BANK OF GREECE
EUROSYSTEM

Working Paper

Alternative Bayesian compression in Vector Autoregressions and related models

Mike G. Tsionas

216

NOVEMBER 2016

ALTERNATIVE BAYESIAN COMPRESSION IN VECTOR AUTOREGRESSIONS AND RELATED MODELS

Mike G. Tsionas
Athens University of Economics and Business

ABSTRACT

In this paper we reconsider large Bayesian Vector Autoregressions (BVAR) from the point of view of Bayesian Compressed Regression (BCR). First, we show that there are substantial gains in terms of out-of-sample forecasting by treating the problem as an error-in-variables formulation and estimating the compression matrix instead of using random draws. As computations can be efficiently organized around a standard Gibbs sampler, timings and computa-tional complexity are not affected severely. Second, we extend the Multivariate Autoregressive Index model to the BCR context and show that we have, again, gains in terms of out-of-sample forecasting. The new techniques are used in U.S data featuring medium-size, large and huge BVARs.

Keywords: Bayesian Vector Autoregressions; Bayesian Compressed Re-gression; Error-in-Variables; Forecasting; Multivariate Autoregressive Index model.

JEL Classifications: C11, C13.

Acknowledgements: The author wishes to thank the Bank of Greece for its hospitality and funding the research in this paper. Excellent research assistance by Xingzhi Yao is gratefully acknowledged.. The views of the paper are of the author and do not necessarily reflect those of the Bank of Greece.

Correspondence:
E. G. Tsionas
Athens University of Economics and Business
76 Patission Str.,10434
Athens,Greece
email: tsionas@otenet.gr

1 Introduction

Recent attempts to improve the out-of-sample forecasting performance of Bayesian Vector Autoregressions (BVAR) include the application of ideas from Bayesian Compressed Regression (BCR). Guhaniyogi and Dunson (2015) propose BCR as a way to deal with the problem of proliferation of parameters when the number of regressors exceeds the number of observations. Their procedure is based on the idea that a compressed version of the regressor matrix can be constructed to replace the original model. Koop, Korobilis and Pettenuzzo (2016, KKP) have applied BCR to BVARs. BVAR problems can be hugely dimensional as the number of (lagged) predictors can easily exceed typical sample sizes in economics.

In Guhaniyogi and Dunson (2015) the compression matrix is drawn randomly, it involves a single parameter and, finally, it is selected using analytical expressions for the marginal likelihood of linear models, when conjugate priors are used. KKP extend this approach in an ingenious way by showing how the error covariance matrix of BVARs can be compressed as well. In this paper we show that, in fact, BCR can be viewed as an errors-in-variables problem. The compression or projection matrix can then be estimated using a standard Gibbs sampler. Although the fact that the projection matrix is not estimated in BCR is important, it is equally important to investigate whether gains can be realized in terms of out-of-sample forecasting by increasing CPU time through formal estimation via the Gibbs sampler. Indeed, it turns out that with the Guhaniyogi and Dunson (2015) prior for the projection matrix, important gains can be realized.

We extend the procedure to allow for time-varying-parameter vector autoregressions (TVP-VAR) and we, additionally, propose a substantive extension of the Multivariate Autoregressive Index model (MAI, Carriero, Kapetanios and Marcellino, 2016). The MAI depends on an expensive Metropolis step which can benefit from BCR and application of the Gibbs sampler.

In Section 2 we present BCR as an errors-in-variables problem. In Section 3 we summarize issues in BVARs and we propose our alternative. In Section 4 we consider the extension of MAI. TVP-VARs are presented in Section 5. Our empirical applications follow closely the ones in KKP and the results are compared and contrasted in Section 6. In section 7 we consider an application to the data by Giannone, Lenza and Primiceri (GLP, 2015) whose focus is on prior selection for Bayesian Vector Autoregressions In the final section, we offer some concluding remarks.

2 Bayesian Compressed Regression as an Error-in -Variables Problem

An alternative way to understand Bayesian compressed regression (BCR) is the following. Suppose we want a "small" model of the form:

$$y_t = \tilde{x}_t' \gamma + u_t, \tag{1}$$

where $u_t \sim N(0, \sigma_u^2)$ but the data is $(y_t, x_t; t = 1, ..., T)$ where $x_t \in \mathbb{R}^k$ with $k \gg T$ and $\tilde{x}_t$ is $m \times 1$ $(m \ll T)$ is a "compressed" or "more realistic" version of the regressors. Supposing

$$\tilde{x}_t = \Phi x_t + \varepsilon_t, \tag{2}$$

where Φ is an $m \times k$ matrix, and $\varepsilon_t \sim N_m(0, \Omega)$, effectively we have:

$$y_t = x_t' \Phi' \gamma + \varepsilon_t' \gamma + u_t \equiv (\Phi x_t)' \gamma + v_t, \tag{3}$$

where $v_t \sim N(0, \gamma' \Omega \gamma + \sigma_u^2)$. The point for BCR is that we can use (1) and (2) as a standard errors-in-the variables formulation and along the way estimate Φ instead of searching for it. Bayesian inference using the Gibbs sampler in (1) and (2) should be particularly easy. For example, the posterior conditional distribution of $\tilde{x}_t$ is:

$$\tilde{x}_t \sim N\left(\tilde{x}_t^*,\ V^*\right), \tag{4}$$

where $\tilde{x}_t^* = \left(\gamma\gamma' + \sigma_u^2\Omega^{-1}\right)^{-1}\left(y_t\gamma + \sigma_u^2\Omega^{-1}\Phi x_t\right)$, $V^* = \sigma_u^2\left(\gamma\gamma' + \sigma_u^2\Omega^{-1}\right)^{-1}\left(y_t\gamma + \sigma_u^2\Omega^{-1}\Phi x_t\right)$.

Of course Φ can be determined via random searches when the marginal likelihood is in closed form but when this is the case the priors are conjugate and they have their own well known problems. Under some sparsity, a LASSO prior can be used to relate "optimally" $\tilde{x}_t$ to x_t. Moreover, we can impose the orthonormality constraint $\Phi'\Phi = I_k$. Regarding a prior for matrix $\Phi = [\Phi_{ij}]$ we have:

$$\begin{gathered} P\left(\Phi_{ij} = \tfrac{1}{\sqrt{\psi}}\right) = \psi^2, \\ P\left(\Phi_{ij} = 0\right) = 2\psi\left(1-\psi\right), \\ P\left(\Phi_{ij} = -\tfrac{1}{\sqrt{\psi}}\right) = \left(1-\psi\right)^2, \end{gathered} \tag{5}$$

where ψ is a parameter (along with m). Guhaniyogi and Dunson (2015) show that their Bayesian compressed regression algorithm produces a predictive density which converges to the true predictive density under broad conditions and especially when sparsity is present. For the "window" m they draw from a uniform distribution, $U\left[2\log k, \min\left(T,k\right)\right]$. For ψ they draw from a uniform distribution in (0.1, 1). AS they correctly mention: "Given that the whole model averaging process is embarrassingly parallelizable over different choices of m, the computation can be done very quickly using a parallel implementation with sufficiently many processors."

Using the Gibbs sampler, suppose Ω is a diagonal matrix for simplicity and $\Omega = \mathrm{diag}\left[\omega_{11}, ..., \omega_{nn}\right]$. Without loss of generality consider the first equation of (2):

$$\tilde{x}_{t,1} = x_t'\varphi_1 + v_t,\ v_t \sim N\left(0, \omega_{11}\right), \tag{6}$$

where $\varphi_1 = [\varphi_{11}, ..., \varphi_{1n}]'$ is the first row of Φ. Under the prior in (5), the posterior conditional distribution of $\varphi_{1,i}$ $(i = 1, ..., n$) is

$$\begin{array}{c}\overline{P}_i(\kappa) \equiv p\left(\varphi_{1,i} = \kappa|\varphi_{1,(-i)}, \omega_{11}, \tilde{x}_{t,1}, x_t\right) \propto \\ \exp\left\{-\frac{1}{2\omega_{11}}\sum_{t=1}^{T}\left(x_{t,1} - \sum_{j=1,j\neq i}^{n}\varphi_{1,j}x_{tj} - \kappa x_{ti}\right)^2\right\} P(\kappa),\end{array} \tag{7}$$

where $\varphi_{1,(-i)}$ denotes all elements of φ_1 except the i^{th} element (Geweke, 1994), $\kappa \in \left\{\frac{1}{\sqrt{\psi}}, 0, -\frac{1}{\sqrt{\psi}}\right\}$ and $P(\kappa)$ is ψ^2, $2\psi\,(1-\psi)$ and $(1-\psi)^2$ respectively as 9n (5). Moreover, we normalize $\overline{P}_i(\kappa) = \frac{\overline{P}_i(\kappa)}{\sum_{\kappa'}\overline{P}_i(\kappa')}$ and we impose orthonormality after drawing all elements of Φ. Regarding priors for γ, σ_u and ω_{11} we use standard flat priors $p\,(\gamma, \sigma_u, \omega_{11}) \propto \sigma_u^{-1}\omega_{11}^{-1}$, and, therefore, their conditional posterior distributions are amenable to random number generation. Our prior for ψ is uniform in $[0.1,\ 1]$. The posterior conditional distribution of ψ is not in any known family but random drawings can be generated using the inverse c.d.f technique where the c.d.f is computed numerically over a grid of twenty points in $[0.1,\ 1]$. Although a Metropolis algorithm can be used as well we have found that it converges much slower and it is difficult to tune especially in medium to large sized problems.

