Model Uncertainty in Panel Vector Autoregressive Models

Gary Koop
University of Strathclyde

Dimitris Korobilis
University of Glasgow

September 28, 2015

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

1 Introduction

This paper develops Bayesian methods for estimation and model selection with large 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.

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. Nevertheless, we focus on panel restrictions as defined in e.g. 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 two countries' VARs are non-

[1] Canova and Ciccarelli (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.

zero. CSHs occur when two countries have VARs with different coefficients - in other words, homogeneity (absense of heterogeneity) arises when the coefficients on the own lagged variables for the two countries are exactly the same.

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 meaning that the total number of restriction to examine is $2^{90+45+45}$. 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 hierarchical priors and Markov Chain Monte Carlo (MCMC) methods so as to avoid the huge computational burden of exhaustively estimating every restricted model. These 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; see, among many others, Fernández, Ley and Steel (2001a,b), Eicher, Papageorgiou and Raftery (2010), and Ley and Steel (2012).

With univariate linear 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)[2] methods have proved popular. Early papers include Cripps, Carter and Kohn (2005) and George, 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

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

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 a group of countries which are, in a sense we describe below, homogeneous. But the grouping 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[3] for country i can be written as:

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

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 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

[3]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.

($N = 7$) 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 shrinks towards zero irrelevant coefficients. Papers such as Carriero, Clark and Marcellino (2015), 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 of 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, define

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

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.

CSHs occur if the VAR coefficients differ across countries.[4] 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.

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 CSHs 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.

[4]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.

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 $\varepsilon_t \sim N(0, \Sigma)$ for $t = 1, .., T$ (uncorrelated over time), α is a vector containing the $K = P(NG)^2$ VAR coefficients and Z_t is an appropriate $NG \times K$ matrix such that the VAR for each country contains P lags of the variables for all countries. That is,

$$Z_t = \begin{pmatrix} z_t & 0 & \cdots & 0 \\ 0 & z_t & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & z_t \end{pmatrix},$$

where $z_t = \left(Y'_{t-1}, Y'_{t-2}, ..., Y'_{t-P}\right)$. 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 jth 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:

$$\alpha_j | \gamma_j \sim \left(1 - \gamma_j\right) N\left(0, \underline{c} \times \tau_j^2\right) + \gamma_j N\left(0, \tau_j^2\right), \tag{4}$$

where $\gamma_j \in \{0, 1\}$ is an unknown parameter estimated from the data. In standard SSVS implementations, the researcher chooses specific values for $\underline{c}$ and τ_j^2 such that the first element in the Normal mixture has a prior variance near zero and the second element has a larger prior 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). Thus, if $\gamma_j = 0$, α_j is shrunk to zero whereas if $\gamma_j = 1$ it is not. In Bayesian estimation of the model, it is conventional to use a Bernoulli prior for γ_j (e.g. $\Pr\left(\gamma_j\right) = \pi_j$ with $\pi_j = \frac{1}{2}$ expresses a view that the jth coefficient is, a priori, equally likely to be excluded or included). In this paper, we extend this standard approach by using hierarchical priors for π_j and τ_j^2. In particular, we assume these are unknown parameters with Beta and Gamma priors, respectively. In this way, we lessen concerns about prior sensitivity. Complete details, including choices of $\underline{c}$ and the prior hyperparameters in the Beta and Gamma priors are given in the Technical Appendix.

The basic ideas underlying our S^4 algorithm can be expressed in terms of (4). In this section, we outline how we do this, partly relying on a three country example. 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. Similarly, while equation (4) pertains to a scalar α_j, our algorithms tries to set to zero all G^2 elements that pertain to DIs between countries j and k.

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. Consider the case of three countries ($N = 3$) with one variable each ($G = 1$), then the vector of coefficients α is simply $\alpha = (\alpha_1, \alpha_2, \alpha_3)'$. Applying the SSVS prior of equation (4) in this simple case gives

$$\alpha_j|\gamma_j \sim \left(1 - \gamma_{i,j}^{CSH}\right) N\left(\alpha_k, \underline{c} \times \tau_j^2\right) + \gamma_{i,j}^{CSH} N\left(0, \tau_j^2\right).$$

It is immediately evident that it is hard to handle all possible combinations of coefficients α_j being equal to coefficients α_k (note also that for the general case with $G > 1$, these will be vectors and not scalars). In order to deal with this issue, we introduce 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, in our simple example with $N = 3$ and $G = 1$ we have 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 $2^{N(N-1)/2}$ possible combinations of CSH restrictions if we write the posterior of α as

$$\alpha|- \sim N\left(\prod_{j=1}^{N-1}\prod_{k=j+1}^{N} \Gamma_{j,k}\mu_\alpha, D_\alpha\right),$$

where μ_α and D_α are the posterior mean and variance of α , and the $-$ shows that α is conditional on some quantities such as data and other parameters (exact formulas are available in the Technical Appendix). This formula imposes that if $\gamma_{jk}^{CSH} = 0$, then $\mu_{\alpha,j} = \mu_{\alpha,k}$ and at the same time $D_{\alpha,jj}$ is very small, thus, shrinking the whole posterior of α_j towards the posterior of α_k. The simple formula above allows for the application all possible CSH restrictions using fast sparse matrix multiplications. The alternative would be to index all possible pairs (α_j, α_k) and check whether their associated index γ_{jk}^{CSH} is zero or one, which is computationally infeasible for large N.

