$ and $\mathbf{H}_x^\dagger$ is the Moore–Penrose pseudo-inverse of $\mathbf{H}_x$.

Finally, note that although spatial interpolation is not a main task in lattice data applications, the reconstruction of missing data (i.e., partial or total reconstruction of the multivariate time series of one—or more—State) is an important issue in general. This can be simply done by exploiting the conditional expectation of the GMRF and following Section 6.2 in Ippoliti, Valentini and Gamerman (2012).

8. Spatio-temporal analysis of US housing prices. Public policy interventions in housing markets are widespread and a key question is the extent to which these policies achieve their desired objectives and whether there are any unintended consequences. Especially for its relationship with mortgage behavior, in recent years, real housing prices have been of great concern for many financial institutions. Understanding the impact of specific factors on real housing prices is thus of great interest for governments, real estate developers and investors. In this paper, we examine if the total personal income (TPI) and the unemployment rate (UR) have some impact on the housing price index (HPI). The data, introduced in Section 1.1, consist of quarterly time series on 48 States (excluding Alaska and Hawaii) from 1984 (first quarter) to 2011 (fourth quarter). However, in this study, the last 10 quarters have been excluded from the estimation procedure and used only for forecast purposes.

In order to consider per capita personal income (PCI), the annual population series (U.S. Census Bureau) is converted into a quarterly series through geometric interpolation. Moreover, we consider real per capita personal income (RPCI) and housing price index (RHPI) dividing PCI and HPI by a State level general price index. However, since there is no US State level consumer price index (CPI), following Holly, Pesaran and Yamagata (2010), we have constructed a State level general price index based on the CPIs of the cities or areas. All the variables are analyzed on a logarithmic scale. Henceforth, the variables are denoted as $Y = \log(\text{RHPI})$, $X_1 = \log(\text{RPCI})$ and $X_2 = \log(\text{UR})$.

MODEL SPECIFICATION: *Measurement equations*. To provide a full specification of the inverse covariance matrix of each factor loading, we make use of a

contiguity or *adjacency* matrix $\mathbf{W}$. We assume here that $\mathbf{W}$ has zero diagonal elements and nonnegative off-diagonal elements which reflect the dependency between States $\mathbf{s}_i$ and $\mathbf{s}_j$—that is, the neighborhood set $\mathcal{S}_i$. Hence, to postulate plausible relationships between two States, as in Holly, Pesaran and Yamagata (2010), we assume that $\mathbf{W}$ is a binary proximity matrix which assigns uniform weights to all neighbors of State $\mathbf{s}_i$, that is,

$$\{\mathbf{W}\}_{i,j} = \begin{cases} 1, & \text{if States } \mathbf{s}_i \text{ and } \mathbf{s}_j \text{ share a common border,} \\ 0, & \text{otherwise.} \end{cases}$$

Then, since the general model described in Section 3 is overparameterized, it is necessary to impose some parameter restrictions. For example, because Y is univariate (i.e., $n_y = 1$), each column of $\mathbf{H}_y$ (i.e., $\mathbf{h}_{y_j}$) is treated as a univariate GMRF with conditional mean

$$\mathrm{E}\big[h_{y_j}(\mathbf{s}_i)|\mathbf{R}_{-i}\big] = \mu_i^{(h_{y_j})} + \theta_{h_{y_j}} \sum_{u \in \mathcal{S}_i} \big(h_{y_j}(\mathbf{s}_u) - \mu_u^{(h_{y_j})}\big)$$

and conditional variance

$$\mathrm{VAR}\big[h_{y_j}(\mathbf{s}_i)|\mathbf{R}_{-i}\big] = \psi^{(h_{y_j})^2}.$$

On the other hand, since X is a bivariate process—that is, $n_x = 2$ and $\mathbf{X}(\mathbf{s}, t) = [X_1(\mathbf{s}, t), X_2(\mathbf{s}, t)]'$—we assume that for $i, u = 1, \ldots, n$, $\mathbf{T}_i^{(h_{x_j})} = \mathbf{T}^{(h_{x_j})}$ is a (2×2) conditional covariance matrix and

$$\mathbf{F}^{(h_{x_j})} = \mathbf{F}_{iu}^{(h_{x_j})} = -\begin{bmatrix} \theta_{x_1}^{(j)} & \theta_{x_1,x_2}^{(j)} \\ \theta_{x_2,x_1}^{(j)} & \theta_{x_2}^{(j)} \end{bmatrix}.$$

Hence, the covariance matrix can be written as

$$\mathbf{\Sigma}^{(h_{x_j})} = \big(\mathbf{I}_N \otimes \mathbf{T}^{(h_{x_j})^{1/2}}\big)\big[\mathbf{I}_{\tilde{n}_x} + \mathbf{W}^U \otimes \tilde{\mathbf{F}}^{(h_{x_j})} + \mathbf{W}^L \otimes \tilde{\mathbf{F}}^{(h_{x_j})'}\big]^{-1}\big(\mathbf{I}_N \otimes \mathbf{T}^{(h_{x_j})^{1/2}}\big),$$

where $\mathbf{W}^U$ and $\mathbf{W}^L$ denote the upper- and lower-triangular parts of $\mathbf{W}$, respectively. Conditions for which $\mathbf{\Sigma}^{(h_{x_j})}$ is positive definite depend on the parameter space of the spatial interaction parameters in $\mathbf{F}^{(h_{x_j})}$. However, restricting $\mathbf{\Sigma}^{(h_{x_j})^{-1}}$ to be strictly diagonally dominant or adding a penalty if some of the eigenvalues are negative will ensure positive definiteness [for a discussion on this point see Sain and Cressie (2007)].

Since interpreting the spatial parameters in $\mathbf{F}^{(h_{x_j})}$ requires some care, more information on the impact of the choice of $\mathbf{F}^{(h_{x_j})}$ can be obtained by examining the conditional covariance of two neighboring locations (given the rest)

$$\mathbf{\Sigma}_{iu|-iu}^{h_{x_j}} = \begin{bmatrix} \mathbf{T}_i^{(h_{x_j})^{-1}} & \mathbf{T}^{(h_{x_j})^{-1}}\mathbf{F}^{(h_{x_j})} \\ \big(\mathbf{T}^{(h_{x_j})^{-1}}\mathbf{F}^{(h_{x_j})}\big)' & \mathbf{T}^{(h_{x_j})^{-1}} \end{bmatrix}^{-1}$$

or, analogously, the conditional correlation matrix

$$(12) \qquad \mathbf{\Omega}_{ij|-ij} = \mathbf{\Delta}^{-1/2} \mathbf{\Sigma}_{iu|-iu}^{h_{x_j}} \mathbf{\Delta}^{-1/2},$$

where $\mathbf{\Delta} = \text{diag}(\mathbf{\Sigma}_{iu|-iu}^{h_{x_j}})$.

The parameters for the priors on $\boldsymbol{\beta}^{(h_{x_i})}$, $\mathbf{T}_i^{(h_{x_i})^{-1}}$ and $\tilde{\mathbf{F}}_{iu}^{(h_{x_i})}$ are set as follows: $\sigma_\beta^2 = 100$, $\varrho_x = 20$, $\mathbf{S}_x = \mathbf{I}$ and $\varsigma = 0.05$. The design matrix $\mathcal{D}^*$ is specified to represent a constant mean in space and we also consider $\mathbf{m}_y(t) = \mathbf{m}_y$ and $\mathbf{m}_x(t) = \mathbf{m}_x$.

MODEL SPECIFICATION: *State equation*. Motivated by the debate on the possible existence of cointegration between RHPI, RPCI and UR, we consider the cointegrated model specification as shown in Section 5.1. The temporal lag of the state equations has been fixed to 2 (i.e., $p^* = 2$), and an increasing number of common factors, that is, $2 \le m, l \le 12$, have been considered for the model specification. Then the maximum possible number of cointegrating relationships is defined as $r_d^* = m - 1$ and $r_f^* = l - 1$. Other modeling details, including prior hyperparameter values, are defined in Section 6 and Appendix B.

Together with the model specification described above, hereafter denoted as M_0, other simpler models representing a simplification of M_0 were also considered for comparison purposes. Specifically, to have an idea of the relative importance of the different specifications used in M_0 (e.g., correlated factor loadings and cointegrated factors), three models with the following assumptions were considered: (i) uncorrelated factor loadings and a simple VAR specification (i.e., without cointegration) for the state equation (M_1), (ii) uncorrelated factor loadings and cointegrated factors (M_2), (iii) correlated factor loadings and a simple VAR specification (i.e., without cointegration) for the factors (M_3). Finally, a fourth model (M_4) which is relatively simple to estimate [see, e.g., Lutkepohl (2005)] but with a completely different structure is also considered:

$$Y(\mathbf{s}_i, t) = \mathbf{c}(\mathbf{s}_i, t)' \boldsymbol{\beta}(\mathbf{s}_i) + u_y(\mathbf{s}_i, t),$$

where $\mathbf{c}(\mathbf{s}_i, t)$ is the vector containing the covariates X_1 and X_2 (including the intercept), $\boldsymbol{\beta}(\mathbf{s}_i)$ is the corresponding vector of (site-specific) regression coefficients and $u_y(\mathbf{s}_i, t)$ is a VAR(2) process where the noise part of the model is assumed to be distributed as a univariate GMRF (i.e., the noise is uncorrelated in time but it is allowed to be spatially correlated). The introduction of a spatial (GMRF) prior on the regression coefficients is also considered in the parametrization.