3 Issues in Bayesian VARs

Koop, Korobilis and Pettenuzzo (2016, KKP) have applied BCR to Bayesian VARs. Bayesian VAR problems can be hugely dimensional as the number of (lagged) predictors can easily exceed typical sample size in economics. We summarize their work as follows. The VAR model is:

$$Y_t = BY_{t-1} + \epsilon_t, t = 1, ..., T, \tag{8}$$

where Y_t is an $n \times 1$ vector time series, $\epsilon_t \sim iidN(0, \Omega)$ and B is an $n \times n$ vector of coefficients. As KKP correctly note: "Note that, with $n = 100$, the uncompressed VAR will have 10, 000 coefficients in B and 5,050 in Ω. In a VAR(13), such as the one used in this paper, the former number becomes 130, 000. It is easy to see why computation can become daunting in large VARs and why there is a need for shrinkage." They also note that the natural conjugate prior has some well-known restrictive properties in VARs (Koop and Korobilis, 2009, pp. 279-280). For this reason, the need for new priors arises. One of their concerns is the following:

> "In the context of the compressed VAR, working with a Φ of dimension $m \times n$ [....] with only n columns instead of n^2 would likely be much too restrictive in many empirical contexts. For instance, it would imply that to delete a variable in one equation, then that same variable would have to be deleted from all equations. In macroeconomic VARs, where the first own lag in each equation is often found to have important explanatory power, such a property seems problematic. It would imply, say, that lagged inflation could either be included in every equation or none when what we might really want is for lagged inflation to be included in the inflation equation but not most of the other equations in the VAR."

The problem can be solved in a natural way as we have regressors that we do not want to compress This can be done as in the following alternative to (1):

$$y_t = w_t' \beta + \tilde{x}_t' \gamma + u_t, \tag{9}$$

where w_t is a vector of regressors that we do not wish to compress. Additionally, another additional issue with the natural conjugate BCVAR is that it allows the error covariance matrix Ω to be unrestricted. In high dimensional VARs, Ω contains a large number of parameters and we may want a method which allows for their compression. A natural alternative is to consider a VAR:

$$Y_t = \sum_{l=1}^{L} A_l \tilde{Y}_{t-l} + \Pi W_t + \epsilon_t, \tag{10}$$

where W contains the variables that we do not wish to compress, Π is (possibly) diagonal, and we have direct control over the prior for Φ if so desired. We can write the above equation as:

$$Y_t^* = \Gamma \tilde{X}_t + \epsilon_t, \tag{11}$$

where $\left[\tilde{Y}_{t-l}, l = 1, ..., L, W_t\right]$ are similar to $\tilde{x}_t$ in (1). We assume:

$$\epsilon_t \sim N(0, \Omega). \tag{12}$$

Following common practice (see, e.g., Primiceri, 2005, Eisenstat, Chan and Strachan, 2015 and Carriero, Clark and Marcellino, 2016) we use a triangular decomposition of Ω:

$$A\Omega A' = \Sigma\Sigma, \tag{13}$$

where $\Sigma = \text{diag}[\sigma_1, ..., \sigma_n]$ and A is lower triangular with ones on the main diagonal. Following KKP we write $A = I_n + \tilde{A}$ where $\tilde{A}$ is lower diagonal. In this way we can rewrite the VAR in a recursive way as:

$$Y_t = BY_{t-1} + A^{-1}\Sigma E_t, \tag{14}$$

where $E_t \sim N_n(0, I_n)$. Further, we can write it as:

$$Y_t = \Gamma Y_{t-1} + \tilde{A}(-Y_t) + \Sigma E_t \tag{15}$$

or

$$Y_t = \Theta Z_t + \Sigma E_t, \tag{16}$$

where $Z_t = \left[Y'_{t-1}, -Y'_t\right]'$, $\Gamma = AB$ and $\Theta = \left[\Gamma, \tilde{A}\right]$. This has the additional advantage that compression of the elements of the covariance matrix is achieved through $\tilde{A}$ and each equation in (16) has a recursive structure. An additional advantage is that each equation of (16) can be seen as a single equation and, therefore, different compression matrices, Φ, can be used.

4 An Alternative Approach

Following Carriero et al. (2016) we can write (8) as:

$$Y_t = BY_{t-1} + \epsilon_t \equiv ACY_{t-1} + \epsilon_t, \tag{17}$$

where B is $n \times n$, A is $n \times r$ and C is $r \times n$ where $r < n$. This is known as the multivariate autoregressive index (MAI) model and exploits reduced rank as compression regression does. The reason is that we can write $F_t = CY_{t-1}$ from which we have $Y_t = AF_t + \epsilon_t$. Carriero et al. (2016) use an expensive Metropolis-Hastings scheme to draw the elements of C from the posterior.

As a matter of fact one can try random compression for C instead of estimating it. Alternatively, we can treat F_t as latent and assume:

$$F_t = CY_{t-1} + u_t,\ u_t \sim N(0, \Psi), \tag{18}$$

along with

$$Y_t = AF_t + \epsilon_t. \tag{19}$$

Once this is done, it is not difficult to generalize the model in a substantive way by assuming:

$$F_t = RF_{t-1} + CY_{t-1} + u_t,\ u_t \sim N_r(0, \Psi), \tag{20}$$

where R is an $r \times r$ matrix. In (18) or (20) one can set $\Psi = O$ and $R = O$ and use standard random compression for C or estimate the elements of these matrices. Equation (19) can still be put in recursive form as in (16) using (13). In this case we assume that Ψ is diagonal and its diagonal elements are ones provided we have standardized the data (KKP also standardize the data). In this set up it is not difficult to estimate (19) and (20) using Sequential Monte Carlo (SMC) also known as Particle Filtering (PF). The selection of $r \in \{1, 2, ..., \overline{r}\}$ can be made using the marginal likelihood which is a natural byproduct of SMC.

5 Time-Varying Parameters and Models

5.1 Time-Varying Parameters

Time-varying parameter vector autoregressions (TVP-VAR) have proved useful in many contexts. In compressed regression or the errors-in-variables interpretation we propose here (in which case the compression matrix Φ is estimated) it is not difficult to introduce time-varying parameters. For example, in (1) which is a typical equation of the recursive system in (16) we can assume:

$$\gamma_t = d + D\gamma_{t-1} + v_t,\ v_t \sim N(0, \Sigma_\gamma). \tag{21}$$

Typically, it is assumed that $d = 0$ and $D = I$ so that the parameter vector follows a random walk. As the dimensionality of γ in (1) is not expected to be large we can avoid these assumptions and use SMC to perform the computations.

5.2 Summary of Models

In this paper we will consider alternative specifications as follows.

1. Model I, is the recursive VAR in (16) or (15) where compression of the covariance matrix Ω is allowed. Here, we use the Gibbs sampler developed for errors-in-variables problems.

2. Model II, is the alternative MAI in (19) and (18).

3. Model III, is the alternative MAI in (19) and (20) which allows for autoregressive F_t.

4. Models IV, V and VI are models I, II and III with the time-varying parameter structure in (21).

6 Empirical Results

6.1 Data

We use the same dataset as in KKP. To this dataset we use a medium size VAR, a large size VAR and a huge size VAR. As in KKP we use the FRED-MD data-base of monthly US variables from January 1960 through December 2014, see McCracken and Ng (2015) for a description of this macroeconomic data set, which includes a range of variables from a broad range of categories (e.g. output, capacity, employment and unemployment, prices, wages, housing, inventories and orders, stock prices, interest rates, exchange rates and monetary aggregates). We use the 129 variables for which complete data was available, after transforming all variables as in Appendix A.8 of KKP. We present detailed forecasting results for seven variables of interest: industrial production growth (INDPRO), the unemployment rate (UNRATE), total nonfarm employment (PAYEMS), the change in the Fed funds rate (FEDFUNDS), the change in the 10 year T-bill rate (GS10), the finished good producer price inflation (PPIFGS) and consumer price inflation (CPIAUCSL). We estimate VARs of different dimensions, with these seven variables included in all of our specifications. Specifically, we use a Medium VAR with 19 variables, a Large VAR with 46 variables and a Huge VAR with all 129 variables. As in KPP most of our variables have substantial persistence in them and, accordingly, the first own lag in each equation almost always has important explanatory power. Accordingly, we do not compress the first own lag. This is included in every equation, with compression being done on the remaining variables.

6.2 Predictive Accuracy

We try to follow KKP as closely as possible. For this reason we use the first half of the sample, January 1960-June 1987, to obtain initial parameter estimates for all models, which are then used to predict outcomes from July 1987 ($h = 1$) to June 1987 ($h = 12$). The next period, we include data for July 1987 in the estimation sample, and use the resulting estimates to predict the outcomes from August 1987 to July 1988. We proceed recursively in this fashion until December 2014, thus generating a time series of forecasts for each forecast horizon h, with $h = 1, ..., 12$. When $h > 1$, point forecasts are iterated and predictive simulation is used to produce the predictive densities.