Once the PVAR is transformed in this way, sampling from the conditional posterior of the restricted coefficients 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:

$$A_1^{true} = \begin{bmatrix} 0.7 & 0 & 0.2 & 0.2 & 0 & 0 \\ 0 & 0.7 & 0.3 & 0.3 & 0 & 0 \\ 0 & 0 & 0.6 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0 & 0 \\ 0.3 & -0.4 & 0 & 0 & 0.6 & 0.5 \\ 0.2 & 0.4 & 0 & 0 & 0 & 0.5 \end{bmatrix}, \Sigma^{true} = \begin{bmatrix} 1 & 0 & -0.5 & -0.5 & 0 & 0 \\ 0 & 1 & -0.5 & -0.5 & 0 & 0 \\ -0.5 & -0.5 & 1 & 0.5 & 0 & 0 \\ -0.5 & -0.5 & 0.5 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

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.

$$A_1^{S^4} = \begin{bmatrix} .64 & .03 & .25 & .28 & -.02 & .01 \\ .01 & .65 & .37 & .33 & .01 & .00 \\ .00 & .00 & .59 & .51 & .01 & .00 \\ .00 & -.01 & .05 & .51 & .02 & -.01 \\ .34 & -.42 & .00 & .00 & .61 & .50 \\ .23 & .41 & .00 & .01 & -.02 & .46 \end{bmatrix}, \Sigma^{S^4} = \begin{bmatrix} 1.01 & -.02 & -.51 & -.51 & .00 & .00 \\ -.02 & .99 & -.42 & -.44 & .00 & .00 \\ -.51 & -.42 & 1.22 & .55 & -.01 & .00 \\ -.51 & -.44 & .55 & 1.21 & .01 & .01 \\ .00 & .00 & -.01 & .01 & .95 & .00 \\ .00 & .00 & .00 & -.01 & -.00 & 1.02 \end{bmatrix}.$$

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:

$$A_1^{OLS} = \begin{bmatrix} .59 & -.02 & .22 & .19 & -.07 & -.06 \\ .00 & .60 & .39 & .27 & .00 & .08 \\ .06 & .09 & .50 & .51 & .06 & -.05 \\ .08 & .04 & -.09 & .52 & .10 & -.09 \\ .34 & -.44 & -.01 & .00 & .61 & .50 \\ .23 & .42 & .02 & .02 & -.02 & .46 \end{bmatrix}, \Sigma^{OLS} = \begin{bmatrix} 1.02 & .04 & -.59 & -0.53 & .02 & -.07 \\ .04 & .96 & -.44 & -0.46 & .01 & .00 \\ -.59 & -.44 & 1.42 & .60 & .02 & .02 \\ -.53 & -.46 & .60 & 1.30 & .01 & .05 \\ .02 & .01 & -.02 & .01 & .88 & -.01 \\ -.07 & .00 & .02 & .05 & .01 & 1.08 \end{bmatrix}.$$

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. In order to measure how good the performance of each estimator is, we are using the Absolute Percentage Deviation (APD) statistic

$$APD = \frac{1}{K} \sum_{i=1}^{K} \left(\widetilde{\alpha}_i - \alpha_i^{true} \right),$$

where $\widetilde{\alpha}$ is the vectorized form of the matrix A_1 for estimated coefficients, and α^{true} is the vectorized form of the true coefficients A_1^{true}. Under this statistic $APD^{S^4} = 0.04842 < 0.11534 = APD^{OLS}$, thus, on average, the coefficients from the S^4 algorithm are closer to the true coefficients compared to OLS. The same qualitative result is obtained if we compare Σ^{OLS} and Σ^{S^4}.

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 $p\left(A_{1,22} = A_{1,33}\right)$ 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:

$$\begin{array}{c} p\left(A_{1,12} = 0\right) = 0 \\ p\left(A_{1,13} = 0\right) = 1 \\ p\left(A_{1,21} = 0\right) = 1 \\ p\left(A_{1,23} = 0\right) = 1 \\ p\left(A_{1,31} = 0\right) = 0 \\ p\left(A_{1,32} = 0\right) = 1 \end{array}.$$

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:

$$\begin{array}{c} p\left(\Sigma_{12}=0\right)=.238 \\ p\left(\Sigma_{13}=0\right)=.983 \\ p\left(\Sigma_{23}=0\right)=.920 \end{array}.$$

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:

$$\begin{array}{c} p\left(A_{1,11}=A_{1,22}\right)=.072 \\ p\left(A_{1,11}=A_{1,33}\right)=.055 \\ p\left(A_{1,22}=A_{1,33}\right)=.543 \end{array}.$$

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 log of the marginal likelihood (ML) and Deviance information criterion (DIC) for each model. 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 in models involving latent variables such as ours. Note that higher (lower) values of ML (DIC) are associated with better model performance.

Table 2: Model fit (numerical standard errors in parentheses)

Method	M1	M2	M3	M4	M5	M6	M7	M8
ML	-45.39	-45.06	-47.18	-52.04	-50.30	-48.59	-48.99	-51.95
	(0.09)	(0.05)	(0.03)	(0.06)	(0.03)	(0.06)	(0.04)	(0.03)
DIC	-56.12	-53.40	-53.46	-56.82	-50.61	-52.82	-54.14	-50.73
	(0.09)	(0.07)	(0.03)	(0.11)	(0.05)	(0.00)	(0.00)	(0.01)

A key message from Table 2 is that our full approach (M1) does well and the large VAR approach (M8) does poorly, indicating that our S^4 prior which takes into account the panel structure of the model can lead to substantial improvements. Results relating to which types of restrictions are most important are a bit less clear. The DIC results suggest 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, but the marginal likelihood suggests more of a role for the DI restrictions. The ability to impose static interdependencies is of less benefit in this data set since M5 (which only allows for their imposition) does relatively poorly using either model comparison metric.

What kind of restrictions does our preferred M1 model find? Tables 3a, 3b and 3c address this question for DI, CSH and SI restrictions, respectively. 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.[5] For each part of Table 3 we 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 3a and 3c will clearly show where such linkages exist. Country pairs not listed in these tables are found to be not interlinked.