MODEL ESTIMATION. The identifiability constraints associated with the model to be estimated concern the ordering of the States and the connection between the chosen ordering and the specific form of the factor loading matrices $\mathbf{H}_y$ and $\mathbf{H}_x$. Unfortunately, no fixed rules exist to select the States which must be constrained. In the following, we thus discuss a possible strategy which exploits results from a cluster analysis performed (before estimating the model) on the data matrices $\mathbf{Y}$

and $\mathbf{X}$, respectively, of dimensions $(\tilde{n}_y \times T)$ and $(\tilde{n}_x \times T)$. In this case, considering RHPI, the K-Means classification algorithm is repetitively run for a number of clusters equal to m, with $2 \leq m \leq 12$. The States (one for each cluster) to be constrained are thus chosen as the ones that: (possibly) belong to different BEA regions, show the highest mean values of RHPI and/or are far apart from each other (especially when m is larger than the number of BEA regions). For a given l, such that $2 \leq l \leq 12$, the same procedure is also applied to $\mathbf{X}$ and, whenever possible, the same States selected for the housing prices are chosen. Note that especially in cases in which $l > m$, the choice of the States within the clusters obtained for $\mathbf{X}$ can be made independently of RHPI and based on several criteria such as the membership to different BEA regions and/or highest (smallest) mean values of RPCI (UR). When $m > l$, the same criteria can be adopted to choose the States among the ones already constrained in $\mathbf{H}_y$.

For each fitted model, the MCMC algorithm was run for 250,000 iterations. Posterior inference was based on the last 150,000 draws using every 10th member of the chain to avoid autocorrelation within the sampled values. Several MCMC *diagnostics* could be used to test the convergence of the chains [see, e.g., Geweke (1992), Gilks, Richardson and Spiegelhalter (1996), Spiegelhalter et al. (2002) and Jones et al. (2006)]. In our case, convergence of the chains of the model was monitored visually through trace plots as well as using the R-statistic of Gelman (1996) on four chains starting from very different values.

Competing models were compared using the predictive model choice statistic, PMCC, described in Section 6.5. The PMCC criterion suggests that, for M_0, the optimal choice is found with $m = 7$ and $l = 8$. The same number of components is also confirmed for models M_1–M_3. However, compared with M_3, the best of the three alternative models, the PMCC increases 17%, which denotes much worse model fitting properties.

Notice that for M_0, the following States have been constrained in the factor loading matrix $\mathbf{H}_y$: North Carolina, Montana, California, Massachusetts, Texas, Illinois and Arizona. Instead, considering $\mathbf{H}_x$, we have constrained 5 States for UR: North Carolina, California, Massachusetts, Texas and Illinois, and 3 States for RPCI: Arizona, Montana and Massachusetts.

FACTOR LOADINGS AND COMMON LATENT FACTORS. The MCMC estimates of the endogenous components, $g_i(t), i = 1, \ldots, 7$, appear as nonstationary processes, each representing specific features of the large-scale temporal variability of the RHPI series. The first two latent components represent common trends and are characterized by narrow 95% credibility intervals. Specifically, the pattern of the first component, shown in Figure 2(a), highlights a growth of RHPI since the early nineties up to 2006 followed by a sustained decrease. At the national level, prices increased substantially from 2000 to the peak in 2006 and then have been falling very sharply across the country. An exploratory analysis shows that this component tracks the pattern of the national RHPI, although the latter seems to be a bit more volatile, especially in the period 1984–1994. We also notice that
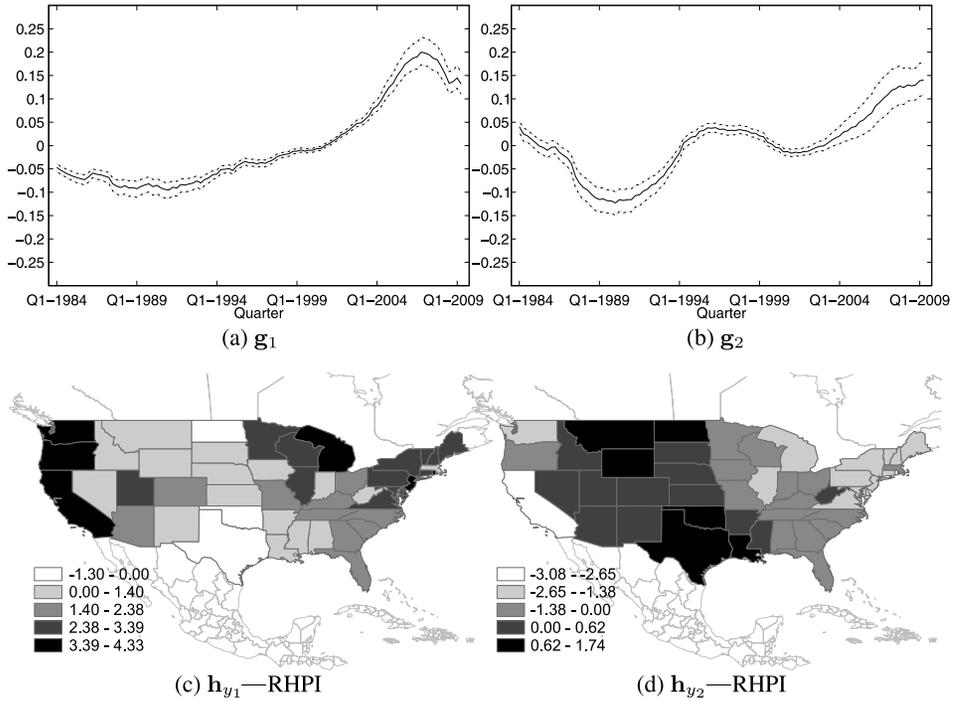
FIG. 2.   *Subplots* (a) *and* (b): *marginal posterior medians for the estimated latent factors* $g_1(t)$ *and* $g_2(t)$ *(continuous line) and their* 95% *credible intervals (dashed line). Subplots* (c) *and* (d): *maps of the posterior medians for the factor loadings* $\mathbf{h}_{y_1}$ *and* $\mathbf{h}_{y_2}$ *related to the real housing price index.*

this component is highly correlated (i.e., the correlation is in general greater than 0.80) with all the State time series with the exception of Connecticut, Texas and Oklahoma, for which the correlation is around 0.50.

The series of the second component, $g_2(t)$, shown in Figure 2(b), is character-ized by a price trough in the mid-1980s and mid-1990s followed by a mild price peak. Then, the late 1990s begin with a dramatic and sustained increase. Examina-tion of the data plotted in Figure 1 shows that this is a typical pattern of the 50% of the States of Plains, Southeast and Rocky Mountain.

The remaining latent variables (not shown here) present some peculiarities for the periods 1984–1990 and 2004–2007 and, compared with the first two factors, are characterized by slightly wider credibility intervals.

Figure 2(c)–(d) show the maps of the estimated first two factor loadings—that is, the first two columns of the measurement matrix $\mathbf{H}_y$. The maps clearly show the presence of clusters of US States. Table 2 also shows the posterior summaries of the between-location conditional correlations estimated [using equation (12)] for each column of $\mathbf{H}_y$. Since the 95% credibility intervals do not overlap zero and all the conditional correlations seem to be statistically significant, the clusters are easily identified by looking at the spatial patterns of the factor loadings.

Posterior summary of the between-location conditional correlations for the columns of the measurement matrix $\mathbf{H}_y$. In brackets we show the 2.5 and 97.5 percentiles used for defining the 95% credible interval limits

	Factor loadings ($\mathbf{H_y}$)						
	1	**2**	**3**	**4**	**5**	**6**	**7**
Median	0.09	0.08	0.08	0.07	0.06	0.08	0.08
95% CI	[0.05, 0.12]	[0.04, 0.12]	[0.02, 0.12]	[0.03, 0.10]	[0.03, 0.12]	[0.02, 0.12]	[0.02, 0.12]

Figure 2(c) shows [using the *natural break method* of ArcMap, ESRI (2009)] the weights of the first factor loading, $\mathbf{h}_{y_1}$. Except for Texas, Oklahoma and North Dakota, these weights are all positive, with the highest loadings observed in the Pacific and Northeast regions, which strongly influence the contiguous regions.

Figure 2(d) also shows an interesting pattern in the loadings. Southwest, Rocky Mountain States, some Plains States and Louisiana have positive loadings, while the other States have negative loadings. The States with highest loadings (Louisiana, New Mexico, Texas, Oklahoma, North Dakota and Wyoming) show a temporal pattern very similar to the second latent variables. On the other hand, the States with lowest values (California, Connecticut, Michigan, New Jersey and Rhode Island) show temporal dynamics which, at least until the end of the nineties, result in the opposite of $g_2(t)$. Many of these States in the last 25 years have been particular beneficiaries of new technologies. These innovations interacting with restrictions on new residential buildings have resulted in real housing prices in these regions deviating from the average across US States over a relatively prolonged period [Holly, Pesaran and Yamagata (2010)]. Also, considering the period 1984–1990, the spatial contrast highlighted in the map of Figure 2(d) clearly confirms that while West–South–Central regions (especially "oil-patch" states such as Texas and Oklahoma) experienced sharp declines, the Northeast and California housing market were booming. Note that this map provides clear evidence of the results described in Table 1 where we have found significant correlations between the States belonging to the East and West regions.

The MCMC estimates of the exogenous components, $f_i(t), i = 1, \ldots, 8$, summarize the dynamics of RPCI and UR variables. The first three of these latent factors, together with their 95% credibility intervals, are shown in Figure 3. These components seem to have a substantial impact on RPCI and UR, although the latter shows more complex dynamics which can be fully understood by examining the behavior of all the estimated factors.