Our measures of predictive precision are relative to an AR(1) benchmark which is a reasonable model as most series are highly persistent. To examine the precision of the h-step-ahead point forecasts for a given model i we use the ratio of mean-squared forecast errors:

$$MSFE_{i,j,h} = \frac{\sum_{\tau=t_1}^{t_2-h} e_{i,j,\tau+h}^2}{\sum_{\tau=t_1}^{t_2} e_{AR1,j,\tau+h}^2}, \tag{22}$$

where t_1 and t_2 denote the start and end of the out-of-sample period, and $e_{i,j,\tau+h}^2$, $e_{AR1,j,\tau+h}^2$ are the squared forecast errors of variable j at time τ from models i and

AR(1) respectively when the forecast horizon is h. We take the results of forecasting for all variables and models DFM (dynamic factor model), FAVAR (factor augmented VAR), BVAR (Bayesian VAR), BCVAR (Bayesian Compressed VAR) and BCVARC (Bayesian Compressed VAR with compression for the covariance matrix Ω as well) directly from KKP. The point forecasts d to compute the forecast errors are computed by averaging over all draws from the various models' h-step-ahead predictive densities.

KPP also follow Christoersen and Diebold (1998) and consider the ratio between the multivariate weighted mean squared forecast error (WMSFE) of model i and the WMSFE of the benchmark AR(1) model as follows:

$$WMSFE_{i,h} = \frac{\sum_{\tau=t_1}^{t_2-h} we_{i,\tau+h}}{\sum_{\tau=t_1}^{t_2-h} we_{AR1,\tau+h}}, \tag{23}$$

where $we_{i,\tau+h} = e'_{i,\tau+h} W e_{i,\tau+h}$ and $we_{AR1,\tau+h} = e'_{AR1,\tau+h} W e_{AR1,\tau+h}$ are time $\tau + h$ weighted forecast errors of model i and the benchmark AR(1) respectively, $e_{i,\tau+h}$ and $e_{AR1,\tau+h}$ are 7×1 vector of forecast errors for our key series and W is a 7×7 weighting matrix selected as in Carriero et al (2011) to be a diagonal matrix with the inverse variance series along the main diagonal.

Following KKP and Geweke and Amisano (2010) we also use the log predictive likelihood differential between model i and the benchmark AR(1) model:

$$ALPL_{i,j,h} = \frac{1}{t_2 - t_1 - h + 1} \sum_{\tau=t_1}^{t_2-h} \left(LPL_{i,j,\tau+h} - LPL_{AR1,j,\tau+j} \right), \tag{24}$$

where $LPL_{i,j,\tau+h}, LPL_{AR1,j,\tau+j}$ are log predictive scores of variable j for model i and benchmark AR(1) respectively at time $\tau+h$, i.e. the log of the h-step-ahead predictive density evaluated at the outcome. Positive values of ALPL indicate that on average model i produces better forecasts compared to the benchmark. The multivariate analogue is as follows:

$$MVALPL_{i,j,h} = \frac{1}{t_2 - t_1 - h + 1} \sum_{\tau=t_1}^{t_2-h} \left(MVLPL_{i,j,\tau+h} - MVLPL_{AR1,j,\tau+j} \right), \quad (25)$$

where $MVLPL_{i,j,\tau+h}$ and $MVLPL_{AR1,j,\tau+h}$ denote the multivariate log predictive likelihoods for model i and benchmark AR(1) respectively, at time $\tau + h$. One difference relative to KKP is that they use a multivariate normal approximation for this quantity, that is they assume joint normality. In this study we use a multivariate kernel density estimator obtained from 10^6 draws for each forecast horizon. This produces different results compared to a multivariate normal approximation but we abstract from further investigation leaving the issue to a future study.

Our empirical results are summarized in Tables 1 through 7. All results that we present are "statistically significant" in the sense of rejecting equal predictive accuracy (Diebold and Mariano, 1995) using the procedures described in KKP (pp. 17-18).

7 Another empirical application

In this section we consider the same data as in Giannone, Lenza and Primiceri (GLP, 2015) whose focus is on prior selection for Bayesian Vector Autoregressions. Their Large BVAR model contains 22 variables (see their Table 1, p. 441) which include GDP, GDP deflator, federal funds rate, consumption, investment, hours, wages a and a number of additional labor market, financial, and monetary variables. Here, we focus attention on two things. First, a comparison in terms of forecasting ability and, secondly and, perhaps more importantly, to obtain estimates of impulse response functions of real, nominal, and financial variables after a monetary shock (one-standard deviation increase in federal funds rate). The results are reported in Table 8 and our impulse response functions are reported in Figures 1-3.

The approach proposed in this study dominates GLP in terms of forecasting performance. This is not surprising as it was found to dominate huge VARs in the previous sections. In turn, GLP show that their approach dominates flat-prior based VARs and reduces considerably the uncertainty associated with impulse reponse functions. The same is the case in this study, as the impulse response functions seem to be very similar to those in GLP but the 95% Bayes probability intervals are quite tight in many cases of interest. We follow the same identification scheme as in GLP and our findings in terms of policy are quite similar: "A 1-standard deviation (approximately 60 basis points) exogenous increase in the federal funds rate generates a substantial contraction in GDP, employment, and all other variables related to economic activity. Monetary aggregates also decrease on impact, indicating strong liquidity effects. Moreover, stock prices decline, the exchange rate appreciates, and the yield curve flattens. Prices decrease with a delay. Notice that with the exception of the CPI, the response of prices does not exhibit the so-called price puzzle, that is, a counterintuitive positive response to a monetary contraction, which is instead typical of VARs with small information sets" (GLP, p. 446).[1]

Concluding remarks

In this paper we have proposed an errors-in-variables interpretation of Bayesian Compressed Vector Autoregression. Koop, Korobilis and Pettenuzzo (2016) have applied BCR to BVARs. BVAR problems can be hugely dimensional as the number of (lagged) predictors can easily exceed typical sample sizes in economics. In Guhaniyogi and Dunson (2015) the compression matrix is drawn randomly, it involves a single parameter and, finally, it is selected using analytical expressions for the marginal

[1]One difference in Figures 2-4 of GLP is that the gray areas refer to error bands for a flat-prior Bayesian VAR whereas in this study they reflect the 95% Bayes probability interval, computed using simulation techniques. Moreover, GLP do not report 95% Bayes probability intervals but 50% and certain other quantiles of the impulse response function distributions.

likelihood of linear models, when conjugate priors are used. KKP extend this approach in an ingenious way by showing how the error covariance matrix of BVARs can be compressed as well. In this paper we show that, in fact, BCR can be viewed as an errors-in-variables problem. The compression or projection matrix can then be estimated using a standard Gibbs sampler. We find that important gains can be realized in terms of out-of-sample forecasting at the cost of increasing CPU timings through formal estimation via the Gibbs sampler. Indeed, it turns out that with the Guhaniyogi and Dunson (2015) prior for the projection matrix, important gains can be realized. We illustrate the new approach in small, medium and huge sized VARs and to the Bayesian VAR analysis of Giannone, Lenza and Primiceri (2015).

References

Banbura, M., D. Giannone, L. Reichlin (2010). Large Bayesian VARs. Journal of Applied Econometrics 25, 71-92.

Carriero, A., Kapetanios, G. and Marcellino, M. (2016). Structural analysis with multivariate autoregressive index models. Journal of Econometrics 192 (2), 332-348.

Diebold, F. X. and Mariano, R. S. (1995). Comparing Predictive Accuracy. Journal of Business and Economic Statistics, 13, 253-263.

Eisenstat, E., Chan, J. and Strachan, R. (2015). Stochastic model specification search for time-varying parameter VARs, Econometric Reviews, forthcoming.

Geweke, J. (1994). Variable Selection and Model Comparison in Regression. Working paper, University of Minnesota and Federal Reserve Bank of Minneapolis.

Geweke, J. and Amisano, G. (2010). Comparing and evaluating Bayesian predictive distributions of asset returns. International Journal of Forecasting, 26(2), 216-230.

Giannone. D., M. Lenza, G. Primiceri (2015). Prior selection for vector autoregressions. Review of Economics and Statistics 97 (2), 436-451.

Guhaniyogi, R. and Dunson, D. (2015). Bayesian compressed regression. Journal of the American Statistical Association, 110, 1500-1514.

Koop, G. and Korobilis, D. (2009). Bayesian multivariate time series methods for empirical macroeconomics. Foundations and Trends in Econometrics, 3, 267-358.

Koop, G., D. Korobilis, D. Pettenuzzo (2016). Bayesian Compressed Vector Autoregressions, working paper.

Litterman, R. (1980). A Bayesian procedure for forecasting with vector autoregression. MIT, Department of Economics working paper.

McCracken, M. and Ng, S. (2015). FRED-MD: A monthly database for macroeconomic research. Federal Reserve Bank of St. Louis, working paper 2015-012A.

Primiceri, G., (2005). Time varying structural vector autoregressions and monetary policy. Review of Economic Studies, 72, 821-852.

Sims, C.A., T. Zha (1998). Bayesian methods for dynamic multivariate models, International Economic Review 39, 949-968.

TECHNICAL APPENDIX

Particle filtering

The particle filter methodology can be applied to state space models of the general form:

$$y_T \sim p(y_t|x_t),\ s_t \sim p(s_t|s_{t-1}), \tag{A.1}$$

where s_t is a state variable. For general introductions see Gordon (1997), Gordon et al. (1993), Doucet et al (2001) and Ristic et al. (2004).

