STRATHCLYDE

DISCUSSION PAPERS IN ECONOMICS

MODEL UNCERTAINTY IN PANEL VECTOR AUTOREGRESSIVE MODELS

BY

GARY KOOP AND DIMITRIS KOROBILIS

NO 14-08

DEPARTMENT OF ECONOMICS
UNIVERSITY OF STRATHCLYDE
GLASGOW

Model Uncertainty in Panel Vector Autoregressive Models

Gary Koop
University of Strathclyde

Dimitris Korobilis
University of Glasgow

August 2014

Abstract

We develop methods for Bayesian model averaging (BMA) or selection (BMS) in Panel Vector Autoregressions (PVARs). Our approach allows us to select between or average over all possible combinations of restricted PVARs where the restrictions involve interdependencies between and heterogeneities across cross-sectional units. The resulting BMA framework can find a parsimonious PVAR specification, thus dealing with overparameterization concerns. We use these methods in an application involving the euro area sovereign debt crisis and show that our methods perform better than alternatives. Our findings contradict a simple view of the sovereign debt crisis which divides the euro zone into groups of core and peripheral countries and worries about financial contagion within the latter group.

Keywords: Bayesian model averaging, stochastic search variable selection, financial contagion, sovereign debt crisis

JEL Classification: C11, C33, C52, G10

1 Introduction

This paper develops Bayesian methods for estimation and model selection with large panel vector autoregressions (PVARs). PVARs are used in several research fields, but are most commonly used by macroeconomists or financial economists working with data for many countries. In such a case, the researcher may want to jointly model several variables for each country using a VAR, but also allow for linkages between countries. Papers such as Dees, Di Mauro, Pesaran and Smith (2007) and Canova and Ciccarelli (2009) emphasize that PVARs are an excellent way to model the manner in which shocks are transmitted across countries and to address issues such as financial contagion that have played an important role in recent years.[1] As the global economy becomes more integrated, examining such issues is increasingly important for the modern applied economist.

In that respect, we consider the case where we have N countries, each with G macroeconomic variables observed for T periods. In such a setup, the PVAR is the ideal tool for examining the international transmission of macroeconomic or financial shocks. A major difference between a PVAR and a univariate dynamic panel regression is that the VAR specification can explicitly allow an endogenous variable of interest (e.g. the i-th macroeconomic variable for the j-th country) to depend on several lags of: i) the endogenous variable itself; ii) other macroeconomic variables of that country; and iii) macroeconomic variables of all other $N-1$ countries. Thus, the PVAR can uncover all sorts of dynamic or static dependencies between countries or the existence of heterogeneity in coefficients on the macroeconomic variables of different countries. Additionally, given the autoregressive structure of a PVAR, concerns about endogeneity are eliminated and the usual macroeconomic exercises involving multiple-period projections in the future (e.g. forecast error variance decompositions, or impulse responses) can be implemented.

However, this flexibility of the PVAR comes at a cost. The researcher working with an unrestricted PVAR with P lags must estimate $K = (NG)^2 P$ autoregressive coefficients, coefficients on any deterministic terms, and the $\frac{NG(NG+1)}{2}$ free parameters in the error covariance matrix. In most cases, when the number of countries N is moderate or large, the number of parameters might exceed the number of observations available for estimation. Accordingly, interest centers on various restricted PVAR models.

Many such restrictions are possible (and the methods developed in this paper can easily be generalized to deal with any of them), but we focus on ones used, e.g., in Canova and Ciccarelli (2013). These restrictions pertain to the absence of dynamic interdependencies (DI), static interdependencies (SI) and cross-section heterogeneities (CSH). DIs occur when one country's lagged variables affect another country's variables. SIs occur when the correlations between the errors in

[1] Canova and Ciccerelli (2013) offers an excellent survey of the PVAR literature. The reader is referred to this paper for an extensive list of papers using PVAR methods.

two countries' VARs are non-zero. CSHs occur when two countries have VARs with different coefficients (i.e. homogeneity arises when the coefficients on the own lagged variables for the two countries are exactly the same).[2]

The total number restrictions on DIs, SIs and CSHs we may wish to impose is potentially huge. For instance, in our empirical work we have 10 countries in the PVAR which leads to 90 DI restrictions to examine, 45 SI restrictions, and 45 CSH restrictions which can be imposed in any combination. Thus, the researcher is faced with an over-parameterized unrestricted model and a large number of potentially interesting restricted models. This situation is familiar in the BMA literature. Following this literature we rely on Markov Chain Monte Carlo (MCMC) methods so as to avoid the huge computational burden of exhaustively estimating every restricted model. MCMC methods allow for the joint estimation of the PVAR parameters in each model along with the probabilities attached to each model. Such algorithms are far from new in the literature. There are several similar approaches used in traditional regression models, with notable early contributions by George and McCulloch (1993, 1997) and Raftery, Madigan and Hoeting (1997). In economics, BMA algorithms using MCMC methods have been influential, particularly in the problem of finding relevant predictors for economic growth (e.g., among many others, Fernández, Ley and Steel, 2001a,b; Eicher, Papageorgiou and Raftery, 2010; Ley and Steel, 2012).

With regression models, there is a single dependent variable and the restrictions considered are typically simple ones (e.g. a coefficient is set to zero). With VARs, one has a vector of dependent variables, but the existing literature has still worked with simple restrictions. In the VAR literature, stochastic search variable selection (SSVS)[3] methods have proved popular. The pioneering paper is George,

[2]In the panel regression literature it is common to assume homogeneity of coefficients across countries (i.e. the coefficients on the explanatory variables are the same in each country). In the PVAR literature, where the explanatory variables are lags of the dependent variables, this is harder to justify. Consider, for instance, a differential Taylor rule such as the one usually defined in the exchange rate literature (Molodtsova and Pappel, 2009) involving interest rates (i_t), inflation (π_t) and the output gap ($\overline{y}_t$):

$$i_t - i_t^* = \rho_0 + \rho_1 \pi_t + \rho_1^* \pi_t^* + \rho_2 \overline{y}_t + \rho_2^* \overline{y}_t^* + \rho_3 i_{t-1} + \rho_3^* i_{t-1}^*,$$

where variables without stars denote domestic quantities, and variables with stars foreign quantities. Homogeneity (pooling) is usually defined as the case where inflation and the output gap have the same coefficients in both countries, i.e. $\rho_1 = \rho_1^*$ and $\rho_2 = \rho_2^*$. It is not customary to consider the case where the coefficients on lagged interest rates are equal (i.e. $\rho_3 = \rho_3^*$) since it is expected that dynamics of variables of different countries will not be homogeneous. Additionally, working with random effects or mean group estimators is not relevant in high dimensional PVARs, unless we can assume sufficiently large number of observations T; see Canova and Ciccarelli (2013) for a discussion of such estimation issues.

[3]We use this as a general term for methods which use a hierarchical prior which allows a variable to be selected (i.e. its coefficient estimated in an unconstrained manner) or not selected (i.e. its

Sun and Ni (2008) and recent VAR extensions and applications include Koop (2013) and Korobilis (2013). With PVARs, we have many dependent variables and the restrictions can be more complicated. From an econometric perspective, the contribution of this paper lies in extending previous VAR methods to deal with the PVAR and the more complicated DI, SI and CSH restrictions. Since we are not selecting a single variable, as the V in SSVS implies, but rather a particular specification of a restricted PVAR, we name our algorithm Stochastic Search Specification Selection (S^4) for PVARs.

The other contribution of the paper is to use these methods in an empirical study of financial contagion during the recent euro area sovereign debt crisis. Using data on sovereign bond spreads, bid ask spreads and industrial production for euro area countries, we use our PVAR methods to investigate the nature and extent of spillovers within the euro area. We do find there are extensive links between countries. However, these links do not correspond to a conventional division of euro area countries into core and periphery countries and an accompanying fear of financial contagion within the periphery countries. We do find two groups of countries which are, in a sense we describe below, homogeneous. But the division does not correspond closely with the conventional core/periphery grouping. Furthermore, we find spillovers from one country to another, but these spillovers are largely within the core countries or reflect core countries shocks propagated to periphery countries, rather than the reverse.