The first factor, $f_1(t)$, shows a cyclical behaviour with a slightly positive trend in the period 1986–2000. The series exhibits a trough in the period 2000–2006 followed by a sustained decrease. The 2000–2006 pattern has roots in the prior turmoil in the financial markets. In fact, the period 2000–2001 is characterized
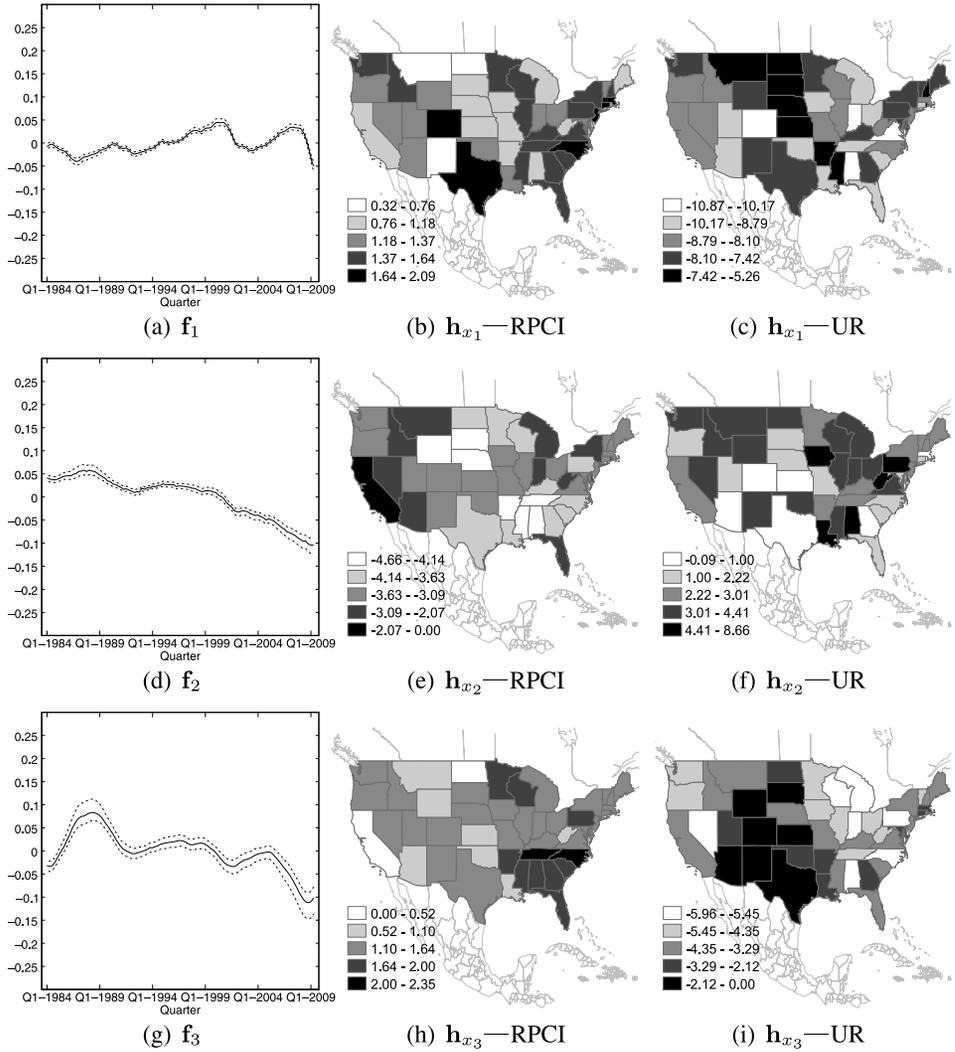
FIG. 3. *Subplots* (a), (d) *and* (g): *marginal posterior medians for the estimated latent factors* $f_1(t)$, $f_2(t)$ *and* $f_3(t)$ (*continuous line*) *and their 95% credible intervals* (*dashed line*). *Subplots* (b), (e) *and* (h): *maps of the posterior medians for the factor loadings* $\mathbf{h}_{x_1}$, $\mathbf{h}_{x_2}$ *and* $\mathbf{h}_{x_3}$ *related to the real per capita personal income variable. Subplots* (c), (f) *and* (i): *maps of the posterior medians for the factor loadings* $\mathbf{h}_{x_1}$, $\mathbf{h}_{x_2}$ *and* $\mathbf{h}_{x_3}$ *related to the unemployment rate variable.*

by a rapid decline of high tech industries, a collapse of the stock market and a slow level of technology investment. The relaxed monetary policy adopted by the Federal Reserve had thus lead to an increase of RPCI and a decrease of UR up to 2007.

The factor loadings related to $f_1(t)$, shown in Figure 3(b) and Figure 3(c), are all positive for RPCI and negative for UR. Figure 3(b) clearly shows groups of States

with common spatial patterns. Specifically, we notice the presence of two clusters: the first involves several States from the Great Lakes, Southeast and New England, while the second is mainly characterized by Oregon and some States of the Mountain region (Arizona, Utah, Nevada and Wyoming). Also, the highest values are related to those States (Colorado, Connecticut, Georgia, Massachusetts, New Jersey, North Carolina and Texas) whose RPCI shows the same cyclical pattern of $f_1(t)$ in the period 1995–2009.

Figure 3(c), related to UR, shows quite a big cluster of States forming a ridge from Montana to Mississippi. For these States the variations of UR are less pronounced with respect to those showing the smallest loadings (e.g., Alabama, Colorado, Indiana and Virginia).

The dynamics of RPCI and UR in the first period of the series is captured by the third latent factor $f_3(t)$ shown in Figure 3(g). The figure shows that the early nineties are characterized by a trough of UR and a hill for the RPCI.

Figure 3(h) shows a huge cluster with values of the loadings in the range 1.10–1.64; the highest values are observed in the Southeast region for which the oscillations of RPCI are a bit more pronounced than other States.

Figure 3(i) shows that the States for which the trough of UR is more pronounced are characterized by lowest values of the loadings. Notice that this figure also shows a reasonable correspondence with Figure 2(d).

The second factor, $f_2(t)$, shows a decreasing trend associated with negative values of $\mathbf{h}_{x_2}$—RPCI—and (mainly) positive values of $\mathbf{h}_{x_2}$—UR. The maps of the factor loading clearly provide information on those States which have experienced a positive trend for RPCI (e.g., Alabama, Arkansas, Mississippi, Nebraska, South Dakota, Tennessee and Wyoming) as well as a downward trend for UR (see, e.g., Alabama, Iowa, Louisiana, Pennsylvania and West Virginia).

The spatial structure of the factor loadings is also confirmed by the the posterior summaries of their within- and between-location conditional correlations and cross-correlations (see Table 3). The 95% credibility intervals suggest that most parts of these correlations can be considered as nonzero. Also, the conditional spatial dependence of each factor loading is positive, while both the between- and the within-location conditional cross-correlations are negative.

MODEL ESTIMATION: COINTEGRATION. As noted in the introduction, there has been quite a long debate in the literature about whether there is cointegration between real housing prices and fundamentals. The idea is that in the absence of cointegration there are no fundamentals driving real housing prices and the absence of an equilibrium relationship would essentially increase the presence of bubbles [Case and Shiller (2003), Holly, Pesaran and Yamagata (2010)]. Here, we test the existence of this cointegrating relationship in a latent space, avoiding to take account of the effect of the cross-sectional dependence [see Holly, Pesaran and Yamagata (2010) for a discussion on this point]. In terms of cointegrated ranks, following Jochmann et al. (2013), our posteriors for r_f, r_d, r_c, r_{c_1} and r_{c_2} are obtained by considering the draws of their respective matrices (i.e., $\mathbf{\Pi}_f$, $\mathbf{\Pi}_{gd}$,

TABLE 3

Posterior summary of the within- and between-location conditional correlations and cross-correlations for the first three factor loadings columns related to the unemployment rate and real per capita personal income variables. In brackets we show the 2.5 and 97.5 percentiles used for defining the 95% credible interval limits

	Conditional correlation				
	Within-location RPCI vs UR	Between-location RPCI	Between-location RPCI vs UR	Between-location UR vs RPCI	Between-location UR
$\mathbf{h}_{x_1}$	−0.22 [−0.44, −0.09]	0.06 [0.01, 0.09]	−0.04 [−0.07, −0.02]	−0.03 [−0.07, −0.01]	0.05 [0.03, 0.08]
$\mathbf{h}_{x_2}$	0.02 [−0.29, 0.12]	0.05 [0.01, 0.10]	−0.00 [−0.06, 0.05]	−0.01 [−0.06, 0.06]	0.07 [0.02, 0.09]
$\mathbf{h}_{x_3}$	−0.27 [−0.38, −0.02]	0.08 [0.02, 0.12]	−0.02 [−0.07, 0.03]	−0.04 [−0.07, −0.01]	0.07 [0.03, 0.10]

$\mathbf{A}\mathbf{B}_2' + \mathbf{A}_2\mathbf{B}_f'$, $\mathbf{A}\mathbf{B}_2'$ and $\mathbf{A}_2\mathbf{B}_f'$; see Appendix A) and taking the number of singular values greater than 0.05.

These are shown in Table 4 where we note that there is a strong support for an exogenous cointegrated rank of either 4 or 5; for r_d there is a hint of a rank equal to 5, but small probabilities are also observed for 4 and 6. Finally, since there is evidence that $r_c < r_{c_1} + r_{c_2}$, we may conclude that a cointegration structure is confirmed between the endogenous and exogenous processes. Such a result thus supports the idea about the existence of a convergence to a stable equilibrium relationship and, hence, about the absence of a US housing price bubble for the period considered in the study.