Given the data Y_t the posterior distribution $p(s_t|Y_t)$ can be approximated by a set of (auxiliary) particles $\left\{s_t^{(i)}, i = 1, ..., .N\right\}$ with probability weights $\left\{w_t^{(i)}, i = 1, ..., N\right\}$ where $\sum_{i=1}^{N} w_t^{(i)} = 1$. The predictive density can be approximated by:

$$p(s_{t+1}|Y_t) = \int p(s_{t+1}|s_t)p(s_t|Y_t)ds_t \simeq \sum_{i=1}^{N} p(s_{t+1}|s_t^{(i)})w_t^{(i)}, \tag{A.2}$$

and the final approximation for the filtering density is:

$$p(s_{t+1}|Y_t) \propto p(y_{t+1}|s_{t+1})p(s_{t+1}|Y_t) \simeq p(y_{t+1}|s_{t+1}) \sum_{i=1}^{N} p(s_{t+1}|s_t^{(i)})w_t^{(i)}. \tag{A.3}$$

The basic mechanism of particle filtering rests on propagating $\left\{s_t^{(i)}, w_t^{(i)}, i = 1, \ldots, N\right\}$ to the next step, viz. $\left\{s_{t+1}^{(i)}, w_{t+1}^{(i)}, i = 1, \ldots, N\right\}$ but this often suffers from the weight degeneracy problem. If parameters $\theta \in \Theta \in \Re^k$ are available, as is often the case, we follow Liu and West (2001) and parameter learning takes place via a mixture of multivariate normals:

$$p(\theta|Y_t) \simeq \sum_{i=1}^{N} w_t^{(i)} N(\theta|a\theta_t^{(i)} + (1-a)\bar{\theta}_t, b^2 V_t), \tag{A.4}$$

where $\bar{\theta}_t = \sum_{i=1}^{N} w_t^{(i)}\theta_t^{(i)}$, and $V_t = \sum_{i=1}^{N} w_t^{(i)}(\theta_t^{(i)} - \bar{\theta}_t)(\theta_t^{(i)} - \bar{\theta}_t)'$. The constants a

and b are related to shrinkage and are determined via a discount factor $\delta \in (0,1)$ as $a = (1-b^2)^{1/2}$ and $b^2 = 1 - [(3\delta - 1)/2\delta]^2$. See also Casarin and Marin (2007).

Andrieu and Roberts (2009), Flury and Shephard (2011) and Pitt et al. (2012) provide the Particle Metropolis-Hastimgs (PMCMC) technique which uses an unbiased estimator of the likelihood function $\hat{p}_N(Y|\theta)$ as $p(Y|\theta)$ is often not available in closed form.

Given the current state of the parameter $\theta^{(j)}$ and the current estimate of the likelihood, say $L^j = \hat{p}_N(Y|\theta^{(j)})$, a candidate θ^c is drawn from $q(\theta^c|\theta^{(j)})$ yielding $L^c = \hat{p}_N(Y|\theta^c)$. Then, we set $\theta^{(j+1)} = \theta^c$ with the Metropolis - Hastings probability:

$$A = \min\left\{1, \frac{p(\theta^c)L^c}{p(\theta^{(j)}L^j} \frac{q(\theta^{(j)}|\theta^c}{q(\theta^c|\theta^{(j)})}\right\}, \tag{A.5}$$

otherwise we repeat the current draws: $\left\{\theta^{(j+1)}, L^{j+1}\right\} = \left\{\theta^{(j)}, L^j\right\}$.

Hall, Pitt and Kohn (2014) propose an auxiliary particle filter which rests upon the idea that adaptive particle filtering (Pitt et al., 2012) used within PMCMC requires far fewer particles that the standard particle filtering algorithm to approximate $p(Y|\theta)$. From Pitt and Shephard (1999) we know that auxiliary particle filtering can be implemented easily once we can evaluate the state transition density $p(s_t|s_{t-1})$. When this is not possible, Hall, Pitt and Kohn (2014) present a new approach when, for instance, $s_t = g(s_{t-1}, u_t)$ for a certain disturbance. In this case we have:

$$p(y_t|s_{t-1}) = \int p(y_t|s_t)p(s_t|s_{t-1})ds_t, \tag{A.6}$$

$$p(u_t|s_{t-1}; y_t) = p(y_t|s_{t-1}, u_t)p(u_t|s_{t-1})/p(y_t|s_{t-1}). \tag{A.7}$$

If one can evaluate $p(y_t|s_{t-1})$ and simulate from $p(u_t|s_{t-1}; y_t)$ the filter would be fully adaptable (Pitt and Shephard, 1999). One can use a Gaussian approximation for the first-stage proposal $g(y_t|s_{t-1})$ by matching the first two moments of

$p(y_t|s_{t-1})$. So in some way we find that the approximating density $p(y_t|s_{t-1}) = N\left(\mathbb{E}(y_t|s_{t-1}), \mathbb{V}(y_t|s_{t-1})\right)$. In the second stage, we know that $p(u_t|y_t, s_{t-1}) \propto p(y_t|s_{t-1}, u_t)p(u_t)$. For $p(u_t|y_t, s_{t-1})$ we know it is multimodal so suppose it has M modes are $\hat{u}_t^m$, for $m = 1, \ldots, M$. For each mode we can use a Laplace approximation. Let $l(u_t) = log\left[p(y_t|s_{t-1}, u_t)p(u_t)\right]$. From the Laplace approximation we obtain:

$$l(u_t) \simeq l(\hat{u}_t^m) + \tfrac{1}{2}(u_t - \hat{u}_t^m)'\nabla^2 l(\hat{u}_t^m)(u_t - \hat{u}_t^m). \tag{A.8}$$

Then we can construct a mixture approximation:

$$g(u_t|x_t, s_{t-1}) = \sum_{m=1}^{M} \lambda_m (2\pi)^{-d/2} |\Sigma_m|^{-1/2} \exp\left\{\tfrac{1}{2}(u_t - \hat{u}_t^m)'\Sigma_m^{-1}(u_t - \hat{u}_t^m\right\}, \tag{A.9}$$

where $\Sigma_m = -\left[\nabla^2 l(\hat{u}_t^m)\right]^{-1}$ and $\lambda_m \propto \exp\left\{l(u_t^m)\right\}$ with $\sum_{m=1}^{M} = 1$. This is done for each particle s_t^i. This is known as the Auxiliary Disturbance Particle Filter (ADPF).

An alternative is the independent particle filter (IPF) of Lin et al. (2005). The IPF forms a proposal for s_t directly from the measurement density $p(y_t|s_t)$ although Hall, Pitt and Kohn (2014) are quite right in pointing out that the state equation can be very informative.

In the standard particle filter of Gordon et al. (1993) particles are simulated through the state density $p(s_t^i|s_{t-1}^i)$ and they are re-sampled with weights determined by the measurement density evaluated at the resulting particle, viz. $p(y_t|s_t^i)$.

The ADPF is simple to construct and rests upon the following steps:

For $t = 0, \ldots, T-1$ given samples $s_t^k \sim p(s_t|Y_{1:t})$ with mass π_t^k for $k = 1, ..., N$.

1) For $k = 1, \ldots, N$ compute $\omega_{t|t+1}^k = g(y_{t+1}|s_t^k)\pi_t^k$, $\pi_{t|t+1}^k = \omega_{t|t+1}^k / \sum_{i=1}^{N} \omega_{t|t+1}^i$.

2) For $k = 1, \ldots, N$ draw $\tilde{s}_t^k \sim \sum_{i=1}^{N} \pi_{t|t+1}^i \delta_{s_t}^i(ds_t)$.

3) For $k = 1, \ldots, N$ draw $u_{t+1}^k \sim g(u_{t+1}|\tilde{s}_t^k, y_{t+1})$ and set $s_{t+1}^k = h(s_t^k; u_{t+1}^k)$.

4) For $k = 1, \ldots, N$ compute

$$\omega_{t+1}^{k} = \frac{p(y_{t+1}|s_{t+1}^{k})p(u_{t+1}^{k})}{g(y_{t+1}|s_{t}^{k})g(u_{t+1}^{k}|\tilde{s}_{t}^{k}, y_{t+1})}, \pi_{t+1}^{k} = \frac{\omega_{t+1}^{k}}{\sum_{i=1}^{N} \omega_{t+1}^{i}}. \tag{A.10}$$

It should be mentioned that the estimate of likelihood from ADPF is:

$$p(Y_{1:T}) = \prod_{t=1}^{T} \left(\sum_{i=1}^{N} \omega_{t-1|t}^{i} \right) \left(N^{-1} \sum_{i=1}^{N} \omega_{t}^{i} \right). \tag{A.11}$$