Consider first the cross-sectional homogeneities. This is the category of restrictions which is most often rejected. 36 of the 45 possible restrictions are not imposed. By examining which countries are not listed in Table 3b, it can be seen that there are several countries with VARs which are sufficiently heterogeneous so as to reject most or all CSH restrictions with all other countries. That is, Belgium, Finland and France have no homogeneities with any other countries and Greece and Italy have homogeneities with only one other country. The remaining five countries (Austria, Ireland, Netherlands, Portugal and Spain), with some exceptions, do tend to have homogeneous VARs. This latter group of countries does contain some of the periphery countries, but also some of the core countries. So we are not finding a conventional core versus periphery division.

[5]The estimated probabilities for each restriction tend to be quite definitive (i.e. near either zero or one), such that changing this threshold fairly substantially either up or down has little impact on our findings.

We stress that our definition of cross-sectional homogeneities only involves own country variables and not linkages between countries. For instance, a finding that Italy and Austria are homogeneous means that a VAR containing only Italian variables and a VAR containing only Austrian variables have very similar estimated coefficients. Such a finding would say nothing about how other-country variables impact on Italy or Austria. Nevertheless, it is striking that we are finding such homogeneity in some cases, but that the resulting grouping does not coincide with the conventional core versus periphery division.

Table 3c shows that many static interdependencies exist. The main pattern here is that France plus the small countries of Austria, Belgium and Finland have SIs with every other country. 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 none of the peripheral countries exhibits SIs with any country other than Austria, Belgium, Finland and France.

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 3a) 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 Spain, but not vice versa. The main pattern is that the peripheral countries lagged dependent variables rarely appear in any of the core countries' VARs. That is, there are many DIs in Table 3a, but it is only rarely 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 the "From" columns of Table 3a at all. 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 3a. Countries where Dynamic Interdependency Restrictions do not hold

No	To	From	No	To	From	No	To	From
1	AUT	FIN	19	IRL	AUT	40	PRT	AUT
2	BEL	AUT	20	IRL	BEL	41	PRT	BEL
3	BEL	FIN	21	IRL	FIN	42	PRT	FIN
4	BEL	FRA	22	IRL	FRA	43	PRT	FRA
5	FIN	AUT	23	IRL	GRC	44	PRT	GRC
6	FIN	BEL	24	IRL	ITA	45	PRT	IRL
7	FIN	FRA	25	IRL	NLD	46	PRT	ITA
8	FRA	AUT	26	ITA	AUT	47	PRT	NLD
9	FRA	BEL	27	ITA	BEL	48	PRT	ESP
10	FRA	FIN	28	ITA	FIN	49	ESP	AUT
11	FRA	GRC	29	ITA	FRA	50	ESP	BEL
12	FRA	IRL	30	ITA	GRC	51	ESP	FIN
13	GRC	AUT	31	ITA	IRL	52	ESP	FRA
14	GRC	BEL	32	NLD	AUT	53	ESP	GRC
15	GRC	FIN	33	NLD	BEL	54	ESP	IRL
16	GRC	FRA	34	NLD	FIN	55	ESP	ITA
17	GRC	IRL	35	NLD	FRA	56	ESP	NLD
18	GRC	ITA	36	NLD	GRC	57	ESP	PRT
			37	NLD	IRL			
			38	NLD	ITA			
			39	NLD	ESP			

Table 3b. Countries where Cross-Sectional Homogeneity Restrictions do not hold

No	C1	C2	No	C1	C2	No	C1	C2	No	C1	C2
1	AUT	BEL	10	BEL	IRE	19	FIN	NLD	28	GRC	NLD
2	AUT	FIN	11	BEL	ITA	20	FIN	PRT	29	GRC	PRT
3	AUT	FRA	12	BEL	NLD	21	FIN	ESP	30	GRC	ESP
4	AUT	GRC	13	BEL	PRT	22	FRA	GRC	31	IRE	ITA
5	AUT	IRE	14	BEL	ESP	23	FRA	IRE	32	IRE	NLD
6	AUT	ESP	15	FIN	FRA	24	FRA	ITA	33	IRE	PRT
7	BEL	FIN	16	FIN	GRC	25	FRA	NLD	34	ITA	NLD
8	BEL	FRA	17	FIN	IRE	26	FRA	PRT	35	ITA	PRT
9	BEL	GRC	18	FIN	ITA	27	FRA	ESP	36	ITA	ESP

Table 3c. Countries where Static Interdependency Restrictions do not hold

No	C1	C2	No	C1	C2	No	C1	C2
1	AUT	BEL	11	BEL	FRA	21	FIN	ITA
2	AUT	FIN	12	BEL	GRC	22	FIN	NLD
3	AUT	FRA	13	BEL	IRE	23	FIN	PRT
4	AUT	GRC	14	BEL	ITA	24	FIN	ESP
5	AUT	IRE	15	BEL	NLD	25	FRA	GRC
6	AUT	ITA	16	BEL	PRT	26	FRA	IRE
7	AUT	NLD	17	BEL	ESP	27	FRA	ITA
8	AUT	PRT	18	FIN	FRA	28	FRA	NLD
9	AUT	ESP	19	FIN	GRC	29	FRA	PRT
10	BEL	FIN	20	FIN	IRE	30	FRA	ESP

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.[6] The black line in the figures is the posterior median of the impulse responses and the shaded region is the credible interval from the 16th to 84th percentile. To aid in comparability, we have used the same Y-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 Spain from a Greek shock since the bands cover zero completely. However, our panel S^4 approach predicts that there is a slight impact on the Spanish bond rate in the medium term. This is consistent with the finding in Table 3a that dynamic interdependencies existed from Greece to Spain.

[6] A brief comparison of impulse responses produced by some other models is given in the Empirical Appendix.