To provide further evidence that our approach is yielding sensible results, the use of Bayes factors using a non-SSVS prior [Sugita (2009), Kass and Raftery (1995)] confirms that, conditionally on $m = 7$ and $l = 8$, results for r_f and r_d are similar to those presented here.

TABLE 4

Posterior of cointegration ranks r_f, r_d, r_c, r_{c_1} and r_{c_2}

	Estimated probabilities for effective ranks					
	1	2	3	4	5	6
r_f	0.00	0.00	0.01	0.34	0.61	0.04
r_d	0.00	0.00	0.00	0.18	0.70	0.12
r_c	0.00	0.00	0.00	0.02	0.48	0.50
r_{c_1}	0.00	0.00	0.10	0.63	0.27	0.00
r_{c_2}	0.00	0.00	0.03	0.45	0.50	0.02

UNCONDITIONAL AND CONDITIONAL FORECASTS. To test the predictive performance of the SD-SEM model, the last 10 quarters have been excluded from the estimation procedure and used only for forecast purposes. Hence, we consider the forecast for a horizon of $k = 10$ periods corresponding to the quarters Q3-2009–Q4-2011. Also, predictions of RHPI are obtained by following two settings:

 (i) *unconditional predictions*: we only use past information; hence, X is not available for the forecast period;

(ii) *conditional predictions*: the exogenous variables X_1 and X_2 are assumed known in the period in which temporal forecasts of RHPI are required.

For each State, both unconditional and conditional forecasts (together with 95% credible intervals) of the housing price index are shown in Figure 4. In general, compared with true values, good prediction results can be achieved and, as expected, the conditional (on the known values of X) approach exhibits more encouraging out-of-sample properties of the model, with data points being more accurately predicted.

To provide some measures of goodness of prediction for the estimated models, Table 5 gives details on the root mean squared prediction error, RMSE = $\sqrt{\text{mean}\{(\tilde{Y}(\mathbf{s}, t) - E[\tilde{Y}(\mathbf{s}, t)_{\text{rep}}])^2\}}$, the mean absolute error deviation, MAE = $\text{mean}\{|\tilde{Y}(\mathbf{s}, t) - E[\tilde{Y}(\mathbf{s}, t)_{\text{rep}}]|\}$ [where $\tilde{Y}$ is the variable at the original scale and the mean is taken over the $(N \times k)$ observations], the coverage probabilities (CP) and the average width (AIW) of the prediction intervals. We note that in the conditional case model M_0 shows much smaller values for RMSE, MAE and AIW; on the other hand, the coverage probabilities of the 95% intervals are larger than the nominal rate. Models M_1, M_2 and M_3 provide very similar results and provide some hints on the role played by the spatially autocorrelated factor loadings and cointegrated factors. In general, model M_0 works better than M_1, M_2 and M_3 for which the average width of the prediction intervals are wider. We note that in introducing the spatial correlation the AIW reduces substantially. The same effect, albeit with different intensity, can be observed assuming cointegrated factors and this can be detected by contrasting models M_0–M_3 and M_1–M_2.

By making the series stationary through a first difference transformation, the best result of model M_4 is characterized by an RMSE of 13.575 and a MAE of 11.372. This result is obtained by using a GMRF prior on the regression coefficients. We also note that for this model the regressors, X_1 and X_2, are assumed as known for the forecast period. Producing unconditional predictions under model M_4, in fact, is not straightforward since it requires further adjustments for predicting the process X.

MULTIPLIER ANALYSIS. We conclude the analysis by providing some results from multiplier analysis [Lutkepohl (2005)] which is helpful to describe how the housing price index reacts over time to exogenous impulses. In this case, we can
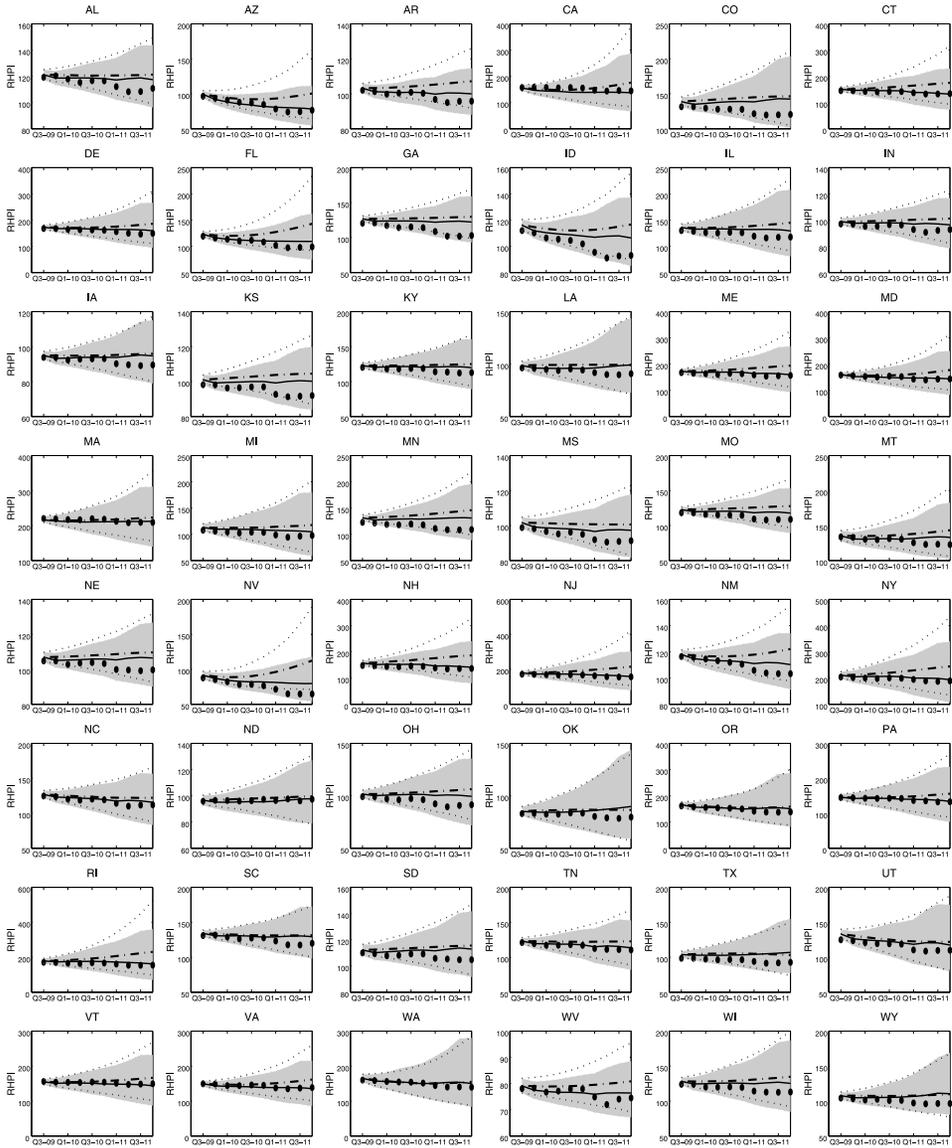
FIG. 4. *Unconditional forecasts* (*dashed line*), *conditional forecasts* (*continuous line*) *and true data* (•) *at the* 48 *United States*; *the* 95% *credible interval limits for the unconditional forecasts are represented by dotted lines. The* 95% *credible interval limits for the conditional forecasts are represented by the shaded area. Each subplot also shows the initials of the State.*

check if past values on either RPCI or UR, observed on a specific State, contain useful information to predict the variation of RHPI, in addition to the information on its past values. It can be shown (see Appendix C) that the dynamic multipliers, $\mathbf{\Gamma}_k$, which reflect the marginal impacts of changes in the predictors X_1 and X_2, are

TABLE 5

Root mean squared prediction errors (RMSE), mean absolute deviations (MAE), coverage probabilities (CP) and average width (AIW) of the prediction intervals, for unconditional and conditional forecasts of RHPI. The statistics are computed for the estimated models M_0, M_1, M_2, M_3 and M_4

Model	Type of prediction	RMSE	MAE	CP 95% interval	AIW 95% interval
M_0	Unconditional	16.081	11.704	0.958	59.762
	Conditional	7.223	5.558	0.989	54.723
M_1	Unconditional	17.294	12.950	1.000	140.052
	Conditional	9.497	6.614	1.000	138.140
M_2	Unconditional	17.496	12.942	0.989	112.814
	Conditional	9.904	7.414	0.998	112.086
M_3	Unconditional	17.150	12.759	0.969	77.042
	Conditional	9.331	6.695	0.985	76.445
M_4	Unconditional	–	–	–	–
	Conditional	13.575	11.372	0.920	53.180

defined as

$$\boldsymbol{\Gamma}_k = \mathbf{H}_y \mathfrak{J} \mathcal{Q}^k \mathcal{B} \mathbf{H}_x^\dagger, \qquad k = 0, 1, \ldots,$$

where, at the kth period (quarter), the $\gamma_{ij,k}$ element of the $(N \times \tilde{n}_x)$ matrix $\boldsymbol{\Gamma}_k$ represents the response of the housing price in the ith State to a given shock in the predictor X_l, $l = 1, 2$, in State j, provided the effect is not contaminated by other shocks to the system. The matrices $\mathfrak{J}$, $\mathcal{Q}$ and $\mathcal{B}$, which contribute to determine the multipliers, are defined in Appendix C.

The impulse responses of RHPI to a 1% shock in the exogenous variables, RPCI and UR, in each State, show some interesting features. However, since many possible interactions among States and variables can be envisaged, in the following we provide a summary of the results as well as a visual impression of some of the dynamic interrelationships existing in the system. Note that following Sims and Zha (1999) and Primiceri (2005), the credibility intervals of the impulse response coefficients are discussed at the 16th and 84th percentiles which, under normality, correspond to the bounds of a one-standard-deviation.