Particle Metropolis adjusted Langevin filters

Nemeth, Sherlock and Fearnhead (2014) provide a particle version of a Metropolis adjusted Langevin algorithm (MALA). In Sequential Monte Carlo we are interested in approximating $p(s_t|Y_{1:t}, \theta)$. Given that:

$$p(s_t|Y_{1:t}, \theta) \propto g(y_t|x_t, \theta) \int f(s_t|s_{t-1}, \theta) p(s_{t-1}|y_{1:t-1}, \theta) ds_{t-1}, \tag{A.12}$$

where $p(s_{t-1}|y_{1:t-1}, \theta)$ is the posterior as of time $t-1$. If at time $t-1$ we have a set set of particles $\left\{s_{t-1}^{i}, i = 1, \ldots, N\right\}$ and weights $\left\{w_{t-1}^{i}, i = 1, \ldots .N\right\}$ which form a discrete approximation for $p(s_{t-1}|y_{1:t-1}, \theta)$ then we have the approximation:

$$\hat{p}(s_{t-1}|y_{1:t-1}, \theta) \propto \sum_{i=1}^{N} w_{t-1}^{i} f(s_t|s_{t-1}^{i}, \theta). \tag{A.13}$$

See Andrieu et al. (2010) and Cappe at al. (2005) for reviews. From (A.13) Fernhead (2007) makes the important observation that the joint probability of sampling particle s_{t-1}^{i} and state s_t is:

$$\omega_t = \frac{w_{t-1}^{i} g(y_t|s_t, \theta) f(s|s_{t-1}^{i}, \theta)}{\xi_t^{i} q(s_t|s_{t-1}^{i}, y_t, \theta)}, \tag{A.14}$$

where $q(s_t|s_{t-1}^i, y_t, \theta)$ is a density function amenable to simulation and

$$\xi_t^i q(s_t|s_{t-1}^i, y_t, \theta) \simeq cg(y_t|s_t, \theta) f(s_t|s_{t-1}^i, \theta), \tag{A.15}$$

and c is the normalizing constant in (A.12).

In the MALA algorithm of Roberts and Rosenthal (1998)[2] we form a proposal:

$$\theta^c = \theta^{(s)} + \lambda z + \tfrac{\lambda^2}{2} \nabla \mathrm{log} p(\theta^{(s)}|Y_{1:T}), \tag{A.16}$$

where $z \sim N(0, I)$ which should result in larger jumps and better mixing properties, plus lower autocorrelations for a certain scale parameter λ. Acceptance probabilities are:

$$a(\theta^c|\theta^{(s)}) = \min\left\{1, \frac{p(Y_{1:T}|\theta^c)q(\theta^{(s)}|\theta^c)}{p(Y_{1:T}|\theta^{(s)})q(\theta^c|\theta^{(s)})}\right\}. \tag{A.17}$$

Using particle filtering it is possible to create an approximation of the score vector using Fisher's identity:

$$\nabla \log p(Y_{1:T}|\theta) = E\left[\nabla \log p(s_{1:T}, Y_{1:T}|\theta)|Y_{1:T}, \theta\right], \tag{A.18}$$

which corresponds to the expectation of:

$$\nabla \log p(s_{1:T}, Y_{1:T}|\theta) = \nabla \log p(|s_{1:T-1}, Y_{1:T-1}|\theta) + \nabla \log g(y_T|s_T, \theta) + \nabla \log f(s_T|s|_{T-1}, \theta),$$

over the path $s_{1:T}$. The particle approximation to the score vector results from replacing $p(s_{1:T}|Y_{1:T}, \theta)$ with a particle approximation $\hat{p}(s_{1:T}|Y_{1:T}, \theta)$. With particle i at time t-1 we can associate a value $\alpha_{t-1}^i = \nabla \log p(s_{1:t-1}^i, Y_{1:t-1}|\theta)$ which can be updated recursively. As we sample κ_i in the APF (the index of particle at time $t-1$ that is

[2]The benefit of MALA over Random-Walk-Metropolis arises when the number of parameters n is large. This happens because the scaling parameter λ is $O(n^{-1/2})$for Random-Walk-Metropolis but it is $O(n^{-1/6})$ for MALA, see Roberts et al. (1997) and Roberts and Rosenthal (1998)

propagated to produce the ith particle at time t) we have the update:

$$\alpha_t^i = a_{t-1}^{\kappa_i} + \nabla \log g(y_t|s_t^i, \theta) + \nabla \log f(s_t^i|s_{t-1}^i, \theta). \tag{A.19}$$

To avoid problems with increasing variance of the score estimate $\nabla \log p(Y_{1:t}|\theta)$ we can use the approximation:

$$\alpha_{t-1}^i \sim N(m_{t-1}^i, V_{t-1}). \tag{A.20}$$

The mean is obtained by shrinking α_{t-1}^i towards the mean of α_{t-1} as follows:

$$m_{t-1}^i = \delta \alpha_{t-1}^i + (1-\delta) \sum_{i=1}^N w_{t-1}^i \alpha_{t-1}^i, \tag{A.21}$$

where $\delta \in (0,1)$ is a shrinkage parameter. Using Rao-Blackwellization one can avoid sampling α_t^i and instead use the following recursion for the means:

$$m_t^i = \delta m_{t-1}^{\kappa_i} + (1-\delta) \sum_{i=1}^N w_{t-1}^i m_{t-1}^i + \nabla \log g(y_t|s_t^i, \theta) + \nabla \log f(s_t^i|s_{t-1}^{\kappa_i}, \theta), \tag{A.22}$$

which yields the final score estimate:

$$\nabla \log \hat{p}(Y_{1:t}|\theta) = \sum_{i=1}^N w_t^i m_t^i. \tag{A.23}$$

As a rule of thumb Nemeth, Sherlock and Fearnhead (2014) suggest taking $\delta = 0.95$. Furthermore, they show the important result that the algorithm should be tuned to the asymptotically optimal acceptance rate of 15.47% and the number of particles must be selected so that the variance of the estimated log-posterior is about 3. Additionally, if measures are not taken to control the error in the variance of the score vector, there is no gain over a simple random walk proposal.

Of course, the marginal likelihood is:

$$p(Y_{1:T}|\theta) = p(y_1|\theta) \prod_{t=2}^{T} p(y_t|Y_{1:t-1}, \theta), \tag{A.24}$$

where

$$p(y_t|Y_{1:t-1}, \theta) = \int g(y_t|s_t) \int f(s_t|s_{t-1}, \theta) p(s_{t-1}|Y_{1:T-1}, \theta) ds_{t-1} ds_t, \tag{A.25}$$

provides, in explicit form, the predictive likelihood.

Additional References

Andrieu, C., and G.O. Roberts. (2009) The pseudo-marginal approach for efficient computation. Ann. Statist., 37, 697–725.

Andrieu, C., A. Doucet, and R. Holenstein (2010), "Particle Markov chain Monte Carlo methods," Journal of the Royal Statistical Society: Series B, 72, 269–342.

Cappé, O., E. Moulines, E., and T. Rydén (2005). Inference in Hidden Markov Models. Springer, Berlin.

Casarin, R., J.-M. Marin (2007). Online data processing: Comparison of Bayesian regularized particle filters. University of Brescia, Department of Economics. Working Paper n. 0704.

Doucet, A., N. de Freitas, and N. Gordon (2001). Sequential Monte Carlo Methods in Practice. New York: Springer.

Fearnhead, P. (2007). Computational methods for complex stochastic systems: a review of some alternatives to MCMC. Statistics and Computing, 18(2):151-171.

Flury, T., and N. Shephard, (2011) Bayesian inference based only on simulated likelihood: particle filter analysis of dynamic economic models. Econometric Theory 27, 933-956.

Gordon, N.J. (1997). A hybrid bootstrap filter for target tracking in clutter. IEEE Transactions on Aerospace and Electronic Systems 33: 353– 358.

Gordon, N.J., D.J. Salmond, and A.F.M. Smith (1993). Novel approach to non-linear/non-Gaussian Bayesian state estimation. IEEE-Proceedings-F 140: 107–113.

Hall, J., M.K. Pitt, and R. Kohn (2014). Bayesian inference for nonlinear structural time series models. Journal of Econometrics 179 (2), 99–111.

Liu, J., M. West (2001). Combined parameter and state estimation in simulation-based filtering. In: Doucet, A., de Freitas, N., Gordon, N. (Eds.), Sequential Monte Carlo Methods in Practice. Springer-Verlag.

Nemeth, C., P. Fearnhead (2014). Particle Metropolis adjusted Langevin algorithms for state-space models. Pre-print arXiv:1402.0694v1.

Pitt, M.K., R.S. Silva, P. Giordani, and R. Kohn (2012). On some properties of Markov chain Monte Carlo simulation methods based on the particle filter. J. Econom. 171(2), 134–151.

B. Ristic, S. Arulampalam, and N. Gordon (2004). Beyond Kalman Filters: Particle Filters for Applications. Norwood, MA: Artech House.