The paper is organized as follows. In the following section we define the PVAR and the restrictions of interest. The third section describes our S^4 methods for doing BMA and BMS with PVARs (with additional details provided in the Technical Appendix). The fourth section contains a brief Monte Carlo study showing that our methods are effective at choosing PVAR restrictions. Section 5 contains our empirical application and the sixth section concludes.

2 Panel VARs

Let y_{it} denote a vector of G dependent variables for country i ($i = 1, .., N$) at time t ($t = 1, .., T$) and $Y_t = (y'_{1t}, .., y'_{Nt})'$. A VAR[4] for country i can be written as:

$$y_{it} = A_{1,i} Y_{t-1} + ... + A_{P,i} Y_{t-P} + \varepsilon_{it} \quad (1)$$

where $A_{p,i}$ are $G \times NG$ matrices for each lag $p = 1, ..., P$, and ε_{it} are uncorrelated over time and are distributed as $N(0, \Sigma_{ii})$ with Σ_{ii} covariance matrices of dimension $G \times G$. Additionally, we define $cov(\varepsilon_{it}, \varepsilon_{jt}) = E(\varepsilon_{it}, \varepsilon_{jt}) = \Sigma_{ij}$ to be the covariance

coefficient set to zero or shrunk to being nearly zero). Other terminologies such as "spike and slab" priors are sometimes used.

[4]For ease of exposition, the formulae in this section for our VARs do not include deterministic terms or exogenous variables. These can be added with straightforward extensions of the formulae.

matrix between the errors in the VARs of country i and country j. We refer to this specification as the unrestricted PVAR.

Note that the unrestricted PVAR is very general and that lagged variables from any country can influence any other country (e.g. lagged values of country 1 variables can impact on current country 2 variables) and the magnitude of such influences are completely unrestricted (e.g. events in country 1 can have different impacts on country 2 than on country 3). Similarly, contemporaneous relationships, modelled through the error covariance matrices, are unrestricted so that, e.g., shocks in country 1 can be strongly correlated with shocks in country 2, but weakly correlated with shocks in country 3.

Unrestricted PVARs such as (1) can suffer from concerns about over-parameterization due to the high dimensionality of the parameter space. For instance, Canova and Ciccarelli (2009) use data on four dependent variables ($G = 4$) for the G-7 countries ($N = 7$) and one lag ($P = 1$). An unrestricted PVAR with such choices would have 784 VAR coefficients and 406 error variances and covariances to estimate.

One strand of the macro VAR literature relies on shrinkage and model selection methods to deal with such high dimensional parameter spaces. For example, Banbura et al. (2010) uses the Minnesota prior (Littermann, 1986) to estimate VARs of large dimension and imposes shrinkage towards zero of irrelevant coefficients. Papers such as Carriero, Clark and Marcellino (2011), Carriero, Kapetanios and Marcellino (2009), Giannone, Lenza, Momferatou and Onorante (2010), Gefang (2013), Koop (2013) and Korobilis (2013) use similar shrinkage and model selection methods to estimate VARs with hundreds or even thousands of coefficients. BMS and BMA applications in this strand of the literature simply restrict each individual coefficient to be zero (or not). However, in a PVAR there are a variety of restrictions of interest which reflect the panel nature of the data. These are ignored in conventional large VAR approaches. Therefore, there can be gains in not treating a PVAR in the same manner as a standard large VAR. Canova and Ciccarelli (2013) provide an excellent survey of the various restrictions used and what their implications are. In the introduction, we explained briefly the DI, SI and CSH restrictions considered in this paper. Here we provide precise definitions.

DIs refer to links across countries through PVAR coefficients. In (1), the endogenous variables for each country depend on the lags of the endogenous variables for every country. It is often of interest to investigate if DIs exist and, if not, to estimate restricted PVARs which lack such interdependencies. To formally define DIs between countries j and k, we partition the PVAR coefficient matrices (for $p = 1, .., P$) into $G \times G$ matrices $A_{p,jk}$ which control whether lags of country k dependent variables enter the VAR for country j. That is,

$$A_p = \begin{bmatrix} A_{p,1} \\ A_{p,2} \\ \vdots \\ A_{p,N} \end{bmatrix} = \begin{bmatrix} A_{p,11} & A_{p,12} & \cdots & A_{p,1N} \\ A_{p,21} & A_{p,22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & A_{p,(N-1)N} \\ A_{p,N1} & \cdots & A_{p,N(N-1)} & A_{p,NN} \end{bmatrix}. \tag{2}$$

Within the unrestricted VAR, we can define $N(N-1)$ restrictions which imply there are no DIs from country k to j by imposing the restriction that $A_{1,jk} = .. = A_{P,jk} = 0$ for $j,k = 1,..,N$ and $j \neq k$. Note that the algorithm developed in this paper will allow for selection between a large number of restricted models since we are allowing for every possible configuration of DIs between countries. Using the G-7 countries as an example, our algorithm could select a restricted PVAR that has France exhibiting DIs with Germany, USA and Italy but not Canada, Japan and the UK. Another restricted PVAR would have France exhibiting DIs with Germany, USA, Italy and Canada but not Japan and the UK, etc.. Allowing for every country to have DIs with any or all of the $N-1$ remaining countries leads to $N(N-1)$ restricted PVARs that our algorithm can choose between when investigating DIs. Note that it is possible for such linkages between two countries to flow in one direction only. For instance, it is possible that lagged German variables influence French variables (and, thus, there are DIs from Germany to France), but that lagged French variables do not influence German variables (and, thus, there are no DIs from France to Germany).

SIs are modelled through the error covariance matrix. If $\Sigma_{jk} = 0$, then there are no SIs between countries j and k. We can define $\frac{N(N-1)}{2}$ restricted PVARs which impose $\Sigma_{jk} = 0$ for $j,k = 1,..,N$ and $j \neq k$. In contrast to the DI restrictions, these are always symmetric. For instance, if there are SI's from Germany to France, they will also exist from France to Germany.

CSH occurs if the VAR coefficients differ across countries.[5] Such homogeneity occurs between two countries if $A_{p,jj} = A_{p,kk}$ for $j \neq k$ and $p = 1,..,P$. Thus, we can construct $\frac{N(N-1)}{2}$ restricted PVARs which impose homogeneity between two different countries. We could also consider restrictions which impose homogeneity of error covariances, but we do not do so in practice since such restrictions are less likely to be reasonable in macroeconomic and financial applications than homogeneity restrictions involving VAR coefficients.

Table 1 contains a list of the restrictions considered in this paper.

[5]Note that our definition of cross-country homogeneity involves only the VAR part of the model for each country. For instance, it says country 1's lagged dependent variables influence country 1's variables in the same manner as country 2's lagged dependent variables influence country 2's variables. It does not involve restricting, say, country 3's lagged dependent variables to have the same impact on country 1 as on country 2. Such an alternative could be handled by simply re-defining the restrictions.

Table 1: Possible Specification Restrictions in PVARs		
Name	Restriction	Number
No DIs from country k to j	$A_{1,jk} = ... = A_{P,jk} = 0$	$N(N-1)$
No SIs between countries k and j	$\Sigma_{jk} = 0$	$\frac{N(N-1)}{2}$
No CSH between countries k and j	$A_{p,jj} = A_{p,kk}$ $\forall\, p = 1,..,P$	$\frac{N(N-1)}{2}$

Note that the number of restrictions we have described is potentially huge. And there are many other restrictions which might be interesting in the context of a particular empirical application. For instance, global VARs (see, e.g., Dees, Di Mauro, Pesaran and Smith, 2007) can be obtained by imposing restrictions on A_p such that only cross-country averages enter the PVAR. In the empirical work of this paper, we will not consider global VAR restrictions, but note that they can easily be accommodated in our approach.