One interesting feature is that a shock in RPCI in the States belonging to New England (with the exception of Connecticut and New Jersey) does not seem to produce evident effects on RHPI. The same holds for a RPCI shock in Mideast States whose effects seem to disappear after one quarter. It thus seems that past values of RPCI, in these regions, do not help in forecasting RHPI throughout the US. At the same time, apart from New Hampshire and Maryland, the prices in New England and the Mideast do not seem to react to a RPCI shock in any other region. The housing prices in Michigan, Ohio and Illinois, belonging to the Great Lakes,

also seem to behave similarly. Note that this similarity in behavior was also found by Apergis and Payne (2012) in a study on housing price convergence.

On the other hand, there is stronger evidence of the relationships between UR shock effects in the States of New England and the Mideast and RHPI responses in several States, mainly belonging to the Southeast, Plains and Southwest regions. Also, RHPI forecasts in New England and Mideast regions can be improved by exploiting UR information on other States. In any case, considering the infra-regional responses (i.e., RHPI responses of New England and Mideast States to a UR shock produced in any State belonging to the same region), we note that UR effects on the variation of RHPI disappear after one period.

Regarding the remaining BEA regions, a 1% shock to either RPCI or UR seems to highlight effects on the housing prices involving quite a large network of States, particularly in the second quarter. Analyzing the impulse responses for longer periods, we note that the network of relevant relationships between the States becomes sparser. However, the most persistent effects on RHPI, which also involve a large numbers of States belonging to the Southeast, Plains, Rocky Mountain, Southwest and Far West regions, are associated to RPCI shocks in Nevada, Arizona, Georgia, Alabama and Mississippi, and to UR shocks in Illinois, South Carolina, Florida, Alabama, Iowa, South Dakota and Nebraska.

Moreover, the States whose RHPI responses are more persistent to RPCI shocks in any other State of the aforementioned regions are Florida and Nevada, while the States whose responses are more persistent to UR shocks are New Mexico, Arizona, Arkansas and Mississippi.

If we consider the sign of the impulse response coefficients, we note that, in general, a positive shock to RPCI is associated to a positive effect on RHPI. Some exceptions are observed in the first period where we can find negative coefficients. On the other hand, the scenario appears to be different for the UR case, in which we note both positive and negative effects on RHPI even for longer periods. Although we may expect that unemployment has an adverse effect on real estate prices, previous studies have nevertheless found unemployment to be positively related to housing prices. For a discussion on this point we refer the reader, for example, to Vermeulen and Van Ommeren (2009), Clayton, Miller and Peng (2010) and Moench and Ng (2011).

Finally, to provide a flavor of the type of relationship, Figure 5 shows posterior mean housing price responses (solid line) in Nevada, Oregon, Arizona, New Mexico, Utah, Idaho and California to a 1% shock to RPCI in Nevada. Figure 6, instead, shows the responses in Florida, Tennessee, Alabama, Mississippi, Arkansas, West Virginia, North Carolina and Georgia to a 1% shock to UR in Florida. The shaded regions indicate the credibility intervals corresponding to 68 and 90 percent. Overall, the plots suggest that State-level responses follow a similar pattern (consistently with the ripple effect) and, in most cases, the effects tend to decay over two years, especially for UR shocks.
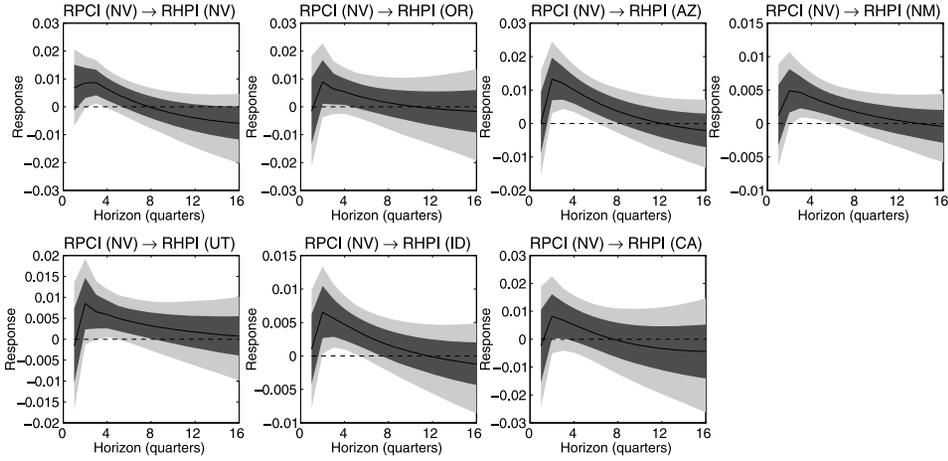
FIG. 5. *Posterior mean impulse responses (solid line) of RHPI to a RPCI shock in Nevada. The credibility intervals at* 68% *and* 90% *are represented by shaded areas. The responses are observed in Nevada (NV), Oregon (OR), Arizona (AZ), New Mexico (NM), Utah (UT), Idaho (ID) and California (CA).*

9. Discussion. In this paper we have discussed the modeling of spatio-temporal multivariate processes observed on a lattice by means of a Bayesian spatial dynamic structural equation model. We have used ideas from factor analysis to frame and exploit both the spatial and the temporal structure of the observed processes.
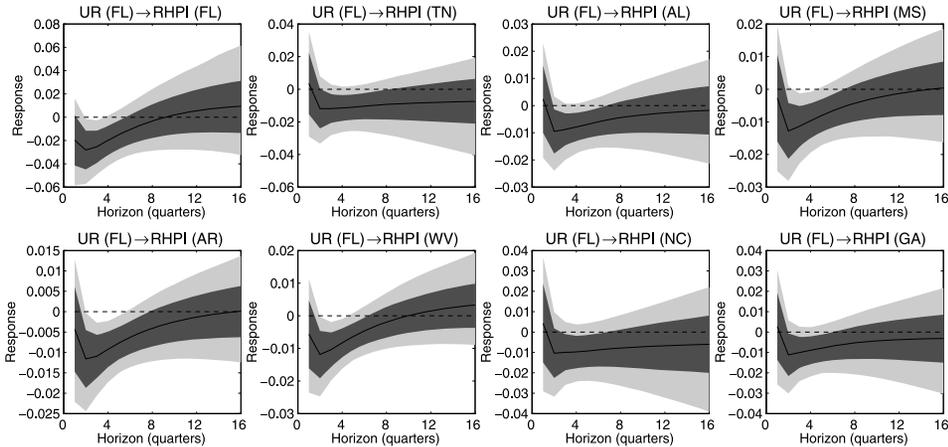
FIG. 6. *Posterior mean impulse responses (solid line) of RHPI to a UR shock in Florida. The credibility intervals at* 68% *and* 90% *are represented by shaded areas. The responses are observed in Florida (FL), Tennessee (TN), Alabama (AL), Mississippi (MS), Arkansas (AR), West Virginia (WV), North Carolina (NC) and Georgia (GA).*

It can be shown that the SD-SEM encompasses a large class of spatial-temporal models that are commonly used and, more importantly, differs from them in two major aspects: (i) it avoids the curse of dimensionality commonly present in large spatio-temporal data and (ii) it facilitates the formation of spatial clusters which further avoids dimensionality issues.

The model has been implemented in a Bayesian setup using MCMC sampling. The MCMC chains of the parameters were monitored to detect possible problems in convergence although no such problems were found in the implementation.

The model was applied to study the impact that the real per capita personal income and the unemployment rate may have on the real housing prices in the USA using State level data. Forecasting the future economic conditions and understanding the relations between the observed variables have been two important aspects covered by our model. The spatial variation is brought into the model through the columns of the factor loading matrix and the estimated conditional correlations and cross-correlations gave significant evidence of spatial dependence associated with contiguity. The spatial patterns of the factor loadings revealed several clusters of interest showing common dynamics.

The time series dynamics have been captured by common dynamic factors. An error correction model specification, with a cointegrating relationship between the common latent factors, was found useful once we took proper account of both heterogeneity and cross-sectional dependence. Overall, results support the hypothesis that real housing prices have been rising in line with fundamentals (real incomes and unemployment rates), and there seems no evidence of housing price bubbles at the national level.

Results from multiplier analysis were also helpful to describe how the housing price index reacts over time to exogenous impulses. We have found that, consistently with the ripple effect, the RHPI responses show a similar pattern for neighboring States. The responses seem to be more persistent to UR shocks, while the effects of a RPCI shock decay more rapidly such that the system appears to approach faster to the initial equilibrium conditions.

A further important advantage of the model formulation is that it enables consideration of cases in which the temporal series of X are longer than those of Y. As noticed in Section 7, this was particularly useful to improve the temporal predictions by conditioning on known values of the predictor providing a set of plausible scenarios for RHPI.

Of course, we acknowledge that other possibilities could be considered for modeling the spatial structure and an example is provided by Wang and Wall (2003). An alternative scheme could also lead to the specification of common factors with a spatio-temporal structure. In this case, one may follow the methodology proposed in Debarsy, Ertur and LeSage (2012) to quantify dynamic responses over time and space as well as space–time diffusion impacts.

Finally, in this paper we have focused exclusively on normally distributed data. However, nonlinear and non-Gaussian spatio-temporal models have been extensively used in various areas of science, from epidemiology to meteorology and

environmental sciences, among others. In this case, assuming the measurements belong to the exponential family of distributions, a *generalized* spatial dynamic structural equation model represents a natural extension of the SD-SEM discussed here. This extension will be a topic for future work.