Table 1: Out-of-sample point forecast performance, medium VAR, comparison with KKP

variable	BCVAR	BCVARC	this study	BCVAR	BCVARC	this study
	$h=1$			$h=2$		
PAYEMS	0.830	0.838	0.313 model I	0.728	0.732	0.342 model I
CPIAUCSL	0.958	0.967	0.402 model I	0.940	0.936	0.413 model I
FEDFUNDS	1.023	0.962	0.381 model I	0.974	0.945	0.397 model I
INDPRO	0.828	0.889	0.377 model IV	0.931	0.929	0.385 model IV
UNRATE	0803	0.848	0.503 model I	0.844	0.869	0.518 model I
PPIFGS	0.970	0.993	0.281 model I	1.029	1.012	0.294 model I
GS10	0.996	1.013	0.355 model I	1.003	1.003	0.362 model I
	$h=3$			$h=6$		
PAYEMS	0.683	0.687	0.515 model I	0.747	0.738	0.588 model I
CPIAUCSL	0.982	0.978	0.482 model I	1.003	0.995	0.495 model I
FEDFUNDS	1.017	1.001	0.411 model I	0.991	0.986	0.425 model I
INDPRO	0.939	0.949	0.402 model IV	0.970	0.957	0.423 model IV
UNRATE	0.871	0.866	0.521 model I	0.939	0.946	0.535 model I
PPIFGS	1.050	1.042	0.303 model I	1.059	1.043	0.387 model I
GS10	1.046	1.032	0.488 model I	1.036	1.038	0.523 model I
	$h=9$			$h=12$		
PAYEMS	0.838	0.843	0.717 model I	0.934	0.935	0.682 model I
CPIAUCSL	0.979	0.961	0.691 model I	1.016	1.012	0.662 model I
FEDFUNDS	0.921	0.950	0.657 model I	0.991	0.996	0.682 model I
INDPRO	0.967	0.978	0.672 model IV	0.974	0.975	0.693 model IV
UNRATE	0.954	0.951	0.683 model I	0.968	0.968	0.694 model I
PPIFGS	1.055	1.042	0.623 model I	1.070	1.053	0.663 model I
GS10	1.005	1.016	0.717 model I	1.029	1.023	0.744 model I

Notes: The table reports the ratio between the MSFE of a given model and the MSFE of a benchmark AR(1) for the medium-size VAR. This is computed as

$$MSFE_{ij} = \frac{\sum_{\tau=t_1}^{t_2-h} e^2_{i,j,\tau+h}}{\sum_{\tau=t_1}^{t_2-h} e^2_{AR1,j,\tau+h}},$$

for model i, variable j (shown in columns of the table). The forecast horizon is $h \in \{1,2,3,6,9,12\}$. All forecasts are generated out-of-sample using recursive estimates of the model as in KKP, starting in 1987:07 and ending in 2014:12. The results for BCVAR and BCVARC (the second compresses the covariance matrix of the VAR as well) are taken from KKP, Table 1, page 29.

Table 2: Out-of-sample point forecast performance, Large VAR, comparison with KKP

variable	BCVAR	BCVAR$_C$	this study	BCVAR	BCVAR$_C$	this study
	$h=1$			$h=2$		
PAYEMS	0.831	0.864	0.332 model I	0.747	0.762	0.381 model I
CPIAUCSL	0.951	0.942	0.452 model VI	0.911	0.898	0.464 model VI
FEDFUNDS	0.949	0.944	0.381 model I	0.963	0.924	0.397 model I
INDPRO	0.820	0.904	0.327 model IV	0.907	0.935	0.366 model I
UNRATE	0.809	0.897	0.313 model I	0.857	0.893	0.371 model I
PPIFGS	0.967	0.991	0.414 model I	1.013	1.006	0.488 model I
GS10	0.997	1.002	0.380 model I	0.996	1.009	0.482 model I
	$h=3$			$h=6$		
PAYEMS	0.717	0.732	0.481 model I	0.764	0.783	0.490 model I
CPIAUCSL	0.923	0.926	0.592 model I	0.897	0.885	0.599 model I
FEDFUNDS	1.001	0.989	0.493 model I	0.998	0.963	0.496 model I
INDPRO	0.927	0.938	0.515 model I	0.975	0.971	0.588 model IV
UNRATE	0.906	0.930	0.516 model I	0.927	0.962	0.552 model I
PPIFGS	1.004	1.007	0.503 model VI	1.001	0.993	0.517 model VI
GS10	1.050	1.047	0.552 model I	1.031	1.022	0.583 model I
	$h=9$			$h=12$		
PAYEMS	0.858	0.863	0.582 model I	0.962	0.956	0.661 model I
CPIAUCSL	0.848	0.841	0.617 model I	0.880	0.860	0.623 model I
FEDFUNDS	0.970	1.025	0.603 model I	1.010	0.997	0.620 model I
INDPRO	0.987	0.988	0.519 model IV	0.998	1.000	0.525 model IV
UNRATE	0.979	0.987	0.615 model I	0.998	0.985	0.623 model I
PPIFGS	0.973	0.973	0.600 model I	0.989	0.981	0.633 model I
GS10	0.995	1.022	0.588 model I	1.012	1.000	0.611 model I

Notes: The table reports the ratio between the MSFE of a given model and the MSFE of a benchmark AR(1) for the Large size VAR. This is computed as

$$MSFE_{ij} = \frac{\sum_{\tau=t_1}^{t_2-h} e^2_{i,j,\tau+h}}{\sum_{\tau=t_1}^{t_2-h} e^2_{AR1,j,\tau+h}},$$

for model i, variable j (shown in columns of the table). The forecast horizon is $h \in \{1, 2, 3, 6, 9, 12\}$. All forecasts are generated out-of-sample using recursive estimates of the model as in KKP, starting in 1987:07 and ending in 2014:12. The results for BCVAR and BCVAR$_C$ (the second compresses the covariance matrix of the VAR as well) are taken from KKP, Table 2, page 30.

Table 3: Out-of-sample point forecast performance, Huge VAR, comparison with KKP

variable	BCVAR	BCVAR_C	this study	BCVAR	BCVAR_C	this study
	$h = 1$			$h = 2$		
PAYEMS	0.777	0.796	0.617 model I	0.640	0.671	0.622 model I
CPIAUCSL	0.928	0.935	0.690 model I	0.887	0.892	0.712 model I
FEDFUNDS	0.965	1.013	0.604 model I	0.962	0.892	0.610 model I
INDPRO	0.844	0.902	0.707 model IV	0.945	0.920	0.710 model IV
UNRATE	0.810	0.860	0.718 model I	0.852	0.852	0.721 model I
PPIFGS	0.974	1.012	0.723 model VI	1.013	1.019	0.733 model I
GS10	1.009	1.015	0.650 model I	1.005	1.044	0.682 model IV
	$h = 3$			$h = 6$		
PAYEMS	0.611	0.622	0.733 model I	0.668	0.706	0.787 model I
CPIAUCSL	0.912	0.904	0.731 model I	0.931	0.916	0.742 model I
FEDFUNDS	0.967	0.987	0.781 model I	0.991	0.988	0.813 model I
INDPRO	0.950	0.938	0.749 model IV	0.967	0.983	0.812 model IV
UNRATE	0.876	0.882	0.803 model I	0.924	0.943	0.822 model I
PPIFGS	1.034	1.048	0.813 model I	1.063	1.041	0.852 model I
GS10	1.049	1.064	0.815 model I	1.022	1.042	0.833 model I
	$h = 9$			$h = 12$		
PAYEMS	0.766	0.760	0.870 model I	0.848	0.866	0.877 model I
CPIAUCSL	0.895	0.885	0.863 model I	0.901	0.872	0.869 model I
FEDFUNDS	0.969	0.995	0.892 model I	1.023	1.035	0.903 model I
INDPRO	0.975	0.990	0.814 model IV	0.989	1.012	0.822 model I
UNRATE	0.951	0.957	0.833 model VI	0.979	0.989	0.841 model I
PPIFGS	1.047	1.035	0.859 model I	1.073	1.042	0.862 model I
GS10	1.006	1.024	0.843 model VI	1.013	1.006	0.849 model I

Notes: The table reports the ratio between the MSFE of a given model and the MSFE of a benchmark AR(1) for the Huge size VAR. This is computed as

$$MSFE_{ij} = \frac{\sum_{\tau=t_1}^{t_2-h} e^2_{i,j,\tau+h}}{\sum_{\tau=t_1}^{t_2-h} e^2_{AR1,j,\tau+h}},$$

for model i, variable j (shown in columns of the table). The forecast horizon is $h \in \{1, 2, 3, 6, 9, 12\}$. All forecasts are generated out-of-sample using recursive estimates of the model as in KKP, starting in 1987:07 and ending in 2014:12. The results for BCVAR and BCVAR_C (the second compresses the covariance matrix of the VAR as well) are taken from KKP, Table 3, page 31.