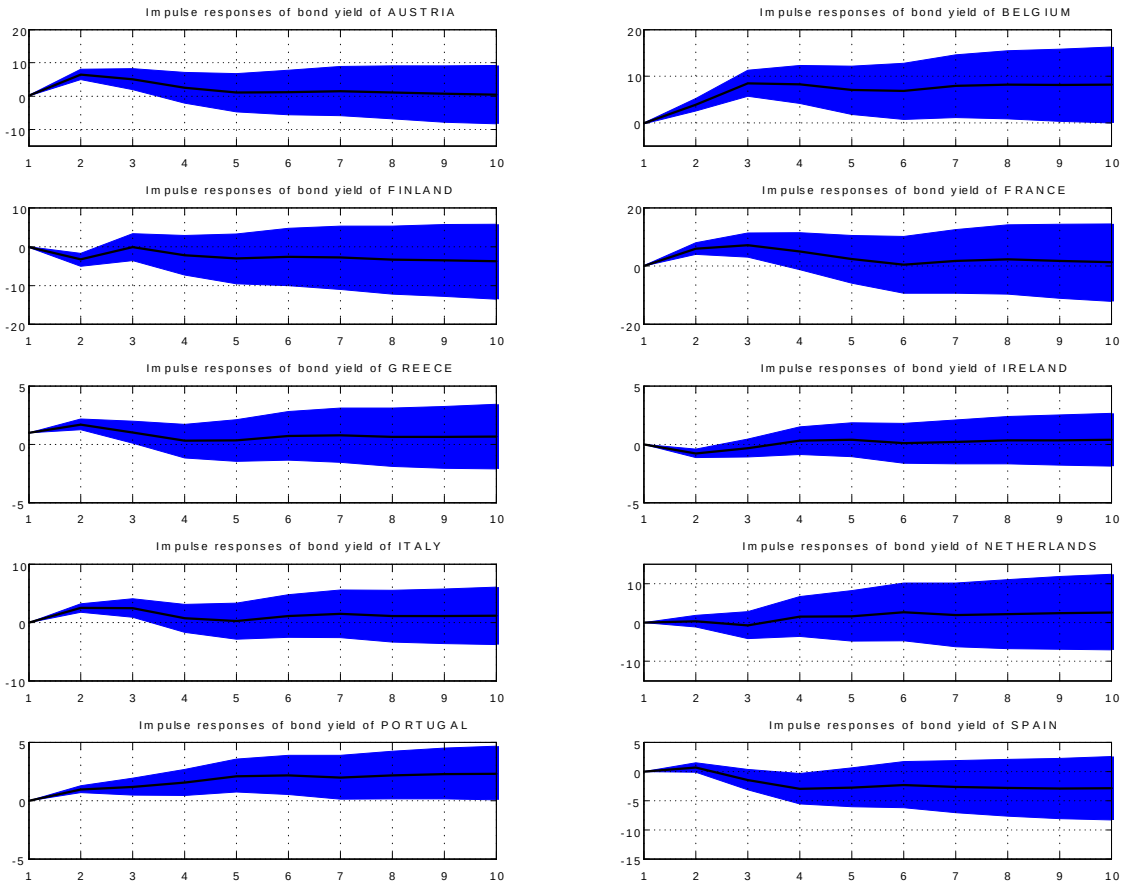

Figure 1: Responses to a shock to Greek bond yields from the unrestricted model, M8.

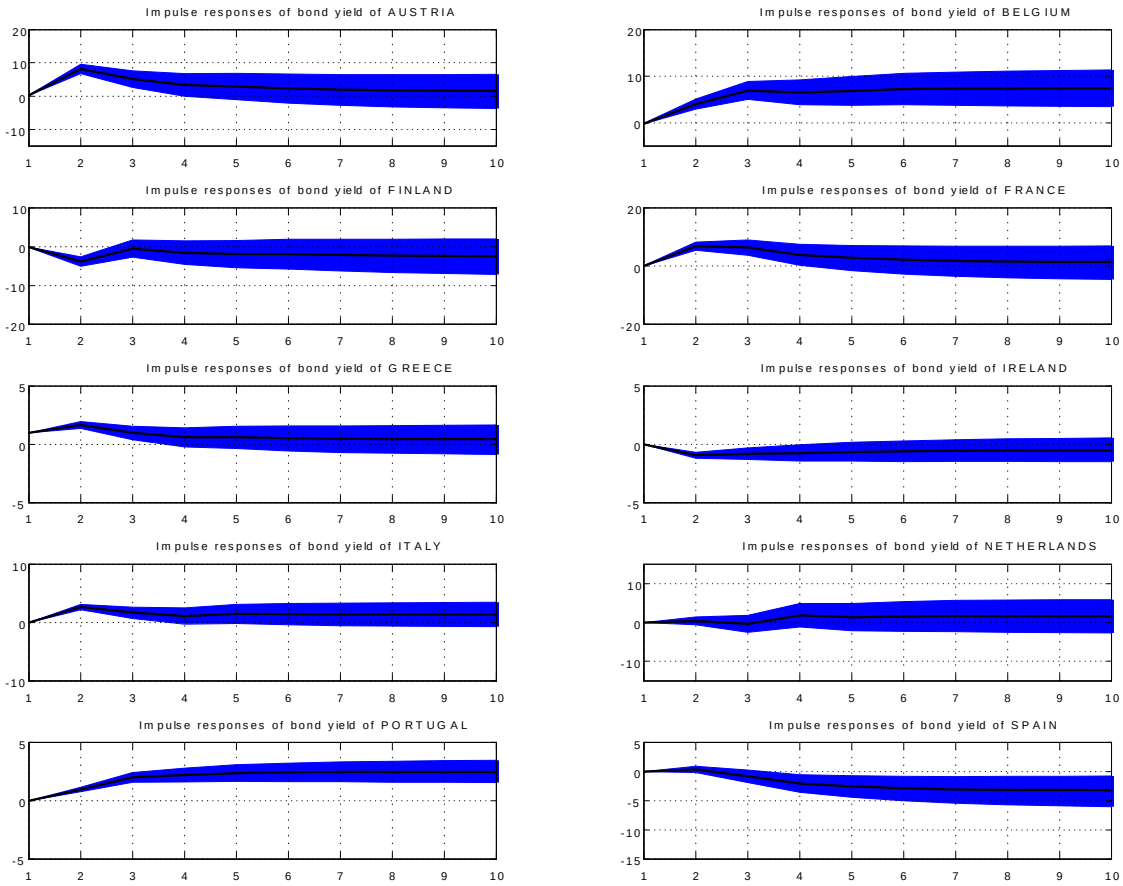

Figure 2: Responses to a shock to Greek bond yields from our model, M1.