APPENDIX A: COINTEGRATED LATENT FACTORS AND THEIR VECTOR ERROR CORRECTION REPRESENTATION

Let $\tilde{\boldsymbol{\Phi}}(z)$ denote the characteristic polynomial associated with the vector ECM shown in (11) and let c be the number of unit roots of $\mathrm{Det}[\boldsymbol{\Phi}(z)]$. Let also that $\mathrm{rank}(\tilde{\mathbf{A}}) = r$, with $r = m + l - c$. Then, we assume that the latent exogenous variables, $\mathbf{f}(t)$, are cointegrated with cointegrating rank r_f so that $r > r_f$ and $r_f < l$.

Let $\mathbf{Q}(\sum_{i=1}^{p} \boldsymbol{\Phi}_i)\mathbf{P} = \mathbf{J}$ be the Jordan canonical form of $\sum_{i=1}^{p} \boldsymbol{\Phi}_i$, where $\mathbf{Q} = \mathbf{P}^{-1}$ an $((m+l) \times (m+l))$ matrix, $\mathbf{J} = \mathrm{diag}(\mathbf{I}_{m-r_d}, \boldsymbol{\Lambda}_{r_d}, \mathbf{I}_{l-r_f}, \boldsymbol{\Lambda}_{r_f})$ and $r_d \equiv r - r_f$ [Ahn and Reinsel (1990) and Cho (2010)]. Because of the exogeneity of $\mathbf{f}(t)$, the matrices $\tilde{\mathbf{A}}$ and $\tilde{\boldsymbol{\Phi}}_i$ are upper block triangular matrices, that is,

$$\tilde{\mathbf{A}} = \begin{bmatrix} \tilde{\mathbf{A}}_1 & \tilde{\mathbf{A}}_{12} \\ \mathbf{0} & \tilde{\mathbf{A}}_2 \end{bmatrix} \quad \text{and} \quad \tilde{\boldsymbol{\Phi}}_i = \begin{bmatrix} \tilde{\boldsymbol{\Phi}}_{1i} & \tilde{\boldsymbol{\Phi}}_{12i} \\ \mathbf{0} & \tilde{\boldsymbol{\Phi}}_{2i} \end{bmatrix}.$$

Then, consider the following matrix partition:

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_1 & \mathbf{P}_{12} \\ \mathbf{0} & \mathbf{P}_2 \end{bmatrix}, \qquad \mathbf{Q} = \mathbf{P}^{-1} = \begin{bmatrix} \mathbf{P}_1^{-1} & -\mathbf{P}_1^{-1}\mathbf{P}_{12}\mathbf{P}_2^{-1} \\ \mathbf{0} & \mathbf{P}_2^{-1} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_{12} \\ \mathbf{0} & \mathbf{Q}_2 \end{bmatrix},$$

with $\mathbf{Q}_1' = [\mathbf{Q}_1^{(1)} \ \mathbf{Q}_1^{(2)}]$, $\mathbf{P}_1 = [\mathbf{P}_1^{(1)} \ \mathbf{P}_1^{(2)}]$, $\mathbf{Q}_{12}' = [\mathbf{Q}_{12}^{(1)} \ \mathbf{Q}_{12}^{(2)}]$, $\mathbf{P}_{12} = [\mathbf{P}_{12}^{(1)} \ \mathbf{P}_{12}^{(2)}]$, $\mathbf{Q}_2' = [\mathbf{Q}_2^{(1)} \ \mathbf{Q}_2^{(2)}]$ and $\mathbf{P}_2 = [\mathbf{P}_2^{(1)} \ \mathbf{P}_2^{(2)}]$.

Note that $\mathbf{Q}_1^{(1)}, \mathbf{P}_1^{(1)}$ are $(m \times (m - r_d))$, $\mathbf{Q}_1^{(2)}, \mathbf{P}_1^{(2)}$ are $(m \times r_d)$, $\mathbf{Q}_2^{(1)}, \mathbf{P}_2^{(1)}$ are $(l \times (l - r_f))$, $\mathbf{Q}_2^{(2)}, \mathbf{P}_2^{(2)}$ are $(l \times r_f)$, $\mathbf{Q}_{12}^{(1)}$ is $(l \times (m - r_d))$, $\mathbf{P}_{12}^{(1)}$ is $(m \times (l - r_f))$, $\mathbf{Q}_{12}^{(2)}$ is $(l \times r_d)$ and $\mathbf{P}_{12}^{(2)}$ is $(m \times r_f)$. Then, we may write

$$\tilde{\mathbf{A}} = -\mathbf{P}(\mathbf{I} - \mathbf{J})\mathbf{Q} = -\begin{bmatrix} \mathbf{P}_1^{(2)} & \mathbf{P}_{12}^{(2)} \\ \mathbf{0} & \mathbf{P}_2^{(2)} \end{bmatrix}\begin{bmatrix} \mathbf{I} - \boldsymbol{\Lambda}_{r_d} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} - \boldsymbol{\Lambda}_{r_f} \end{bmatrix}\begin{bmatrix} \mathbf{Q}_1^{(2)'} & \mathbf{Q}_{12}^{(2)'} \\ \mathbf{0} & \mathbf{Q}_2^{(2)'} \end{bmatrix}$$

$$= -\begin{bmatrix} \mathbf{P}_1^{(2)}(\mathbf{I} - \boldsymbol{\Lambda}_{r_d})\mathbf{Q}_1^{(2)'} & \mathbf{P}_1^{(2)}(\mathbf{I} - \boldsymbol{\Lambda}_{r_d})\mathbf{Q}_1^{(2)'}\mathbf{P}_{12}\mathbf{Q}_2 + \mathbf{P}_{12}^{(2)}(\mathbf{I} - \boldsymbol{\Lambda}_{r_f})\mathbf{Q}_2^{(2)'} \\ \mathbf{0} & \mathbf{P}_2^{(2)}(\mathbf{I} - \boldsymbol{\Lambda}_{r_f})\mathbf{Q}_2^{(2)'} \end{bmatrix},$$

and equation (11) can thus be rewritten as

$$(13) \qquad \Delta\mathbf{g}(t) = \mathbf{A}\mathbf{B}'\mathbf{d}(t-1) + \mathbf{A}_2\mathbf{B}_f'\mathbf{f}(t-1) + \sum_{i=1}^{p-1} \mathbf{K}_i \Delta\mathbf{d}(t-i) + \boldsymbol{\xi}(t),$$

$$(14) \qquad \Delta\mathbf{f}(t) = \mathbf{A}_f\mathbf{B}_f'\mathbf{f}(t-1) + \sum_{i=1}^{p-1} \tilde{\boldsymbol{\Phi}}_{2i} \Delta\mathbf{f}(t-j) + \boldsymbol{\eta}(t),$$

where $\mathbf{A} = -\mathbf{P}_1^{(2)}(\mathbf{I} - \mathbf{\Lambda}_{r_d})$, $\mathbf{B} = [\mathbf{I} \ -\mathbf{P}_{12}\mathbf{Q}_2]'\mathbf{Q}_1^{(2)}$, $\mathbf{A}_f = -\mathbf{P}_2^{(2)}(\mathbf{I} - \mathbf{\Lambda}_{r_f})$, $\mathbf{A}_2 = -\mathbf{P}_{12}^{(2)}(\mathbf{I} - \mathbf{\Lambda}_{r_f})$, $\mathbf{B}_f = \mathbf{Q}_2^{(2)}$ and $\mathbf{K}_i = [\tilde{\mathbf{\Phi}}_{1i} \ \tilde{\mathbf{\Phi}}_{12i}]$. Note that if $\mathbf{P}_{12}$ and $\mathbf{P}_{12}^{(2)}$ are $\mathbf{0}$, then a separated cointegrated structure exists for $\mathbf{g}(t)$ and $\mathbf{f}(t)$.

Let $\mathbf{B} = [\mathbf{B}_1' \ \mathbf{B}_2']'$ where $\mathbf{B}_1 = \mathbf{Q}_1^{(2)}$ and $\mathbf{B}_2 = -\mathbf{Q}_2'\mathbf{P}_{12}'\mathbf{Q}_1^{(2)}$, then $\tilde{\mathbf{A}}$ can be rewritten as

$$\tilde{\mathbf{A}} = -\begin{bmatrix} \mathbf{AB}_1' & \mathbf{AB}_2' + \mathbf{A}_2\mathbf{B}_f' \\ \mathbf{0} & \mathbf{A}_f\mathbf{B}_f' \end{bmatrix}.$$

Also, let $r_f = \text{rank}(\mathbf{A}_f\mathbf{B}_f')$, $r_d = \text{rank}(\mathbf{AB}')$, $r_c = \text{rank}(\mathbf{AB}_2' + \mathbf{A}_2\mathbf{B}_f')$, $r_{c_1} = \text{rank}(\mathbf{AB}_2')$ and $r_{c_2} = \text{rank}(\mathbf{A}_2\mathbf{B}_f')$. Then, it follows that if $\text{rank}(\mathbf{AB}_2' + \mathbf{A}_2\mathbf{B}_f') = 0$, no cointegration structure exists between the endogenous and exogenous processes $\mathbf{g}(t)$ and $\mathbf{f}(t)$.

APPENDIX B: THE SSVS PRIOR FOR THE VECTOR ECM

Since $\mathbf{\Pi}_{gd} = \mathbf{AB}'$, $\mathbf{\Pi}_{gf} = \mathbf{A}_2\mathbf{B}_f'$ and $\mathbf{\Pi}_f = \mathbf{A}_f\mathbf{B}_f'$ are not unique, in this paper we follow the approach proposed by Jochmann et al. (2013) and Koop, León-González and Strachan (2010) to elicit the SSVS priors on the cointegration space. A summary of this approach is provided below.