Table 4: Out-of-sample density forecast performance, Medium VAR, Comparison with KKP

variable	BCVAR	BCVAR$_C$	KKP best	this study	BCVAR	BCVAR$_C$	KKP best	this study
	$h = 1$				$h = 2$			
PAYEMS	0.086	0.083	0.218 BVAR	0.144 model I	0.158	0.163	0.366 BVAR	0.213 model I
CPIAUCSL	0.003	**0.156**		0.221 model I	-0.263	**-0.247**		-0.128 model I
FEDFUNDS	0.006	0.005	0.131 BVAR	0.315 model I	0.022	0.022	0.115 BVAR	0.414 model I
INDPRO	-0.063	0.028	0.046 FAVAR	0.107 model IV	0.084	**0.109**		0.122 model I
UNRATE	0.105	0.081	0.167 BVAR	0.215 model I	0.077	0.062	0.131 BVAR	0.233 model I
PPIFGS	-0.071	0.020	0.025 DFM	0.103 model I	**0.019**	-0.063		0.027 model I
GS10	-0.001	-0.007	0.015 BVAR	0.022 model I	**-0.008**	-0.016		0.014 model I
	$h = 3$				$h = 6$			
PAYEMS	0.172	0.185	0.364 BVAR	0.415 model I	0.144	0.168	0.245 BVAR	0.371 model I
CPIAUCSL	-0.095	-0.017	0.043 FAVAR	0.082 model IV	-0.220	-0.249	0.004 FAVAR	0.124 model I
FEDFUNDS	0.014	0.014	0.115 BVAR	0.312 model I	0.017	0.011	0.119 BVAR	0.303 model I
INDPRO	0.125	0.073	0.144 DFM	0.285 model IV	-0.014	0.038	0.052 FAVAR	0.181 model I
UNRATE	0.065	0.062	0.109 BVAR	0.232 model I	0.042	0.040	0.058 BVAR	0.179 model I
PPIFGS	**0.049**	-0.098		0.105 model VI	-0.172	-0.100	0.003 FAVAR	0.108 model VI
GS10	0.013	0.003	0.014 FAVAR	0.214 model I	-0.003	-0.013	0.003 DFM	0.181 model I
	$h = 9$				$h = 12$			
PAYEMS	**0.096**	0.084		0.232 model I	0.074	**0.089**		0.314 model I
CPIAUCSL	-0.083	-0.184	0.220 FAVAR	0.415 model I	-0.254	-0.312	-0.037 FAVAR	0.182 model I
FEDFUNDS	0.008	0.005	0.119 BVAR	0.239 model I	-0.006	-0.008	0.109 BVAR	0.111 model I
INDPRO	**-0.012**	-0.077		-0.01 model IV	0.128	**0.149**		0.150 model I
UNRATE	**0.048**	0.036		0.047 model I	0.024	0.020	0.033 BVAR	0.032 model I
PPIFGS	-0.070	0.060	0.106 FAVAR	0.109 model VI	-0.144	-0.108	0.120 FAVAR	0.122 model I
GS10	0.011	0.001	0.041 BVAR	0.053 model I	-0.003	-0.014	0.010 BVAR	0.019 model I

Notes: The Table reports the average log predictive likelihood (ALPL) difference between a given model and the benchmark AR(1) for the Medium size VAR. This is defined as:

$$ALPL_{ijh} = \frac{1}{t_2 - t_1 - h + 1} \sum_{\tau=t_1}^{t_2-h} \left(LPL_{i,j,\tau+h} - LPL_{AR1,j,\tau+h}\right),$$

where $LPL_{i,j,\tau+h}$ and $LPL_{AR1,j,\tau+h}$ are the log predictive likelihood values of variable j at time τ and forecast horizon h generated by model i and the benchmark AR(1) model, respectively. All density forecasts are generated out-of-sample using recursive estimates of the model as in KKP, starting in 1987:07 and ending in 2014:12. The results for BCVAR and BCVAR$_C$ (the second compresses the covariance matrix of the VAR as well) are taken from KKP, Table 4, page 32. The column "KKP best" corresponds to the highest average log predictive likelihood in KKP, Table 4, page 32. Numbers in bold indicate that the corresponding model performed best in KKP. DFM is the dynamic factor model, FAVAR is the factor-augmented Vector Autoregression and BVAR is the Bayesian Vector Autoregression of KKP.

Table 5: Out-of-sample density forecast performance, Large VAR, Comparison with KKP

variable	BCVAR	BCVAR$_C$	KKP best	this study	BCVAR	BCVAR$_C$	KKP best	this study
	$h = 1$				$h = 2$			
PAYEMS	0.076	0.063	0.254 BVAR	0.412 model I	0.140	0.137	0.406 BVAR	0.713 model I
CPIAUCSL	**0.104**	-0.026		0.232 model I	0.032	**0.182**		0.351 model I
FEDFUNDS	0.004	-0.002	0.147 BVAR	0.271 model I	0.000	0.009	0.018 FAVAR	0.187 model I
INDPRO	0.011	-0.047	0.020 DFM	0.183 model I	0.065	0.090	0.181 BVAR	0.282 model IV
UNRATE	0.089	0.051	0.170 BVAR	0.315 model I	0.083	0.058	0.150 BVAR	0.225 model I
PPIFGS	**0.096**	-0.024		0.271 model I	**0.053**	-0.048		0.185 model I
GS10	0.006	0.000	0.036 DFM	0.314 model I	**0.021**	-0.003		0.277 model I
	$h = 3$				$h = 6$			
PAYEMS	0.159	0.160	0.416 BVAR	0.817 model I	0.121	0.140	0.300 BVAR	0.616 model I
CPIAUCSL	**-0.073**	-0.075		0.256 model I	**0.087**	-0.061		0.225 model I
FEDFUNDS	0.004	0.006	0.029 BVAR	0.187 model I	0.003	0.012	0.159 BVAR	0.319 model I
INDPRO	**0.010**	-0.021		0.019 model IV	-0.063	-0.149	0.069 DFM	0.082 model VI
UNRATE	0.061	0.054	0.119 BVAR	0.222 model I	0.042	0.023	0.092 BVAR	0.108 model I
PPIFGS	**0.029**	-0.125		0.034 model I	0.007	-0.099	0.049 FAVAR	0.077 model I
GS10	**0.000**	-0.008		0.113 model I	-0.004	-0.009	0.004 BVAR	0.120 model I
	$h = 9$				$h = 12$			
PAYEMS	0.095	0.097	0.165 BVAR	0.188 model I	0.034	0.032	0.063 BVAR	0.171 model I
CPIAUCSL	-0.104	-0.158	-0.040 FAVAR	0.132 model I	**0.059**	-0.058		0.083 model I
FEDFUNDS	-0.002	-0.003	0.145 BVAR	0.203 model I	-0.003	-0.004	0.133 BVAR	0.224 model I
INDPRO	0.027	0.050	0.085 DFM	0.173 model VI	0.078	**0.180**		0.205 model I
UNRATE	0.017	0.013	0.050 BVAR	0.126 model I	0.002	0.008	0.018 DFM	0.019 model I
PPIFGS	-0.047	-0.047	-0.021 FAVAR	-0.022 model I	-0.021	-0.138	0.061 FAVAR	0.073 model I
GS10	0.001	-0.017	0.034 BVAR	0.042 model I	-0.009	-0.022	0.029 BVAR	0.039 model I

Notes: The Table reports the average log predictive likelihood (ALPL) difference between a given model and the benchmark AR(1) for the Large size VAR. This is defined as:

$$ALPL_{ijh} = \frac{1}{t_2 - t_1 - h + 1} \sum_{\tau=t_1}^{t_2-h} \left(LPL_{i,j,\tau+h} - LPL_{AR1,j,\tau+h} \right),$$

where $LPL_{i,j,\tau+h}$ and $LPL_{AR1,j,\tau+h}$ are the log predictive likelihood values of variable j at time τ and forecast horizon h generated by model i and the benchmark AR(1) model, respectively. All density forecasts are generated out-of-sample using recursive estimates of the model as in KKP, starting in 1987:07 and ending in 2014:12. The results for BCVAR and BCVAR$_C$ (the second compresses the covariance matrix of the VAR as well) are taken from KKP, Table 5, page 33. The column "KKP best" corresponds to the highest average log predictive likelihood in KKP, Table 5, page 33. Numbers in bold indicate that the corresponding model performed best in KKP. DFM is the dynamic factor model, FAVAR is the factor-augmented Vector Autoregression and BVAR is the Bayesian Vector Autoregression of KKP.

Table 6: Out-of-sample density forecast performance, Huge VAR, Comparison with KKP

variable	BCVAR	$BCVAR_C$	KKP best	this study	BCVAR	$BCVAR_C$	KKP best	this study
	$h = 1$				$h = 2$			
PAYEMS	0.104	0.102	0.302 BVAR	0.545 model I	0.196	0.196	0.471 BVAR	0.677 model I
CPIAUCSL	0.025	**0.052**		0.235 model I	**0.098**	0.095		0.331 model I
FEDFUNDS	0.014	0.010	0.291 BVAR	0.414 model I	0.013	0.014	0.247 BVAR	0.535 model I
INDPRO	**0.092**	0.026		0.178 model I	0.041	0.179	0.238 DFM	0.338 model I
UNRATE	0.095	0.079	0.157 FAVAR	0.450 model I	0.076	0.079	0.163 BVAR	0.387 model I
PPIFGS	**0.059**	-0.087		0.125 model I	-0.064	**-0.015**		0.092 model I
GS10	-0.001	0.000	0.006 BVAR	0.102 model IV	0.012	-0.001		0.117 model I
	$h = 3$				$h = 6$			
PAYEMS	0.229	0.225	0.447 BVAR	0.561 model I	0.199	0.191	0.296 BVAR	0.376 model I
CPIAUCSL	0.000	**0.121**		0.226 model I	**0.227**	0.142		0.392 model I
FEDFUNDS	0.022	0.016	0.228 BVAR	0.425 model I	0.007	0.013	0.186 BVAR	0.452 model I
INDPRO	0.052	0.043	0.065 BVAR	0.125 model VI	0.056	-0.088	0.082 DFM	0.103 model I
UNRATE	0.067	0.048	0.106 BVAR	0.108 model I	0.036	0.030	0.084 BVAR	0.092 model I
PPIFGS	**0.086**	-0.062		0.091 model I	**0.003**	-0.173		0.106 model I
GS10	0.032	0.009		0.103 model I	**0.012**	-0.005		0.044 model I
	$h = 9$				$h = 12$			
PAYEMS	0.129	0.123		0.211 model I	0.100	**0.110**		0.220 model I
CPIAUCSL	-0.032	0.059	0.212 DFM	0.267 model I	0.016	-0.171	0.060 DFM	0.093 model I
FEDFUNDS	0.014	0.010	0.275 BVAR	0.338 model I	-0.002	-0.001	0.211 BVAR	0.367 model I
INDPRO	0.081	0.050	0.110 DFM	0.287 model IV	0.021	-0.057	0.062 DFM	0.113 model IV
UNRATE	0.026	0.028	0.045 BVAR	0.071 model I	0.029	0.021	0.034 BVAR	0.085 model I
PPIFGS	0.099	0.039		0.066 model I	-0.144	-0.274	-0.130 FAVAR	0.104 model I
GS10	0.008	-0.011	0.039 BVAR	0.091 model I	-0.005	-0.021	0.034 BVAR	0.087 model I