3 Stochastic Search Specification Selection (S^4)

To define our S^4 algorithm, we begin by writing the PVAR more compactly as:

$$Y_t = Z_t\alpha + \varepsilon_t, \tag{3}$$

where $Z_t = I_{NG} \otimes X_t'$, $X_t' = \left(I, Y_{t-1}', Y_{t-2}', ..., Y_{t-P}'\right)'$, $\alpha = (vec(A_1), ..., vec(A_P))'$ is a $K \times 1$ vector containing all the PVAR coefficients, $K = 1, ..., P(NG)^2$, and Y_t, ε_t are $NG \times 1$ vectors (uncorrelated over time) with $\varepsilon_t \sim N(0, \Sigma)$ for $t = 1, .., T$. This is the unrestricted PVAR.

The basic idea underlying SSVS as done, e.g., in George, Sun and Ni (2008), can be explained simply. Let α_j denote the j^{th} element of α. SSVS specifies a hierarchical prior (i.e. a prior expressed in terms of parameters which in turn have a prior of their own) which is a mixture of two Normal distributions:

$$\alpha_j | \gamma_j \sim \left(1 - \gamma_j\right) N\left(0, \underline{\tau}_1^2\right) + \gamma_j N\left(0, \underline{\tau}_2^2\right), \tag{4}$$

where $\gamma_j \in \{0, 1\}$ is an unknown parameter estimated from the data. A Bernoulli prior is used for γ_j. The variable selection property of this prior arises by setting $\underline{\tau}_1^2$ to be small (or zero) and $\underline{\tau}_2^2$ to be large (or infinite). Thus, if $\gamma_j = 1$, the prior has a large variance and, if $\gamma_j = 0$, the prior has a small variance. Large variance priors are relatively noninformative, allowing for a coefficient to be estimated in an unrestricted fashion. Small variance priors are informative, shrinking the coefficient towards the prior mean (which, in this case, is zero). In the limiting case when the prior variance goes to zero, the prior becomes a spike at zero and the accompanying coefficient is set to zero and the associated variable is deleted from the model.

Another way of looking at such an approach is considered in Korobilis (2013) who writes the model as:

$$Y_t = Z_t \Gamma \alpha + \varepsilon_t \tag{5}$$

where Γ is a matrix which can be used to do BMA or BMS In conventional SSVS, Γ is a diagonal matrix with diagonal elements $\gamma_j \in \{0, 1\}$ corresponding to each VAR coefficient.

The basic ideas underlying our S^4 algorithm can be expressed in terms of (4) and (5). In this section, we outline how we do this, partly relying on a three country example. The general case and other additional details are given in the Technical Appendix.

We define the $N(N-1)$ vector γ^{DI}, the $\frac{N(N-1)}{2}$ vector γ^{SI}, and the $\frac{N(N-1)}{2}$ vector γ^{CSH}, which control the DI, SI, and CSH restrictions, respectively and let $\gamma = \left(\gamma^{DI}, \gamma^{SI}, \gamma^{CSH}\right)$.

Handling the DI and SI restrictions is fairly easy, since each involves restricting a specific matrix of parameters to be zero (or not). For the DIs, γ^{DI} is made up of elements $\gamma_{jk}^{DI} \in \{0, 1\}$, $j = 1, ..., N$, $k = 1, ..., N$, $j \neq k$. If $\gamma_{jk}^{DI} = 0$, then the coefficients on the lags of all country k variables in the VAR for country j are set to zero. Using a simple extension of the hierarchical prior in (4) and the methods of George, Sun and Ni (2008), it is straightforward to produce MCMC draws of γ^{DI}. The only difference between our approach and that of George, Sun and Ni (2008) is that each γ_{jk}^{DI} will apply to a whole block of parameters instead of a single parameter.

For the SIs, γ^{SI} is made up of elements $\gamma_{jk}^{SI} \in \{0, 1\}$, $j = 1, ..., N-1$, $k = j+1, ..., N$. If $\gamma_{jk}^{SI} = 0$ then the block of the PVAR error covariance matrix relating to the covariance between countries j and k is set to zero. In contrast to conventional SSVS, γ_{jk}^{SI} will restrict an entire block of the error covariance matrix to be zero, rather than a single element, but this involves only trivial changes to the algorithm of George, Sun and Ni (2008).

Handling restrictions which do not simply restrict a vector or matrix of coefficients to be zero is more complicated, and treatment of this issue is a contribution of this paper. CSH restrictions take this form. There are $\frac{N(N-1)}{2}$ such restrictions and we investigate whether they hold by introducing restriction selection matrices: $\Gamma_{j,k}$ for $j = 1, .., N-1$ and $k = j+1, .., N$. $\Gamma_{j,k}$ contains one dummy variable, $\gamma_{jk}^{CSH} \in \{0, 1\}$, which is used to estimate whether cross-country homogeneity exists between countries j and k. For instance, if we had $N = 3$ and each of the three VARs which make up the PVAR only had one coefficient ($\alpha = (\alpha_1, \alpha_2, \alpha_3)'$), then we can define the matrices:

$$\Gamma_{1,2} = \begin{bmatrix} \gamma_{12}^{CSH} & 1-\gamma_{12}^{CSH} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \Gamma_{1,3} = \begin{bmatrix} \gamma_{13}^{CSH} & 0 & 1-\gamma_{13}^{CSH} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \Gamma_{2,3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \gamma_{23}^{CSH} & 1-\gamma_{23}^{CSH} \\ 0 & 0 & 1 \end{bmatrix},$$

where $\gamma^{CSH} = \left(\gamma_{12}^{CSH}, \gamma_{13}^{CSH}, \gamma_{23}^{CSH}\right)$ is the original vector of CSH restrictions between countries 1 and 2, countries 1 and 3, and countries 2 and 3, respectively. If homogeneity exists between countries 1 and 2 then $\gamma_{12}^{CSH} = 0$ and $\Gamma_{1,2}\alpha = \begin{bmatrix} \alpha_2 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}$ so that the first and second coefficients are restricted to be equal to one another. If instead these two countries are heterogeneous, $\gamma_{12}^{CSH} = 1$ and $\Gamma_{1,2}$ is the identity matrix such that $\Gamma_{1,2}\alpha = \alpha$ and the coefficients are left unrestricted. By defining matrices $\Gamma_{1,3}$ and $\Gamma_{2,3}$ in a similar fashion we can impose analogous restrictions involving country 3.

If we define:

$$\Gamma = \Gamma_{1,2} \times \Gamma_{1,3} \times \Gamma_{2,3}.$$

then we obtain a selection matrix that covers all possible combinations of CSH restrictions. For instance, assume that there is homogeneity between countries 1 and 3 (so that $\gamma_{13}^{CSH} = 0$) and the coefficients of countries 1 and 2, and countries 2 and 3 are heterogeneous ($\gamma_{12}^{CSH} = \gamma_{23}^{CSH} = 1$). In this case, it is easy to see that Γ takes the form

$$\Gamma = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

so that the restricted coefficients matrix is $\Gamma\alpha = (\alpha_3, \alpha_2, \alpha_3)$. In this case, the first and third countries coefficients are the same, thus imposing homogeneity between them. If $\gamma_1^{CSH} = \gamma_2^{CSH} = \gamma_3^{CSH} = 0$ then there is homogeneity among all countries and in this case $\Gamma\alpha = (\alpha_3, \alpha_3, \alpha_3)'$.

We can generalize the procedure above when we have N countries to impose (or not) the $N(N-1)/2$ possible CSH restrictions and the DI restrictions if we write the PVAR of (4) as:

$$\begin{aligned} Y_t &= Z_t \prod_{j=1}^{N-1} \prod_{k=j+1}^{N} \Gamma_{j,k}\alpha + \varepsilon_t \\ &= Z_t \Gamma \alpha + \varepsilon_t. \end{aligned} \tag{6}$$

Once the PVAR is transformed in this way, sampling from the conditional posterior of the restricted coefficients $\widetilde{\alpha} = \prod_{j=1}^{N-1} \prod_{k=j+1}^{N} \Gamma_{j,k}\alpha = \Gamma\alpha$ becomes a straightforward problem. In particular, conditional upon draws of the restriction indicators, we have a particular restricted PVAR. The parameters of this specific PVAR can be drawn using standard formulae for restricted VAR models. We provide additional details in the Technical Appendix.