Specifically, a nonidentified $r_d^* \times r_d^*$ symmetric positive definite matrix $\mathbf{E}$ is introduced with the property, $\mathbf{\Pi}_{gd} = \mathbf{A}\mathbf{E}\mathbf{E}^{-1}\mathbf{B}' \equiv \bar{\mathbf{A}}\bar{\mathbf{B}}'$, where $\bar{\mathbf{A}} = \mathbf{A}\mathbf{E}$ and $\bar{\mathbf{B}} = \mathbf{B}\mathbf{E}^{-1}$. The introduction of the nonidentified matrix $\mathbf{E}$ facilitates posterior computation because the posterior conditional distributions of $\bar{\mathbf{A}}$ and $\bar{\mathbf{B}}$ in the MCMC algorithm are Gaussian [Koop, León-González and Strachan (2010)]. The same holds analogously for $\mathbf{\Pi}_{gf}$ and $\mathbf{\Pi}_f$.

Let $\bar{\mathbf{a}} = \text{vec}(\bar{\mathbf{A}}')$ and $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_{\tilde{m}})$ a parameter vector, where $\tilde{m} = mr_d^*$. Then, we assume that $\bar{\mathbf{a}}|\boldsymbol{\rho} \sim N(\mathbf{0}, \mathbf{V}_0)$, where $\mathbf{V}_0 = \text{diag}(v_1^2, \ldots, v_{\tilde{m}}^2)$, $v_i^2 = (1 - \rho_i)v_{0i}^2 + \rho_i v_{1i}^2$ and ρ_i, the ith element of $\boldsymbol{\rho}$, has a Bernoulli distribution with parameter p_a, that is, $\rho_i \sim \text{Be}(p_a)$. In this paper, we set $p_a = 0.5$, $v_{0i}^2 = 0.1\hat{\sigma}^2(\bar{a}_i)$, $v_{1i}^2 = 10\hat{\sigma}^2(\bar{a}_i)$, where $\hat{\sigma}^2(\bar{a}_i)$ is an estimate of the variance of the ith element of $\bar{\mathbf{a}}$ obtained from a preliminary MCMC run with a noninformative prior.

With appropriate notation, the same assumptions hold for $\bar{\mathbf{a}}_f = \text{vec}(\bar{\mathbf{A}}_f)$, with $\bar{\mathbf{A}}_f = \mathbf{A}_f\mathbf{E}_f$, and $\bar{\mathbf{a}}_2 = \text{vec}(\bar{\mathbf{A}}_2)$, with $\bar{\mathbf{A}}_2 = \mathbf{A}_2\mathbf{E}_f$.

The prior for the cointegrated space is defined through $\bar{\mathbf{b}} \sim N(\mathbf{0}, \mathbf{I})$ and $\bar{\mathbf{b}}_f \sim N(\mathbf{0}, \mathbf{I})$, where $\bar{\mathbf{b}} = \text{vec}(\bar{\mathbf{B}})$, $\bar{\mathbf{b}}_f = \text{vec}(\bar{\mathbf{B}}_f)$ and $\bar{\mathbf{B}}_f = \mathbf{B}\mathbf{E}_f^{-1}$. The SSVS prior for $\mathbf{k} = \text{vec}([\mathbf{K}_1, \ldots, \mathbf{K}_{p^*-1}]')$ is given by $\mathbf{k}|\boldsymbol{\delta} \sim N(\mathbf{0}, \mathbf{D})$, where $\mathbf{D} = \text{diag}(\tau_1^2, \ldots, \tau_{(m+l)(p^*-1)}^2)$, $\tau_i^2 = (1 - \delta_i)\tau_{0i}^2 + \delta_i\tau_{1i}^2$ and $\boldsymbol{\delta}$ is an unknown vector with typical element $\delta_i \sim \text{Be}(p_\tau)$. Here, we set $p_\tau = 0.5$, $\tau_{0i}^2 = 0.1\hat{\sigma}^2(k_i)$, $\tau_{1i}^2 = 10\hat{\sigma}^2(k_i)$, and $\hat{\sigma}^2(k_i)$ is an estimate of the variance of the ith element of $\mathbf{k}$ obtained from a preliminary MCMC run using a noninformative prior. Analogously, we define $\boldsymbol{\phi} = \text{vec}([\tilde{\mathbf{\Phi}}_{21}, \ldots, \tilde{\mathbf{\Phi}}_{2p^*-1}]')$ and assume that $\boldsymbol{\phi}|\boldsymbol{\delta}_\phi \sim N(\mathbf{0}, \mathbf{D}_\phi)$,

where $\mathbf{D}_\phi = \mathrm{diag}(\kappa_{\phi 1}^2, \ldots, \kappa_{\phi m^2(p^*-1)}^2)$, $\kappa_{\phi i}^2 = (1 - \delta_{\phi i})\kappa_{\phi 0i}^2 + \delta_{\phi i}\kappa_{\phi 1i}^2$ and $\boldsymbol{\delta}_\phi$ is an unknown vector with element $\delta_{\phi i} \sim \mathrm{Be}(p_\phi)$. Here we set $p_\phi = 0.5$, $\kappa_{\phi 0i}^2 = 0.1\hat{\sigma}^2(\bar{\phi}_i)$, $\kappa_{\phi 1i}^2 = 10\hat{\sigma}^2(\bar{\phi}_i)$, and $\hat{\sigma}^2(\bar{\phi}_i)$ is an estimate of the variance of the ith element of $\bar{\boldsymbol{\phi}}$ obtained from a preliminary MCMC run using a noninformative prior.

APPENDIX C: MULTIPLIER ANALYSIS

If the model contains integrated variables and the generation mechanism is started at time $t = 0$, it readily follows that [Lütkepohl (2005), page 402–407]

$$(15) \qquad \mathbf{g}(t) = \mathfrak{J}\mathcal{Q}^t\mathbf{g}(0) + \sum_{i=0}^{t-1}\mathfrak{J}\mathcal{Q}^i\mathcal{B}\mathbf{f}(t-i) + \sum_{i=0}^{t-1}\mathfrak{J}\mathcal{Q}^i\mathfrak{J}'\boldsymbol{\xi}(t-i),$$

where $\mathfrak{J}$, $\mathcal{B}$ and $\mathcal{Q}$ are $(m \times (mp+ls))$, $((mp+ls) \times l)$ and $((mp+ls) \times (mp+ls))$ matrices such that

$$\mathfrak{J} = [\mathbf{I} \quad \mathbf{0} \quad \cdots \quad \mathbf{0}],$$

$$\mathcal{B} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbf{I}_l \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}, \qquad \mathcal{Q} = \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_2 & \cdots & \mathbf{C}_p & \mathbf{D}_1 & \cdots & \mathbf{D}_2 & \mathbf{D}_s \\ \mathbf{I}_m & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \cdots & \mathbf{I}_m & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{I}_l & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{I}_l & \mathbf{0} \end{bmatrix}.$$

Then, assuming without loss of generality $\mathbf{m}_x(t) = \mathbf{0}$ and $\mathbf{m}_y(t) = \mathbf{0}$, it follows from the measurement equation (1) that by denoting with $\mathbf{H}_x^\dagger$ the pseudo-inverse of $\mathbf{H}_x$, that is, $\mathbf{H}_x^\dagger = (\mathbf{H}_x'\mathbf{H}_x)^{-1}\mathbf{H}_x'$, for $m < \tilde{n}_x$ and $\mathbf{H}_x'\mathbf{H}_x$ invertible, the least-square estimator of $\mathbf{f}(t)$ is $\hat{\mathbf{f}}(t) = \mathbf{H}_x^\dagger\mathbf{X}(t)$.

Hence, from equations (2) and (15), it follows that the marginal impact of changes of the predictor $\mathbf{X}(t)$ on the dependent variable $\mathbf{Y}(t)$ can be investigated through the coefficient matrices

$$\boldsymbol{\Gamma}_k = \mathbf{H}_y\mathfrak{J}\mathcal{Q}^k\mathcal{B}\mathbf{H}_x^\dagger, \qquad k = 0, 1, \ldots$$

REFERENCES

AHN, S. K. and REINSEL, G. C. (1990). Estimation for partially nonstationary multivariate autoregressive models. *J. Amer. Statist. Assoc.* **85** 813–823. MR1138362

ANSELIN, L. (1988). *Spatial Econometrics*: *Models and Applications*. Kluwer Academic, Dordrecht, The Netherlands.

APERGIS, N. and PAYNE, J. E. (2012). Convergence in U.S. housing prices by state: Evidence from the club convergence and clustering procedure. *Letters in Spatial and Resource Sciences* **5** 103–111.

BANERJEE, S., CARLIN, B. P. and GELFAND, A. E. (2004). *Hierarchical Modeling and Analysis for Spatial Data*. Chapman & Hall/CRC, Boca Raton, FL.

BOX, G. E. P., JENKINS, G. M. and REINSEL, G. C. (1994). *Time Series Analysis*: *Forecasting and Control*, 3rd ed. Prentice Hall Inc., Englewood Cliffs, NJ. MR1312604

BROWN, P. J., VANNUCCI, M. and FEARN, T. (1998). Multivariate Bayesian variable selection and prediction. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **60** 627–641. MR1626005

CAMERON, G., MUELLBAUER, J. and MURPHY, A. (2006). *Was There a British House Price Bubble? Evidence from a Regional Panel. Mimeo*. Oxford Univ. Press, London.

CAPOZZA, D. R., HENDERSHOTT, P. H., MACK, C. and MAYER, C. J. (2002). Determinants of real house price dynamics. NBER Working Paper 9262.