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 is a group of homogeneous countries, but it does 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.p

[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] Carriero, A., Clark, T., Marcellino, M., 2015. Bayesian VARs: Specification choices and forecast accuracy. Journal of Applied Econometrics 30, 46-73.

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

[8] Cripps, E., Carter, C.K., Kohn, R., 2005. Variable selection and covariance selection in multivariate regression models. Handbook of Statistics, Volume 25: Bayesian Thinking: Modeling and Computation. Editors: D. Dey and C.R. Rao. Elsevier Science, Amsterdam: The Netherlands.

[9] 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.

[10] 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.

[11] 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.

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

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

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

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

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

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

[18] 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.

[19] Ishwaran, H., Rao, J. S., 2003. Detecting differentially expressed genes in microarrays using Bayesian model selection. Journal of the American Statistical Association 98, 438-455.

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

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

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

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

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

[25] 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.

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

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

[28] Spiegelhalter, D., Best, N., Carlin, B., 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, thus, 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.

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:[7]

1. DI prior:

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

$$\tau_{ij}^{-2} \sim Gamma\left(1, \underline{\theta}^{DI}\right) \tag{A.3}$$

where the specification selection indicator for this DI restriction has prior

$$\gamma_{ij}^{DI} \sim Bernoulli\left(\pi_{ij}^{DI}\right), \tag{A.4}$$

$$\pi_{ij}^{DI} \sim Beta\left(1, \underline{\varphi}\right), \tag{A.5}$$

for $i = 1, ..., N$, $j = i, ..., N - 1$, and $i \neq j$ (so that DI restrictions do not apply to matrices A_{ii}, A_{jj} etc).

2. CSH prior:

$$vec\left(A_{ii}\right) \sim \left(1 - \gamma_{ij}^{CSH}\right) N\left(A_{jj}, \left(\xi_{ij}^2 \times \underline{c}^{CSH}\right) \times I_{G^2}\right) + \gamma_i^{CSH} N\left(A_{jj}, \xi_{ij}^2 \times I_{G^2}\right), \ \forall\ j \neq i \tag{A.6}$$

$$\xi_{ij}^{-2} \sim Gamma\left(1, \underline{\theta}^{CSH}\right). \tag{A.7}$$

[7] 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.

The specification selection indicator for this CSH restriction has prior:

$$\gamma_{ij}^{CSH} \sim Bernoulli\left(\pi_{ij}^{CSH}\right), \tag{A.8}$$
$$\pi_{ij}^{CSH} \sim Beta\left(1, \underline{\varphi}\right), \tag{A.9}$$

for $i = 1, ..., N$, $j = i, ..., N-1$, and $i \neq j$ (so that A_{ii} and A_{jj} are not the same matrix).

3. SI prior:

$$vec\left(\Psi_{ij}\right) \sim \left(1-\gamma_{ij}^{SI}\right) N\left(0, \kappa_{ij}^2 \times \underline{c}^{SI} \times I_{G^2}\right) + \gamma_{ij}^{SI} N\left(0, \kappa_{ij}^2 \times I_{G^2}\right), \tag{A.10}$$
$$\kappa_{ij}^{-2} \sim Gamma\left(1, \underline{\theta}^{SI}\right) \tag{A.11}$$

where the specification selection indicator for this SI restriction has prior:

$$\gamma_{ij}^{SI} \sim Bernoulli\left(\pi_{ij}^{SI}\right), \tag{A.12}$$
$$\pi_{ij}^{SI} \sim Beta\left(1, \underline{\varphi}\right), \tag{A.13}$$

where $i = 1, ..., N$, $j = i, ..., N-1$, $i > j$.

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

$$\psi_{kl}^{ii} \sim \begin{cases} N\left(0, \underline{\kappa}_2^2\right), \text{ if } k \neq l \\ Gamma\left(\underline{\rho}_1, \underline{\rho}_2\right) \text{ if } k = l \end{cases}, \tag{A.14}$$

where $k, l = 1, ..., G$ index each of the G macro variables of country $i = 1, ..., N$.

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

The prior hyperparameters of the model are $\underline{c}^{DI}, \underline{c}^{CSH}, \underline{c}^{SI}, \underline{\theta}^{DI}, \underline{\theta}^{CSH}, \underline{\theta}^{SI}, \underline{\varphi}, \underline{\kappa}_2^2, \underline{\rho}_1$ and $\underline{\rho}_2$. We have used relatively vague priors where possible, and for other priors we have followed a full Bayes approach that allows to update priors from the data. Nevertheless, prior choices in very large models do matter, especially for structural results such as impulse responses (probably not so much for forecasting). Therefore, we used slightly different prior hyperparameters in the Monte Carlo and empirical exercises, considering always the PVAR size we had to work with. The final decision does indeed depend on our experience with such large VARs, although choices refer only to lower level priors and, thus, sensitivity is somewhat limited. For example, in the prior for the DI restrictions in eq (A.2) we only choose one hyperparameter, $\underline{c}^{DI}$, while all other hyperparameters have their own priors and are updated by the data. The following table summarizes the prior distributions and hyperparameter values used in the two exercises

	Monte Carlo	Empirical
$\underline{c}^{DI}$	1e-6	1e-6
$\underline{c}^{CSH}$	1e-4	1e-5
$\underline{c}^{SI}$	1e-5	1e-5
$\underline{\theta}^{DI}$	50	10
$\underline{\theta}^{CSH}$	60	60
$\underline{\theta}^{SI}$	25	10
$\underline{\varphi}$	1	1
$\underline{\kappa}_2^2$	4	4
$\underline{\rho}_1$	0.01	0.01
$\underline{\rho}_2$	0.01	0.01