Using this MCMC algorithm, we can find the posterior mode for γ and this can be used to select the optimal restricted PVAR, thus doing BMS. Or, if we simply average over all draws provided by the MCMC algorithm we are doing BMA. Our empirical results use the BMA approach.

4 Monte Carlo Study

In order to demonstrate the performance of our algorithm, we carry out a small Monte Carlo study. We consider a case where the number of observations is fairly small relative to the number of parameters being estimated and a variety of restrictions hold. In particular, we generate 1,000 artificial data sets, each with $T = 50$ from a PVAR with $N = 3$, $G = 2$ and $P = 1$. Using the notation of (2), the PVAR parameters are set to the values:

$$A_1 = \begin{bmatrix} 0.7 & 0 & 0.2 & 0.2 & 0 & 0 \\ 0 & 0.7 & 0.3 & 0.3 & 0 & 0 \\ 0 & 0 & 0.2 & 0.4 & 0 & 0 \\ 0 & 0 & 0 & 0.3 & 0 & 0 \\ 0.3 & -0.4 & 0 & 0 & 0.2 & 0.4 \\ 0.2 & 0.4 & 0 & 0 & 0 & 0.3 \end{bmatrix}, \Sigma = \begin{bmatrix} 1 & 0 & -0.5 & -0.5 & 0 & 0 \\ 0 & 1 & -0.5 & -0.5 & 0 & 0 \\ -0.5 & -0.5 & 1 & 0.5 & 0 & 0 \\ -0.5 & -0.5 & 0.5 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

The structure above implies that we have DIs from country 2 to country 1 and from country 1 to country 3. We have SIs between countries 1 and 2 and cross-sectional homogeneity between countries 2 and 3. Put another way, the data generating process imposes the following restrictions that we hope our S^4 algorithm will find:

1. $A_{1,13} = A_{1,21} = A_{1,23} = A_{1,32} = 0$
2. $A_{1,22} = A_{1,33}$
3. $\Sigma_{13} = \Sigma_{23} = 0$

For each of our 1,000 artificial data sets we produce 55,000 posterior draws using our MCMC algorithm and discard the first 5,000 as burn in draws. Results pass standard convergence diagnostics (e.g. inefficiency factors reveal that retaining 50,000 draws is more than enough for accurate posterior inference). The relatively noninformative priors we use are described in the Technical Appendix.

To give the reader an idea of how well our algorithm is estimating the PVAR parameters, the following matrices contain the averages (over the 1,000 artificial data sets) of their posterior means.

$$A_1 = \begin{bmatrix} .67 & .11 & .16 & .26 & -.01 & .00 \\ -.01 & .68 & .30 & .32 & .01 & .00 \\ .00 & .00 & .14 & .37 & .00 & .00 \\ .00 & .00 & .01 & .28 & .00 & .01 \\ .27 & -.38 & -.01 & .00 & .17 & .34 \\ .21 & .39 & .00 & .01 & .02 & .29 \end{bmatrix}, \Sigma = \begin{bmatrix} .98 & -.02 & -.47 & -.48 & .00 & -.01 \\ -.02 & .99 & -.46 & -.45 & -.01 & .00 \\ -.47 & -.46 & 1.24 & .42 & -.01 & .00 \\ -.48 & -.45 & .42 & 1.24 & .01 & -.01 \\ .00 & -.01 & -.01 & .01 & 1.00 & -.01 \\ -.01 & .00 & .00 & -.01 & -.01 & .90 \end{bmatrix}.$$

Considering the relatively small sample size, these posterior means are quite close to the true values used to generate the data sets.

For comparison, the following matrices present ordinary least squares (OLS) estimates averaged over the 1,000 artificial data sets:

$$A_1 = \begin{bmatrix} .64 & .05 & .22 & .21 & -.03 & -.01 \\ .03 & .65 & .37 & .27 & .04 & .03 \\ .02 & .06 & .08 & .44 & .00 & -.05 \\ -.07 & -.04 & -.05 & .20 & .07 & .04 \\ .32 & -.39 & .01 & .03 & .10 & .37 \\ .18 & .31 & -.04 & .05 & .03 & .22 \end{bmatrix}, \Sigma = \begin{bmatrix} 1.02 & -.05 & -.48 & -0.46 & .02 & -.02 \\ -.05 & .96 & -.46 & -0.42 & -.01 & .03 \\ -.48 & -.46 & 1.43 & .41 & -.01 & .02 \\ -.46 & -.42 & .41 & 1.40 & .01 & -.01 \\ .02 & -.01 & -.01 & .01 & 1.00 & -.01 \\ -.02 & .03 & .02 & -.01 & -.01 & .87 \end{bmatrix}.$$

The OLS estimates are similar to the ones produced by our S^4 algorithm. However, note that the OLS estimates do not do as good a job of shrinking to zero the parameters which are truly zero.

We now turn to the issue of how accurate the S^4 algorithm is in picking the correct restrictions. Remember that the restrictions are controlled through the S^4 dummy variables so that, for instance, $\gamma_{2,3}^{CSH} = 0$ indicates that countries 2 and 3 are homogeneous. In our MCMC algorithm, the proportion of draws of $\gamma_{2,3}^{CSH} = 0$ will be an estimate of the posterior probability that countries 1 and 2 are homogeneous and, thus, that $A_{1,22} = A_{1,33}$. Thus, we will use notation where $p(A_{1,22} = A_{1,33})$ is the posterior probability that countries 1 and 2 are homogeneous, averaged over the 1000 artificial data sets (and adopt the same notational convention for the other restrictions).

With regards to the DI restrictions we find the following:

$$\begin{array}{l} p(A_{1,12} = 0) = .108 \\ p(A_{1,13} = 0) = .997 \\ p(A_{1,21} = 0) = .981 \\ p(A_{1,23} = 0) = .994 \\ p(A_{1,31} = 0) = .167 \\ p(A_{1,32} = 0) = .999 \end{array}.$$

It can be seen that the S^4 algorithm is doing a very good job of picking up the correct DI restrictions.

With regards to the SI restrictions we find the following:

$$\begin{array}{c} p\left(\Sigma_{12}=0\right)=.144 \\ p\left(\Sigma_{13}=0\right)=.808 \\ p\left(\Sigma_{23}=0\right)=.836 \end{array}.$$

Here S^4 is also doing a good job in picking up the correct restrictions, although the probabilities are smaller than those found for the DI restrictions.

With regards to the CSH restrictions we find the following:

$$\begin{array}{c} p\left(A_{1,11}=A_{1,22}\right)=.309 \\ p\left(A_{1,11}=A_{1,33}\right)=.304 \\ p\left(A_{1,22}=A_{1,33}\right)=.982 \end{array}.$$

S^4 is doing well at picking out the correct cross-sectional homogeneity restriction between countries 2 and 3.

Overall, we find the results of our Monte Carlo study reassuring. This exercise involved a sample size of only $T = 50$ observations in a PVAR with 57 unknown parameters. Therefore, our S^4 algorithm is doing well at picking the correct restrictions in a case where the number of observations is small relative to the number of parameters. We repeated this exercise with $T = 100$ but do not report results here since the probabilities of the restrictions correctly holding are very nearly one in every case.

5 Empirical Application

The issues of financial contagion and cross-country spillovers between sovereign debt markets in euro area economies have figured prominently in debates about the euro area debt crisis. A few examples of recent papers are Arghyrou and Kontonikas (2012), Bai, Julliard and Yuan (2012), De Santis (2012) and Neri and Ropele (2013). A common strategy in these papers (and many others) is to develop a modelling approach involving sovereign bond spreads (reflecting credit risk considerations), bid-ask spreads (to reflect liquidity considerations) and a macroeconomic variable. Discussion is often framed in terms of core (Germany, Netherlands, France, Austria, Belgium and Finland) and periphery (Greece, Ireland, Portugal, Spain and Italy) countries.