CARTER, C. K. and KOHN, R. (1994). On Gibbs sampling for state space models. *Biometrika* **81** 541–553. MR1311096

CASE, K. E. and SHILLER, R. J. (2003). Is there a bubble in the housing market? *Brookings Papers on Economic Activity* **2** 299–362.

CHO, S. (2010). Inference of cointegrated model with exogenous variables. SIRFE Working Paper 10–A04.

CLAYTON, J., MILLER, N. and PENG, L. (2010). Price-volume correlation in the housing market: Causality and co-movements. *Journal of Real Estate Finance and Economics* **40** 14–40.

CRESSIE, N. A. C. (1993). *Statistics for Spatial Data. Wiley Series in Probability and Mathematical Statistics*: *Applied Probability and Statistics*. Wiley, New York. MR1239641

DAWID, A. P. (1981). Some matrix-variate distribution theory: Notational considerations and a Bayesian application. *Biometrika* **68** 265–274. MR0614963

DEBARSY, N., ERTUR, C. and LESAGE, J. P. (2012). Interpreting dynamic space–time panel data models. *Stat. Methodol.* **9** 158–171. MR2863605

DI GIACINTO, V., DRYDEN, I., IPPOLITI, L. and ROMAGNOLI, L. (2005). Linear smoothing of noisy spatial temporal series. *J. Math. Stat.* **1** 299–311. MR2400983

DURBIN, J. and KOOPMAN, S. J. (2001). *Time Series Analysis by State Space Methods. Oxford Statistical Science Series* **24**. Oxford Univ. Press, Oxford. MR1856951

ELHORST, J. P. (2001). Dynamic models in space and time. *Geographical Analysis* **33** 119–140.

ENGLE, R. F., HENDRY, D. F. and RICHARD, J.-F. (1983). Exogeneity. *Econometrica* **51** 277–304. MR0688727

ESRI. (2009). ArcMap 9.2. Environmental Systems Resource Institute, Redlands, California.

FRÜHWIRTH-SCHNATTER, S. (1994). Data augmentation and dynamic linear models. *J. Time Series Anal.* **15** 183–202. MR1263889

GALLIN, J. (2008). The long run relationship between housing prices and income: Evidence from local housing markets. *Real Estate Economics* **36** 635–658.

GELFAND, A. E. and GHOSH, S. K. (1998). Model choice: A minimum posterior predictive loss approach. *Biometrika* **85** 1–11. MR1627258

GELMAN, A. (1996). Inference and Monitoring Convergence. In *Introducing Markov Chain Monte Carlo*.

GEORGE, E. I., SUN, D. and NI, S. (2008). Bayesian stochastic search for VAR model restrictions. *J. Econometrics* **142** 553–580. MR2408749

GEWEKE, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In *Bayesian Statistics*, 4 (*PeñíScola*, 1991) (J. Bernardo, J. Berger, A. Dawid and A. Smith, eds.) 169–193. Oxford Univ. Press, New York. MR1380276

GILKS, W. R., RICHARDSON, S. and SPIEGELHALTER, D. J., eds. (1996). *Markov Chain Monte Carlo in Practice*. *Interdisciplinary Statistics*. Chapman & Hall, London. MR1397966

GIUSSANI, B. and HADJIMATHEOU, G. (1991). Modeling regional housing prices in the United Kingdom. *Papers In Regional Science* **70** 201–219.

GOURIEROUX, C. S. and MONFORT, A. (1997). *Time Series and Dynamic Models*. Cambridge Univ. Press, Cambridge.

HOLLY, S., PESARAN, M. H. and YAMAGATA, T. (2010). A spatio-temporal model of house prices in the USA. *J. Econometrics* **158** 160–173. MR2671272

IPPOLITI, L., VALENTINI, P. and GAMERMAN, D. (2012). Space-time modelling of coupled spatio-temporal environmental variables. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **61** 175–200. MR2905058

JOCHMANN, M., KOOP, G., LEON-GONZALEZ, R. and STRACHAN, R. (2013). Stochastic search variable selection in vector error correction models with an application to a model of the UK macroeconomy. *J. Appl. Econometrics* **28** 62–81.

JOHANSEN, S. (1988). Statistical analysis of cointegration vectors. *J. Econom. Dynam. Control* **12** 231–254. MR0986516

JONES, G. L., HARAN, M., CAFFO, B. S. and NEATH, R. (2006). Fixed-width output analysis for Markov chain Monte Carlo. *J. Amer. Statist. Assoc.* **101** 1537–1547. MR2279478

KASS, R. E. and RAFTERY, A. E. (1995). Bayes factors. *J. Amer. Statist. Assoc.* **90** 773–795.

KIM, H., SUN, D. and TSUTAKAWA, R. K. (2001). A bivariate Bayes method for improving the estimates of mortality rates with a twofold conditional autoregressive model. *J. Amer. Statist. Assoc.* **96** 1506–1521. MR1946594

KOOP, G., LEÓN-GONZÁLEZ, R. and STRACHAN, R. W. (2010). Efficient posterior simulation for cointegrated models with priors on the cointegration space. *Econometric Rev.* **29** 224–242. MR2747499

KOOP, G. M., STRACHAN, R. W., VAN DIJK, H. and VILLANI, M. (2006). Bayesian approaches to cointegration. In *The Palgrave Handbook of Theoretical Econometrics* 871–898. Palgrave Macmillan, Basingstoke, UK.

KUETHE, T. and PEDE, V. (2011). Regional housing price cycles: A spatio-temporal analysis using US state-level data. *Regional Studies* **45** 563–574.

LOPES, H. F., SALAZAR, E. and GAMERMAN, D. (2008). Spatial dynamic factor analysis. *Bayesian Anal.* **3** 759–792. MR2469799

LOPES, H. F. and WEST, M. (2004). Bayesian model assessment in factor analysis. *Statist. Sinica* **14** 41–67. MR2036762

LÜTKEPOHL, H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer, Berlin. MR2172368

MALPEZZI, S. (1999). A simple error correction model of housing prices. *Journal of Housing Economics* **8** 27–62.

MARDIA, K. V. (1988). Multidimensional multivariate Gaussian Markov random fields with application to image processing. *J. Multivariate Anal.* **24** 265–284. MR0926357

MARDIA, K. V., KENT, J. T. and BIBBY, J. M. (1979). *Multivariate Analysis*. Academic Press, London. MR0560319

MEEN, G. (1999). Regional house prices and the ripple effect: A new interpretation. *Housing Studies* **14** 733–753.

MEEN, G. (2001). *Modelling Spatial Housing Markets*: *Theory, Analysis and Policy*. Kluwer, Dordrecht, The Netherlands.

MOENCH, E. and NG, S. (2011). A hierarchical factor analysis of U.S. housing market dynamics. *Econom. J.* **14** C1–C24.

MUELLBAUER, J. and MURPHY, A. (1997). Booms and busts in the UK housing market. *Econom. J.* **107** 1701–1727.

OSIEWALSKI, J. and STEEL, M. F. J. (1996). A Bayesian analysis of exogeneity in models pooling time-series and cross-sectional data. *J. Statist. Plann. Inference* **50** 187–206. MR1396456

PESARAN, M. H. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica* **74** 967–1012. MR2238209

PFEIFER, P. E. and DEUTSCH, S. J. (1980). A three-stage iterative procedure for space-time modeling. *Technometrics* **22** 35–47.

PFEIFER, P. E. and DEUTSCH, S. J. (1981). Space-time ARMA modeling with contemporaneously correlated innovations. *Technometrics* **23** 401–409.

PRIMICERI, G. E. (2005). Time varying structural vector autoregressions and monetary policy. *Rev. Econom. Stud.* **72** 821–852. MR2148143

ROSENBERG, B. (1973). Random coefficients models: The analysis of a cross-section of time series by stochastically convergent parameter regression. *Annals of Economic and Social Measurement* **60** 399–428.

SAIN, S. R. and CRESSIE, N. (2007). A spatial model for multivariate lattice data. *J. Econometrics* **140** 226–259. MR2395923

SAIN, S. R., FURRER, R. and CRESSIE, N. (2011). A spatial analysis of multivariate output from regional climate models. *Ann. Appl. Stat.* **5** 150–175. MR2810393

SIMS, C. A. and ZHA, T. (1999). Error bands for impulse responses. *Econometrica* **67** 1113–1155. MR1707461

SPIEGELHALTER, D. J., BEST, N. G., CARLIN, B. P. and VAN DER LINDE, A. (2002). Bayesian measures of model complexity and fit. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **64** 583–639. MR1979380

STRICKLAND, C. M., SIMPSON, D. P., TURNER, I. W., DENHAM, R. and MENGERSEN, K. L. (2011). Fast bayesian analysis of spatial dynamic factor models for large space time data sets. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **60** 1–16.

SUGITA, K. (2009). A Monte Carlo comparison of Bayesian testing for cointegration rank. *Economics Bulletin* **29** 2145–2151.

VAN DIJK, B., FRANSES, P. H., PAAP, R. and VAN DIJK, D. J. C. (2011). Modeling regional house prices. *Applied Economics* **43** 2097–2110.

VERMEULEN, W. and VAN OMMEREN, J. (2009). Compensation of regional unemployment in housing markets. *Economica* **76** 71–88.

WANG, F. and WALL, M. M. (2003). Generalized common spatial factor model. *Biostatistics* **4** 569–582.

DEPARTMENT OF ECONOMICS
UNIVERSITY OF CHIETI-PESCARA
VIALE PINDARO, 42
65127 PESCARA
ITALY
E-MAIL: pvalent@unich.it
            ippoliti@unich.it
            lfontan@unich.it