Additionally, as explained in Section 3, we define a matrix $\Gamma = \prod_{i=1}^{N-1} \prod_{j=i+1}^{N} \Gamma_{i,j}$, where $\Gamma_{i,j}$ are $K \times K$ matrices constructed using the CSH restriction indicators γ_{ij}^{CSH}. First note that $\gamma_{ij}^{CSH} = 0$ implies that countries i and j have similar coefficients (i.e. the homogeneity restriction $A_{ii} \approx A_{jj}$ holds), and the opposite is true when $\gamma_{ij}^{CSH} = 1$. The matrix $\Gamma_{i,j}$ is the identity matrix with the exception that its $\{i,i\}$ diagonal element is equal to γ_{ij}^{CSH} and its $\{i,j\}$ non-diagonal element is equal to $\left(1 - \gamma_{ij}^{CSH}\right)$. Therefore, each of the possible $N(N-1)/2$ matrices $\Gamma_{i,j}$ allow us to impose on the PVAR coefficients the CSH restriction between countries i and j, and their product, which is the matrix $\Gamma = \prod_{i=1}^{N-1} \prod_{j=i+1}^{N} \Gamma_{i,j}$, allows us to index all $2^{N(N-1)/2}$ possible CSH restrictions among the N countries. Therefore, if μ_α denotes the posterior mean of the unrestricted vectorized PVAR coefficients (i.e. using a noninformative prior), then $\widetilde{\mu}_\alpha = \Gamma \mu_\alpha = \prod_{i=1}^{N-1} \prod_{j=i+1}^{N} \Gamma_{i,j} \mu_\alpha$ is simply the $K \times 1$ vector of posterior means of the PVAR coefficients with the cross-sectional homogeneity restrictions imposed.

Gibbs sampler algorithm for the S^4 algorithm

1. Sample $vec(A)$ from

$$(vec(A)|-) \sim N(\Gamma \times \mu_\alpha, D_\alpha), \tag{A.15}$$

where $D_\alpha = \left(\Sigma^{-1} \otimes X'X + (V'V)^{-1}\right)^{-1}$ and $\mu_\alpha = D_\alpha \left[(\Sigma^{-1} \otimes X'X)\, \alpha_{OLS}\right]$, where α_{OLS} is the OLS estimate of α, and V is a diagonal matrix which has its respective diagonal block of G^2 elements equal to $\tau_{ij}^2 \times \mathbf{1}$ if $\gamma_{ij}^{DI} = 1$ or equal to $\tau_{ij}^2 \times \underline{c}^{DI} \times \mathbf{1}$ if $\gamma_{ij}^{DI} = 0$, and equal to $\xi_{ij}^2 \times \mathbf{1}$ if $\gamma_{ij}^{CSH} = 1$ or equal to $\xi_{ij}^2 \times \underline{c}^{CSH} \times \mathbf{1}$ if $\gamma_{ij}^{CSH} = 1$, where $\mathbf{1}$ is a $G^2 \times 1$ vector of ones.

2. Sample τ_{ij}^2 from

$$\left(\tau_{ij}^2|-\right) \sim Gamma\left(1+\frac{1}{2}G, \underline{\theta}^{DI}+\frac{1}{2}\sum_{k=1}^{G}\frac{\left[vec\left(A_{ij}\right)_k\right]^2}{\left(\underline{c}^{DI}\right)^{1-\gamma_{ij}^{DI}}}\right)$$

3. Sample ξ_{ij}^2 from

$$\left(\xi_{ij}^2|-\right) \sim Gamma\left(1+\frac{1}{2}G, \underline{\theta}^{CSH}+\frac{1}{2}\sum_{k=1}^{G}\frac{\left[vec\left(A_{ii}\right)_k\right]^2}{\left(\underline{c}^{CSH}\right)^{1-\gamma_{ij}^{CSH}}}\right)$$

4. Sample γ_{ij}^{DI} from

$$\left(\gamma_{ij}^{DI}|-\right) \sim Bernoulli\left(\omega_{ij}^{DI}\right), \tag{A.16}$$

where $\omega_{ij}^{DI} = \frac{u_{2,ij}}{u_{1,ij}+u_{2,ij}}$ with $u_{1,ij} = \phi\left(vec\left(A_{ij}\right)|\mathbf{0}, \underline{\tau}_1^2 I_{G^2}\right)\pi_{ij}^{DI}$ and $u_{2,ij} = \phi\left(vec\left(A_{ij}\right)|\mathbf{0}, \underline{\tau}_2^2 I_{G^2}\right)\left(1-\pi_{ij}^{DI}\right)$, and $\phi\left(x|a,b\right)$ denotes the p.d.f. of the Normal distribution with mean a and variance b evaluated at x.

5. Sample π_{ij}^{DI} from

$$\left(\pi_{ij}^{DI}|-\right) \sim Beta\left(1+\sum\gamma_{ij}^{DI}, \underline{\varphi}+\sum\left(1-\gamma_{ij}^{DI}\right)\right). \tag{A.17}$$

6. Sample γ_{ij}^{CSH} from

$$\left(\gamma_{ij}^{CSH}|-\right) \sim Bernoulli\left(\omega_{ij}^{CSH}\right), \tag{A.18}$$

where $\omega_{ij}^{CSH} = \frac{v_{2,ij}}{v_{1,ij}+v_{2,ij}}$ with $v_{1,ij} = \phi\left(vec\left(A_{ii}\right)|vec\left(A_{jj}\right), \xi_{ij}^2 \times \underline{c}^{CSH} \times I_{G^2}\right)\pi_{ij}^{CSH}$ and $v_{2,ij} = \phi\left(vec\left(A_{ii}\right)|vec\left(A_{jj}\right), \xi_{ij}^2 \times I_{G^2}\right)\left(1-\pi_{ij}^{CSH}\right)$.