Inspired by this literature, we use monthly data from January 1999 through December 2012 on the 10 year sovereign bond yield, the percentage change in industrial production and the average bid-ask spread averaged across sovereign bonds of differing maturities for the core and periphery countries. Following a common practice, we take spreads relative to German values and, hence, we leave Germany out of our set of countries. Because the 10-year bond yields and the associated bid-ask spreads are nonstationary time series, we first difference them. When we produce impulse responses, we transform back to levels so they do measure responses of the spreads themselves. Thus, we have 168 monthly observations for 3 variables for 10 countries. We include an intercept in each equation. Our PVARs have a lag length of one, which is a reasonable assumption for financial variables. Even so, the unrestricted PVAR has 1395 parameters to estimate and is seriously overparameterized. Complete details about the priors are given in the Technical Appendix.

Remember that our full approach involves working with the unrestricted PVAR with the S^4 prior which allows for selection (or not) of restrictions involving dynamic interdependencies (DI), static interdependencies (SI) and cross-section homogeneities (CSH). Inspired by similar choices in Canova and Ciccarelli (2009), we begin estimating the following models:

1. M1: This is the full model with DI, SI and CSH restriction search.
2. M2: This is the model with DI and SI restriction search (no search for CSH).
3. M3: This is the model with DI restriction search (no search for SI and CSH).
4. M4: This is the model with CSH restriction search (no search for DI and SI).
5. M5: This is the model with SI restriction search (no search for DI and CSH).
6. M6: This is the model which reduces the PVAR to 10 individual country VARs (i.e. DI and SI restrictions are imposed and not searched - no CSH restrictions are applied).
7. M7: This is M6 with CSH additionally imposed (i.e. individual country VARs which are also homogeneous).
8. M8: This is the full unrestricted PVAR model without any restriction searches (i.e. treating it as a large VAR).

Models M2 through M8 are obtained by restricting the elements of γ as appropriate. For instance, M2, where we do not search for CSH restrictions, is obtained by setting $\gamma_{jk}^{CSH} = 1$ for all j and k, but otherwise is identical to M1 in every aspect. M6 is obtained by setting $\gamma_{jk}^{DI} = \gamma_{jk}^{SI} = 0$ for all possible j and k, but otherwise is identical to M1, etc. Thus, we can be certain that any differences across models are solely due to differences in which restrictions are imposed.

We begin by presenting information on which of M1 through M8 is supported by the data using two popular methods of model comparison. Table 2 presents the Bayesian information criterion (BIC) and Deviance information criterion (DIC) for each model. BIC was derived as an approximation to the log of the marginal likelihood. DIC was developed in Spiegelhalter, Best, Carlin and van der Linde (2002) and is an increasingly popular model selection criterion when MCMC methods are used.

Table 2: Model fit

Method	M1	M2	M3	M4	M5	M6	M7	M8
BIC	-2365	-1149.86	-1132.80	-2018.24	-271.92	-1444.36	-1885.84	-271.39
DIC	-59.07	-52.22	-52.22	-56.71	-50.28	-53.12	-55.43	-50.29

The main message from Table 2 is that our full approach, M1, does best and the large VAR approach (M8) does worst, indicating that our S^4 prior which takes into account the panel structure of the model can lead to big improvements. More subtly, it can be seen that most of the benefits from using the S^4 prior (relative to a large VAR) comes from the ability to impose cross-sectional homogeneities. This follows from the fact that M4, which only allows for the imposition of such homogeneities, is the second best model and is appreciably better than any of the approaches other than M1. Nevertheless, using an S^4 prior which allows for DI and SI restriction search does have appreciable value since M1 is clearly better than M4. The ability to impose static interdependencies is of little benefit in this data set since M5 (which only allows for their imposition) is only slightly better than M8.

What kind of restrictions does our preferred M1 model find? Table 3 addresses this question. Note that we have 90 possible DI, 45 CSH and 45 SI restrictions. Recall that we impose restrictions through γ which is a vector of dummy variables. We classify a restriction as being imposed if the MCMC algorithm calculates the probability that the appropriate element of γ is zero to be greater than a half. Otherwise we classify the restriction as not being imposed. Our model is imposing a large number of them, so that it is easier for Table 3 to list the cases where the restrictions are *not* imposed. For the case of DI and SI, these unrestricted cases are where there are interlinkages between countries. So an examination of Table 3 will clearly show where such linkages exist. Country pairs not listed in Table 3 are found to be not interlinked.

Consider first the cross-sectional homogeneities. This is the category of restrictions which is most often rejected. 24 of the 45 possible restrictions are not imposed. By examining which countries are not listed in the CSH column of Table 3, it can be seen that the countries are being divided into two groups. S^4 is finding the VAR coefficients of Austria, Belgium, Finland, Netherlands, Portugal and Spain to be similar enough to one another so as to impose the CSH restrictions. And another four (France, Greece, Ireland and Italy) are also homogeneous with one another, but different from the first homogeneous group. Thus, similar to the conventional core versus periphery split used in this literature, we are finding two groups of countries. But our division into two groups bears little resemblance to the core versus periphery division. The first group, in particular, adds the typically-peripheral Portugal and Spain to the core countries. The second group adds France, typically considered a core country, to three peripheral countries. We stress that our definition of cross-sectional homogeneities only involves own country variables and not linkages between countries. For instance, a finding that France and Greece are homogeneous means that a VAR containing only French variables and a VAR containing only Greek variables have very similar estimated coefficients. Such a finding would say nothing about how other-country variables impact on France or Greece. Nevertheless, it is striking that we are finding such homogeneity, but that the resulting grouping does not coincide with the conventional core versus periphery division.

Table 3 shows that many SIs exist. The main pattern here is that the small countries of Austria, Belgium and Finland have SIs with every other country. These three countries account for 24 out of 25 SIs listed in the table. The only other SI is between France and Greece. This finding that small countries are quickly affected by happenings elsewhere in the euro area is sensible. However, it is in contradiction with some versions of the financial contagion story which would argue that events in one peripheral country could quickly spillover to other peripheral countries. Note that, with the single France-Greece exception, none of the peripheral countries exhibits SIs with any country other than Austria, Belgium and Finland.

It is worth stressing that our definition of SIs implies, e.g., that the entire $G \times G$ block of the error covariance matrix relating to covariances between France and Greece is non-zero. So we do not present a more refined study of the nature of these contemporaneous linkages. For instance, we cannot make statements such as: "we are finding SIs between the French and Greek bond yields, but not between French and Greek industrial production." Adding such refinements would be a straightforward extension of our approach, but would lead to a much larger model space.

Finally consider the DIs. Remember that these may go from one country (labelled "From" in Table 3) to another country (labelled "To") but do not have to go in the reverse direction. So we find that lagged French variables can appear in the VAR for Greece, but not vice versa. The main pattern is that the peripheral countries lagged dependent variables never appear in any of the core countries' VARs. That is, there are many DIs in Table 3, but it is never the case that occurrences in peripheral countries are driving variables in core countries (nor other peripheral countries). Another interesting finding is that Portugal does not appear in this column of Table 3 at all and Spain only appears once. Again, we are finding a story which is not consistent with two common views of the euro zone. We are not finding there is a reasonably homogenous group of core and periphery countries. Nor are we finding support for a financial contagion story where happenings in the periphery spill over to the core or other peripheral countries.

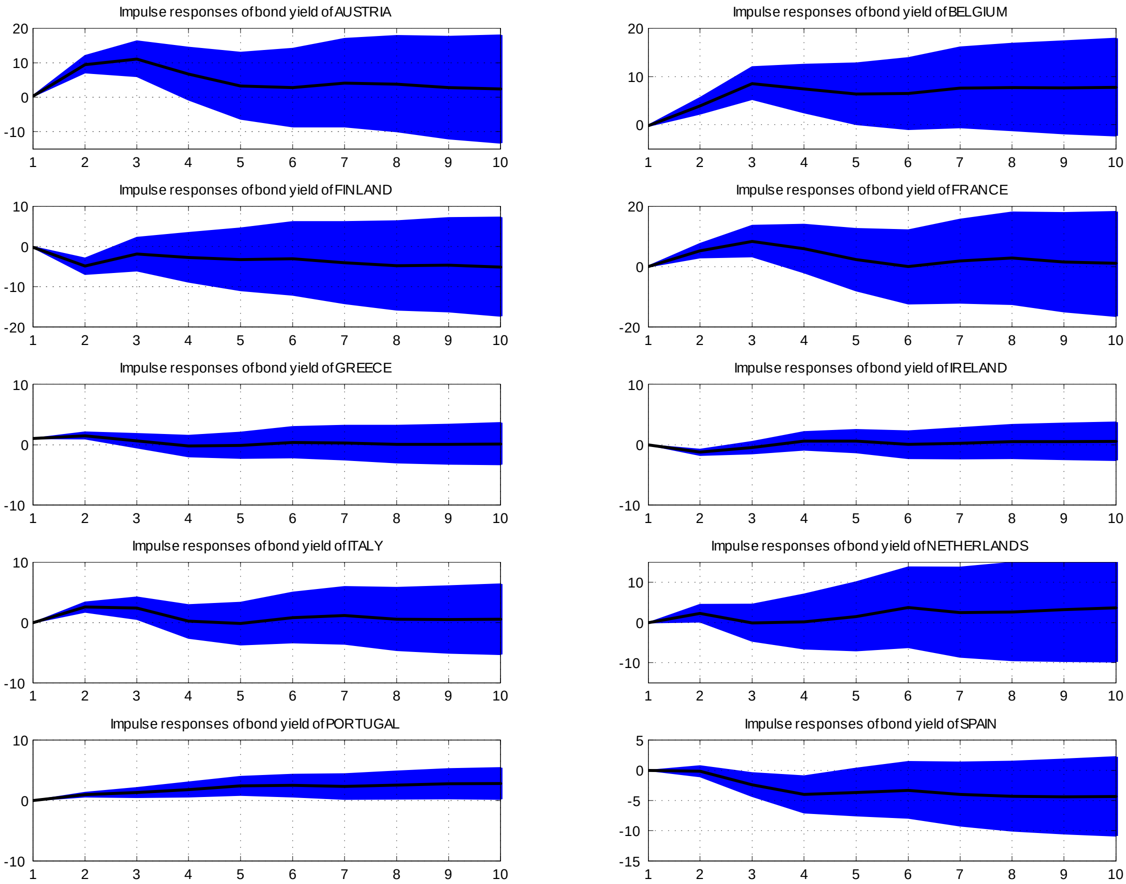
Table 3: Countries Where Restrictions Do Not Hold

	Dynamic Interdependencies			Cross-Sectional homogeneities			Static Interdependencies	
	To	From		First country	Second country		First country	Second country
1	BELGIUM	AUSTRIA	1	AUSTRIA	FRANCE	1	AUSTRIA	BELGIUM
2	BELGIUM	FINLAND	2	AUSTRIA	GREECE	2	AUSTRIA	FINLAND
3	FRANCE	AUSTRIA	3	AUSTRIA	IRELAND	3	AUSTRIA	FRANCE
4	FRANCE	BELGIUM	4	AUSTRIA	ITALY	4	AUSTRIA	GREECE
5	FRANCE	FINLAND	5	BELGIUM	FRANCE	5	AUSTRIA	IRELAND
6	GREECE	AUSTRIA	6	BELGIUM	GREECE	6	AUSTRIA	ITALY
7	GREECE	BELGIUM	7	BELGIUM	IRELAND	7	AUSTRIA	NETHERLANDS
8	GREECE	FINLAND	8	BELGIUM	ITALY	8	AUSTRIA	PORTUGAL
9	GREECE	FRANCE	9	FINLAND	FRANCE	9	AUSTRIA	SPAIN
10	IRELAND	AUSTRIA	10	FINLAND	GREECE	10	BELGIUM	FINLAND
11	IRELAND	BELGIUM	11	FINLAND	IRELAND	11	BELGIUM	FRANCE
12	IRELAND	FINLAND	12	FINLAND	ITALY	12	BELGIUM	GREECE
13	IRELAND	FRANCE	13	FRANCE	NETHERLANDS	13	BELGIUM	IRELAND
14	ITALY	AUSTRIA	14	FRANCE	PORTUGAL	14	BELGIUM	ITALY
15	ITALY	BELGIUM	15	FRANCE	SPAIN	15	BELGIUM	NETHERLANDS
16	ITALY	FINLAND	16	GREECE	NETHERLANDS	16	BELGIUM	PORTUGAL
17	ITALY	FRANCE	17	GREECE	PORTUGAL	17	BELGIUM	SPAIN
18	NETHERLANDS	AUSTRIA	18	GREECE	SPAIN	18	FINLAND	FRANCE
19	NETHERLANDS	BELGIUM	19	IRELAND	NETHERLANDS	19	FINLAND	GREECE
20	NETHERLANDS	FINLAND	20	IRELAND	PORTUGAL	20	FINLAND	IRELAND
21	SPAIN	FRANCE	21	IRELAND	SPAIN	21	FINLAND	ITALY
			22	ITALY	NETHERLANDS	22	FINLAND	NETHERLANDS
			23	ITALY	PORTUGAL	23	FINLAND	PORTUGAL
			24	ITALY	SPAIN	24	FINLAND	SPAIN
						25	FRANCE	GREECE

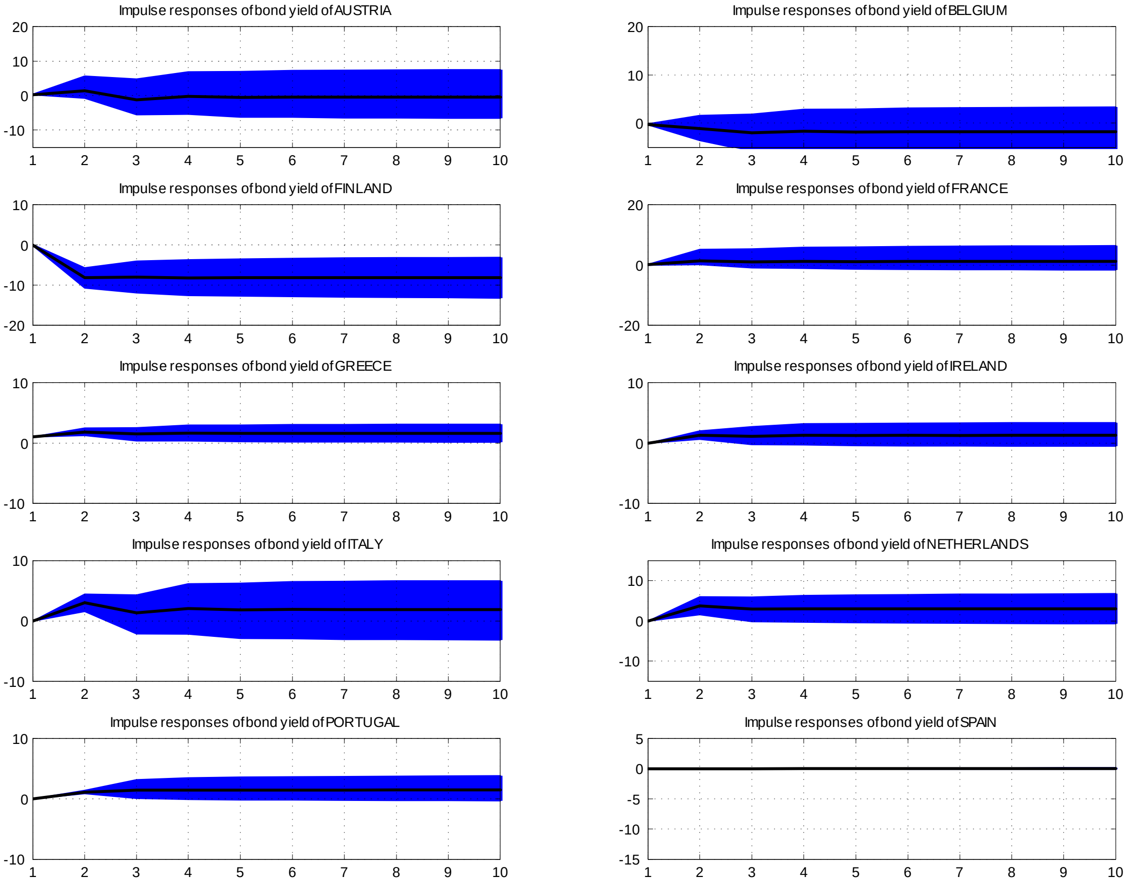
Finally we carry out an impulse response analysis to investigate spillovers of financial shocks across the euro area. For the sake of brevity, we focus on a single shock and ask what would happen to interest rate spreads around the euro area if the Greek 10-year bond rate increased unexpectedly by 1% relative to the German rate. Figures 1 and 2 plot these impulse responses for the unrestricted PVAR model, M8, and our panel S^4 model, M1, respectively. The black line in the figures is the posterior median of the impulse responses and the shaded region is the credible interval from the 16^{th} to 84^{th} percentile. To aid in comparability, we have used the same X-axis scale in the two figures. The results in Figure 2 can be interpreted as BMA results in the sense described at the end of Section 3.