7. Sample π_{ij}^{CSH} from

$$\left(\pi_{ij}^{CSH}|-\right) \sim Beta\left(1+\sum\gamma_{ij}^{CSH}, \underline{\varphi}+\sum\left(1-\gamma_{ij}^{CSH}\right)\right). \tag{A.19}$$

8. Sampling $vec\left(\Psi_{ij}\right)$ and ψ_{kl}^{ii}, follows exactly the algorithm of Appendix A of George, Sun and Ni (2008) as applied to $G \times G$ blocks of the error covariance matrix (as opposed to individual elements).

The reader can also see complete details of our algorithm by looking at our panel VAR MATLAB code, which is available through the website:
https://sites.google.com/site/dimitriskorobilis/matlab/panel_var_restrictions.

Empirical Appendix

We begin by presenting evidence on the convergence of the MCMC algorithms for the eight models used in Section 5. For each model in that section, we use 220,000 MCMC draws. An initial 20,000 draws are discarded and, from the remaining 200,000, every 10th draw is retained leaving us with a chain of 20,000 draws. As MCMC diagnostics, we consider the correlation between MCMC draws which are well separated in the chain (in our case 10 draws apart) and the MCMC inefficiency factor. Since α contains up to 900 parameters and Σ up to 435 parameters, we present results using boxplots of each diagnostic for each parameter. Figures A.1 through A.4 show that the algorithm for M1 (our preferred model) sometimes converges slightly slower than other algorithms, but it is always is converging fast enough such that our 20,000 retained draws should be sufficient to produce accurate estimates of posterior moments. For instance, with M1, inefficiency factors for for the great majority of the elements of α and Σ are less than 5 and there are only a handful of inefficiency factors greater than 50 and none greater than 70.