It can be seen that the main impact of the use of S^4 methods is precision. The impulse responses coming from our S^4 approach are much more precisely estimated than those produced by an unrestricted, over-parameterized PVAR. This improvement in precision can lead to improved policy conclusions. For instance, the unrestricted VAR would suggest there is no effect in the Netherlands from a Greek shock since the bands cover zero completely. However, our panel S^4 approach predicts that there is a slight increase in the Dutch bond rate for several months.

The point estimates of the impulse responses in Figures 1 and 2 tend to be similar to one another. However, there are some differences. For instance, for Ireland the point estimates of the impulse responses from the unrestricted model, M8, are counter-intuitively negative at short horizons, whereas with our approach they are more reasonably positive. A similar thing happens for Spain. Thus, our approach is leading to impulse responses which are not only more precisely estimated, but also more sensible.



Responses to a shock to Greek bond yields from the unrestricted model, M8.



Responses to a shock to Greek bond yields from our model, M1.

6 Conclusions

In a globalized world, PVARs are an increasingly popular tool for estimating cross-country spillovers and linkages. However, unrestricted PVARs are often over-parameterized and the number of potential restricted PVAR models of interest can be huge. In this paper, we have developed methods for dealing with the huge model space that results so as to do BMA or BMS. These methods involve using a hierarchical prior that takes the panel nature of the problem into account and leads to an algorithm which we call S^4.

Our empirical work shows that our methods work well at picking out restrictions and selecting a tightly parameterized PVAR. Our findings are at odds with simple stories which divide the euro zone into a group of core countries and one of peripheral countries and speak of financial contagion within the latter. Instead we are finding a more nuanced story where there are groups of homogeneous countries, but they do not match perfectly with the standard grouping. Furthermore, we do not find evidence of interdependencies within the peripheral countries such as the financial contagion story would suggest.

References

[1] Arghyrou, M., Kontonikas, A., 2012. The EMU sovereign-debt crisis: Fundamentals, expectations and contagion. Journal of International Financial Markets, Institutions and Money 25, 658-677.

[2] Bai, J., Julliard, C., Yuan, K., 2012. Eurozone sovereign bond crisis: Liquidity or fundamental contagion. Manuscript available at http://www.greta.it/credit/credit2012/PAPERS/Speakers/Thursday/08_Bai_Julliard_Yuan.pdf

[3] Banbura, M., Giannone, D., Reichlin, L., 2010. Large Bayesian vector autoregressions. Journal of Applied Econometrics 25, 71-92.

[4] Canova, F., Ciccarelli, M., 2009. Estimating multicountry VAR models. International Economic Review 50, 929-959.

[5] Canova, F., Ciccarelli, M., 2013. Panel Vector Autoregressive models: A survey. European Central Bank Working Paper 1507.

[6] De Santis, R., 2012. The Euro area sovereign debt crisis: Safe haven, credit rating agencies and the spread of fever from Greece, Ireland and Portugal. European Central Bank Working Paper 1419.

[7] Carriero, A., Clark, T., Marcellino, M., 2011. Bayesian VARs: Specification choices and forecast accuracy. Federal Reserve Bank of Cleveland, Working Paper 1112.

[8] Carriero, A., Kapetanios, G., Marcellino, M., 2009. Forecasting exchange rates with a large Bayesian VAR. International Journal of Forecasting 25, 400-417.

[9] Dees, S., Di Mauro, F., Pesaran, M.H., Smith, V., 2007. Exploring the international linkages of the Euro area: A global VAR analysis. Journal of Applied Econometrics 22, 1-38.

[10] Eicher, T., Papageorgiou, C., Raftery, A., 2010. Determining growth determinants: Default priors and predictive performance in Bayesian model averaging. Journal of Applied Econometrics 26, 30-55.

[11] Fernández, C., Ley, E., Steel, M., 2001a. Benchmark priors for Bayesian model averaging. Journal of Econometrics 100, 381-427.

[12] Fernández, C., Ley, E., Steel, M., 2001b. Model uncertainty in cross-country growth regressions. Journal of Applied Econometrics 16, 563-576.

[13] Gefang, D., 2013. Bayesian doubly adaptive elastic-net Lasso for VAR shrinkage. International Journal of Forecasting, forthcoming.

[14] George, E., McCulloch, R., 1993. Variable selection via Gibbs sampling. Journal of the American Statistical Association 88, 881-889.

[15] George, E., McCulloch, R., 1997. Approaches for Bayesian variable selection. Statistica Sinica 7, 339-373.

[16] George, E., Sun, D., Ni, S., 2008. Bayesian stochastic search for VAR model restrictions. Journal of Econometrics 142, 553-580.

[17] Giannone, D., Lenza, M., Momferatou, D., Onorante, L., 2010. Short-term inflation projections: a Bayesian vector autoregressive approach. ECARES working paper 2010-011, Universite Libre de Bruxelles.

[18] Koop, G., 2013. Forecasting with medium and large Bayesian VARs. Journal of Applied Econometrics 28, 177-203.

[19] Korobilis, D., 2013. VAR forecasting using Bayesian variable selection. Journal of Applied Econometrics 28, 204-230.

[20] Ley, E., Steel, M., 2012. Mixtures of g-priors for Bayesian model averaging with economic applications, Journal of Econometrics 171, 251-266.

[21] Litterman, R., 1986. Forecasting with Bayesian vector autoregressions – Five years of experience. Journal of Business and Economic Statistics 4, 25-38.

[22] Molodtsova, T., Papell, D., 2009. Out-of-sample exchange rate predictability with Taylor rule fundamentals. Journal of International Economics 77, 167-180.

[23] Neri, S., Ropele, T., 2013. The macroeconomic effects of the sovereign debt crisis in the euro area. Pages 267-314 in The Sovereign Debt Crisis and the Euro Area, Workshops and Conferences, Bank of Italy.

[24] Raftery, A., Madigan, D., Hoeting, J., 1997. Bayesian model averaging for linear regression models. Journal of the American Statistical Association 92, 179-191.

[25] Smith M., Kohn R. 2002. Parsimonious covariance matrix estimation for longitudinal data. Journal of the American Statistical Association 97: 1141–1153.

[26] Spiegelhalter, D., Best, N., Carlin, B. and van der Linde, A., 2002. Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society Series B 64, 583-639.

Technical Appendix

We write this technical appendix for $P = 1$ (the value used in our empirical work) for notational simplicity. Formulae easily generalize for longer lag lengths. In this case, we can simplify our PVAR notation of (1) and (2). The VAR for country i, $i = 1, ..., N$ is of the form

$$\begin{aligned} y_{it} &= A_i Y_{t-1} + \varepsilon_{it} \\ &= A_{i1} y_{1t-1} + ... + A_{ii} y_{1t-1} + ... + A_{iN} y_{Nt-1} + \varepsilon_{it}, \end{aligned} \tag{A.1}$$

where $E\left(\varepsilon_{it}\varepsilon_{it}'\right) = \Sigma_{ii}$ and $E\left(\varepsilon_{it}\varepsilon_{jt}'\right) = \Sigma_{ij}$, $i \neq j$, $i, j = 1, ..., N$ and Σ is the full error covariance matrix for the entire PVAR. For future reference, we also define the upper triangular matrix Ψ through the equation $\Sigma = \Psi^{-1'}\Psi^{-1}$ which is partitioned into $G \times G$ blocks Ψ_{ii} and Ψ_{ij} conformably with Σ_{ii} and Σ_{ij}, respectively. In addition, we denote the elements of the diagonal blocks of Ψ_{ii} as ψ_{jk}^{ii}. George, Sun and Ni (2008) also parameterize their model in terms of Ψ. Smith and Kohn (2002) provide a justification and derivation of results for the prior we use for Ψ.

6.1 Stochastic Search Specification Selection (S^4): Hierarchical Prior

The DI, SI and CSH restrictions are given in Table 1. They are imposed through the vectors of dummy variables γ_{ij}^{DI}, γ_{ij}^{DI} and γ_{ij}^{CSH} described in Section 3. Our S^4 algorithm is based on a hierarchical prior which allows for their imposition. This is done through the following priors:[6]

1. DI prior:

$$vec\left(A_{ij}\right) \sim \left(1 - \gamma_{ij}^{DI}\right) N\left(0, \underline{\tau}_1^2 \times I\right) + \gamma_k^{DI} N\left(0, \underline{\tau}_2^2 \times I\right), \tag{A.2}$$

   where $\underline{\tau}_1^2$ is small and $\underline{\tau}_2^2$ large so that, if $\gamma_{ij}^{DI} = 0$, A_{ij} is shrunk to be near zero and, and if $\gamma_{ij}^{DI} = 1$, a relatively noninformative prior is used. The specification selection indicator for this DI restriction has prior

$$\gamma_{ij}^{DI} \sim Bernoulli\left(\underline{\pi}^{DI}\right). \tag{A.3}$$

2. CSH prior:

$$vec\left(A_{ii}\right) \sim \left(1 - \gamma_{ij}^{CSH}\right) N\left(A_{jj}, \underline{\xi}_1^2 \times I\right) + \gamma_{ij}^{CSH} N\left(A_{jj}, \underline{\xi}_2^2 \times I\right), \ \forall\, j \neq i, \tag{A.4}$$

[6] In our empirical work, we also include a vector of intercepts in the PVAR. For these, we use a noninformative prior which is a Normal prior with a very large variance.

where $\underline{\xi}_1^2$ is small and $\underline{\xi}_2^2$ is large so that, if $\gamma_{ij}^{CSH} = 0$, A_{ii} is shrunk to be near A_{jj}, and if $\gamma_{ij}^{CSH} = 1$, a relatively noninformative prior is used. The specification selection indicator for this CSH restriction has prior:

$$\gamma_{ij}^{CSH} \sim Bernoulli\left(\underline{\pi}^{CSH}\right), \tag{A.5}$$

for $i = 1, ..., N$, $j = i, ..., N-1$.

3. SI prior:

$$vec\left(\Psi_{ij}\right) \sim \left(1 - \gamma_{ij}^{SI}\right) N\left(0, \underline{\kappa}_1^2 \times I\right) + \gamma_{ij}^{SI} N\left(0, \underline{\kappa}_2^2 \times I\right), \tag{A.6}$$

where $\underline{\kappa}_1^2$ is small and $\underline{\kappa}_2^2$ large so that, if $\gamma_{ij}^{SI} = 0$, Ψ_{ij} (and, thus, Σ_{ij}) is shrunk to be near zero, and if $\gamma_{ij}^{SI} = 1$, a relatively noninformative prior is used. The specification selection indicator for this SI restriction has prior:

$$\gamma_{ij}^{SI} \sim Bernoulli\left(\underline{\pi}^{SI}\right). \tag{A.7}$$

This completes description of the hierarchical prior we use relating to the restrictions. We also require a prior for the error variances, which are not subject to any restrictions. We do this through the following prior:

$$\psi_{kl}^{ii} \sim \begin{cases} N\left(0, \underline{\kappa}_2^2\right), & \text{if } k \neq l \\ G\left(\underline{\rho}_1, \underline{\rho}_2\right) & \text{if } k = l \end{cases}, \tag{A.8}$$

where $G(.,.)$ denotes the Gamma distribution.

These priors depend on prior hyperparameters $\left(\underline{\tau}_1^2, \underline{\tau}_2^2\right)$, $\left(\underline{\xi}_1^2, \underline{\xi}_2^2\right)$, $\left(\underline{\kappa}_1^2, \underline{\kappa}_2^2\right)$, $\left(\underline{\rho}_1, \underline{\rho}_2\right)$ and $\left(\underline{\pi}^{DI}, \underline{\pi}^{SI}, \underline{\pi}^{CSH}\right)$. We set $\underline{\tau}_1^2 = \underline{\xi}_1^2 = \underline{\kappa}_1^2 = 0.01$, thus ensuring tight shrinkage towards the restrictions. For the other hyperparameters we use relatively noninformative choices. We set $\underline{\tau}_2^2 = \underline{\xi}_2^2 = \underline{\kappa}_2^2 = 10$, $\underline{\rho}_1 = \underline{\rho}_2 = 0.01$ and $\underline{\pi}^{DI} = \underline{\pi}^{SI} = \underline{\pi}^{CSH} = \frac{1}{2}$. The last of these implies that, a priori, each restriction is equally likely to hold as not.

6.2 Stochastic Search Specification Selection (S^4): MCMC Algorithm

Our MCMC algorithm requires minor alterations to that given in George, Sun and Ni (2008). In essence, George, Sun and Ni (2008)'s SSVS prior is (conditional on variable selection indicators) a Normal prior which combines with a Normal likelihood in a standard way. Our S^4 prior is also a Normal prior (albeit of a different form than the SSVS prior) which will also combine with a Normal likelihood in a standard way. Hence, we do not write out the formulae in detail but give a heuristic summary of our MCMC algorithm. The reader can find precise details in our MATLAB code available through the website: https://sites.google.com/site/dimitriskorobilis/matlab.

The MCMC algorithm involves the following steps:

1. Sample α from a Normal posterior conditional on Σ, γ^{DI} and γ^{CSH}. In order to impose the CSH restriction we need to create the matrix Γ defined in the main text.[7] For imposing the DI restrictions based on the values of the vector γ^{DI}, use the procedure of George, Sun and Ni (2008).

2. Sample each γ_{ij}^{DI} and γ_{ij}^{CSH} from its Bernoulli posterior conditional on α and Σ. The Bernoulli probability is based on the prior and the value of the likelihood function when γ_{ij}^{DI} $\left(\text{or } \gamma_{ij}^{CSH}\right) = 1$ and when γ_{ij}^{DI} $\left(\text{or } \gamma_{ij}^{CSH}\right) = 0$.

3. Sample ψ_{kl}^{ii} from its Normal conditional posterior (conditional on all other model parameters) if $k \neq l$, and from its Gamma posterior (conditional on all other model parameters) if $k = l$.

4. Sample $vec\left(\Psi_{ij}\right)$ from its Normal conditional posterior distribution (conditional on other parameters) as in George, Sun and Ni (2008).

5. Sample each γ_{ij}^{SI} from from its Bernoulli posterior conditional on α and Σ. The Bernoulli probability is based on the prior and the value of the likelihood function when $\gamma_{ij}^{SI} = 1$ and when $\gamma_{ij}^{SI} = 0$.

6. Calculate Σ using $\Sigma = \Psi^{-1'}\Psi^{-1}$ and go to step 1.

Our empirical results using the euro area data set are produced using 1,100,000 MCMC draws for each model. An initial 100,000 draws are discarded and, from the remaining 1,000,000, every 100th draw is retained. Standard MCMC convergence diagnostics indicate convergence has been achieved.

[7] This is an approximation in the sense that our CSH prior only approximately imposes restrictions whereas by drawing α conditional on Γ we are exactly imposing CSH restrictions.