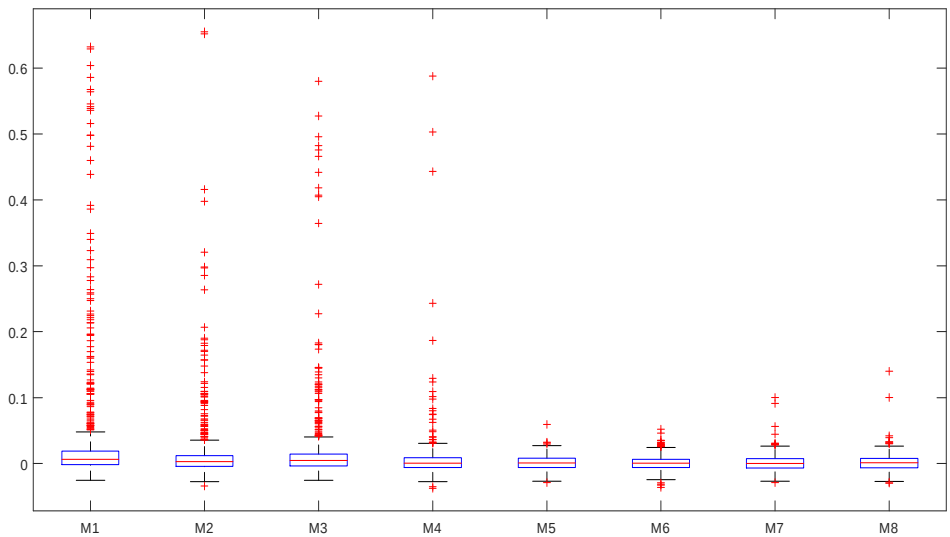

Figure A1: Boxplots of order 10 autocorrelations of MCMC draws of α

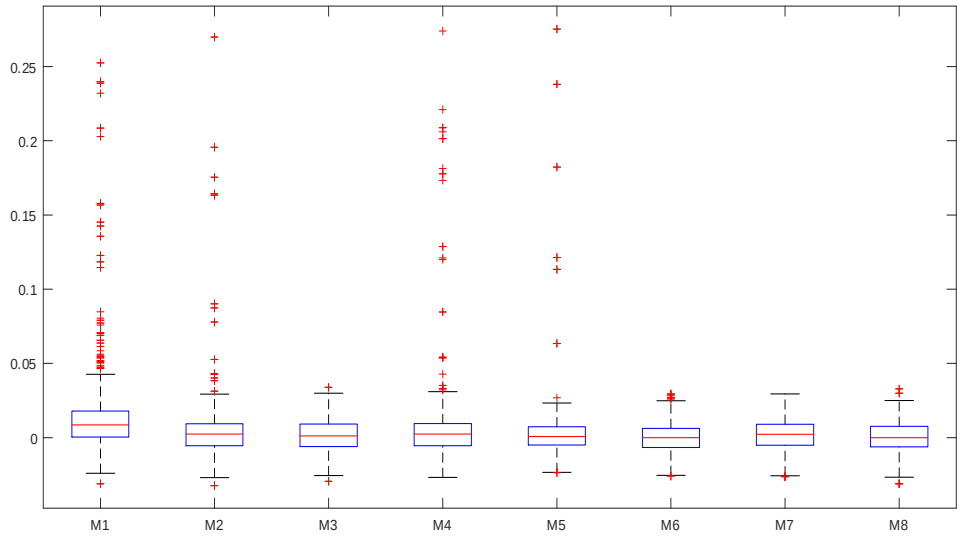

Figure A2: Boxplots of order 10 autocorrelations of MCMC draws of Σ

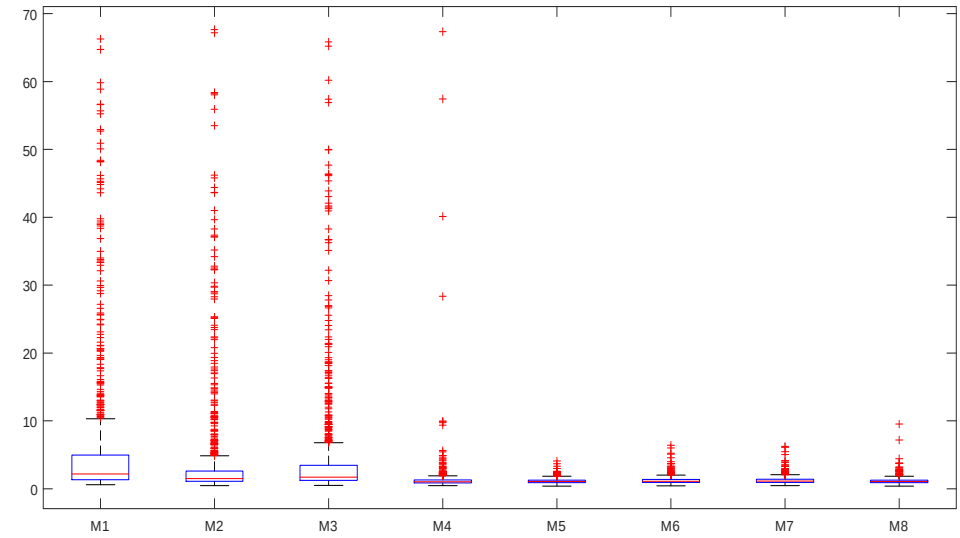

Figure A3: Boxplots of inefficiency factors for elements of α

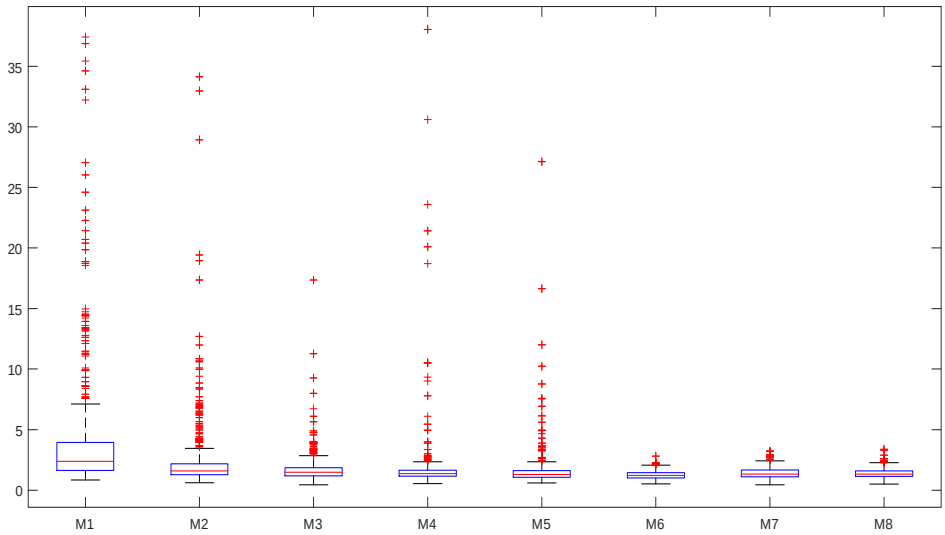

Figure A4: Boxplots of inefficiency factors for elements of Σ

In the body of the text, we presented impulse responses only for M1 (which searched over all types of restrictions) and M8 (the completely unrestricted model). In order to see the impact of some of the restrictions that the other models imply, Figure A.5 presents impulse responses for M1, M6, M7 and M8. Since our impulse response is a measure of the impact of a Greek shock on other countries, all models which do not allow for interdependencies between countries will, by definition, have zero impulse responses for countries other than Greece. Accordingly, Figure A.5 only presents impulse responses for Greece.

It can be seen that M7, which imposes cross-country homogeneity restrictions on all countries leads to an impulse response which is roughly one at all horizons. This is very different, in a counter-intuitive direction, from the results M1 is producing. M6, which uses a VAR involving only Greek variables, is also producing impulse responses at odds with M1. Among other differences, M6 indicates the shock is much more persistent than M1 indicates.

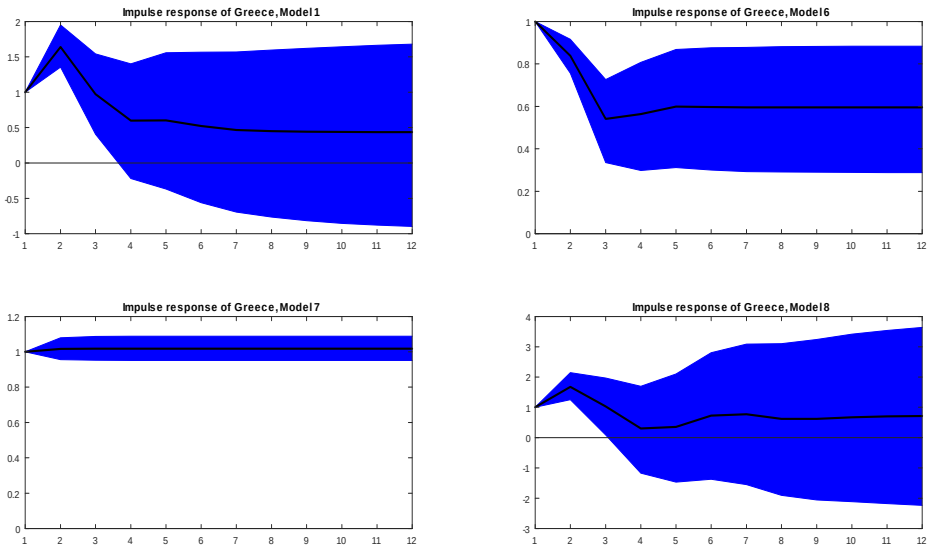

Figure A5: Comparing Greek rate impulse responses for different models