Notes: The Table reports the average log predictive likelihood (ALPL) difference between a given model and the benchmark AR(1) for the Huge size VAR. This is defined as:

$$ALPL_{ijh} = \frac{1}{t_2 - t_1 - h + 1} \sum_{\tau=t_1}^{t_2-h} \left(LPL_{i,j,\tau+h} - LPL_{AR1,j,\tau+h} \right),$$

where $LPL_{i,j,\tau+h}$ and $LPL_{AR1,j,\tau+h}$ are the log predictive likelihood values of variable j at time τ and forecast horizon h generated by model i and the benchmark AR(1) model, respectively. All density forecasts are generated out-of-sample using recursive estimates of the model as in KKP, starting in 1987:07 and ending in 2014:12. The results for BCVAR and $BCVAR_C$ (the second compresses the covariance matrix of the VAR as well) are taken from KKP, Table 6, page 34. The column "KKP best" corresponds to the highest average log predictive likelihood in KKP, Table 6, page 34. Numbers in bold indicate that the corresponding model performed best in KKP. DFM is the dynamic factor model, FAVAR is the factor-augmented Vector Autoregression and BVAR is the Bayesian Vector Autoregression of KKP.

Table 7: Out-of-sample forecast performance: Multivariate results comparison with KKP

	best WMSFE, KKP	this study	best MVALPL, KKP	this study
	medium VAR			
$h = 1$	0.916 BCVAR	0.716 model I	0.979 BVAR	0.732 model I
$h = 2$	0.926 BCVARC	0.747 model I	1.068 BVAR	0.763 model I
$h = 3$	0.940 BVCARC	0.782 model I	1.097 BVAR	0.790 model I
$h = 6$	0.954 BCVARC	0.812 model I	1.030 BVAR	0.794 model IV
$h = 9$	0.957 BCVAR	0.822 model I	1.021 BVAR	0.815 model I
$h = 12$	0.988 FAVAR	0.874 model IV	0.927 BVAR	0.823 model VI
	Large VAR			
$h = 1$	0.906 BCVAR	0.712 model I	0.988 BVAR	0.734 model I
$h = 2$	0.919 BCVAR	0.780 model I	1.011 BCVAR	0.787 model I
$h = 3$	0.934 BCVAR	0.803 model I	1.023 BCVAR	0.793 model I
$h = 6$	0.935 BCVARC	0.833 model I	1.187 BVAR	0.801 model VI
$h = 9$	0.939 BCVAR	0.845 model I	1.198 BVAR	0.892 model I
$h = 12$	0.965 BCVARC	0.885 model IV	1.017 BVAR	0.902 model VI
	Huge VAR			
$h = 1$	0.907 BCVAR	0.702 model I	0.996 BCVAR	0.701 model I
$h = 2$	0.908 BCVARC	0.781 model I	1.139 BCVAR	0.747 model I
$h = 3$	0.916 BCVAR	0.832 model I	1.179 BCVAR	0.798 model I
$h = 6$	0.933 BCVAR	0.845 model I	1.131 BCVAR	0.813 model I
$h = 9$	0.938 BCVAR	0.871 model I	1.076 BCVAR	0.837 model VI
$h = 12$	0.968 BCVARC	0.882 model IV	1.009 BCVAR	0.840 model I

Table 8: MSFE of alternative forecasting methods

		LIT	BGR	SZ	new BCVAR	MAI
one-quarter	real GDP	1.12	1.09	0.96	0.89 model I	0.92 model VI
	GDP deflator	1.46	1.97	0.97	0.91 model I	0.95 model VI
	funds rate	1.19	1.02	0.98	0.92 model I	0.97 model VI
one-year	real GDP	1.32	0.97	0.87	0.90 model I	0.89 model IV
	GDP deflator	1.55	2.73	1.04	0.91 model I	1.03 model IV
	funds rate	1.03	0.83	0.92	0.91 model I	0.90 model V

Notes: The table reports MSFE of alternative methods relative to the MSFE of the hierarchical Bayes model in Giannone et al (2015). LIT is the method in Litterman (1980). SZ is the method in Sims and Zha (1998), BGR is the method in Banbura et al (2010). "new BCVAR" is the BCVAR with the optimal parameters proposed in this study. MAI is the best from the class of MAI models. Columns LIT, BGR and SZ are taken from Table 2 in Giannone et al (2015), p. 444.

Figure 1

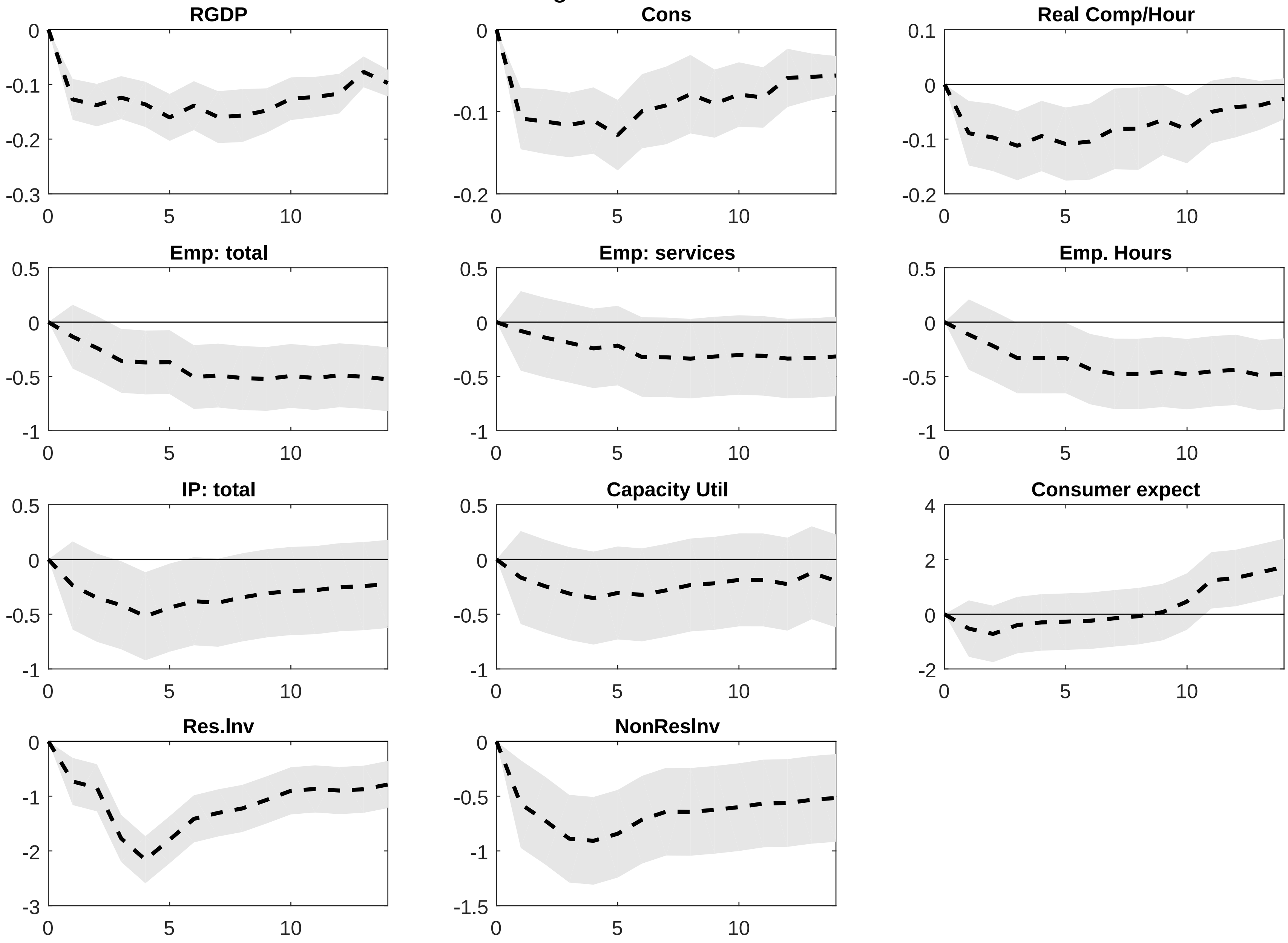

RGDP
Cons
Real Comp/Hour
Emp: total
Emp: services
Emp. Hours
IP: total
Capacity Util
Consumer expect
Res.Inv
NonResInv

Figure 2

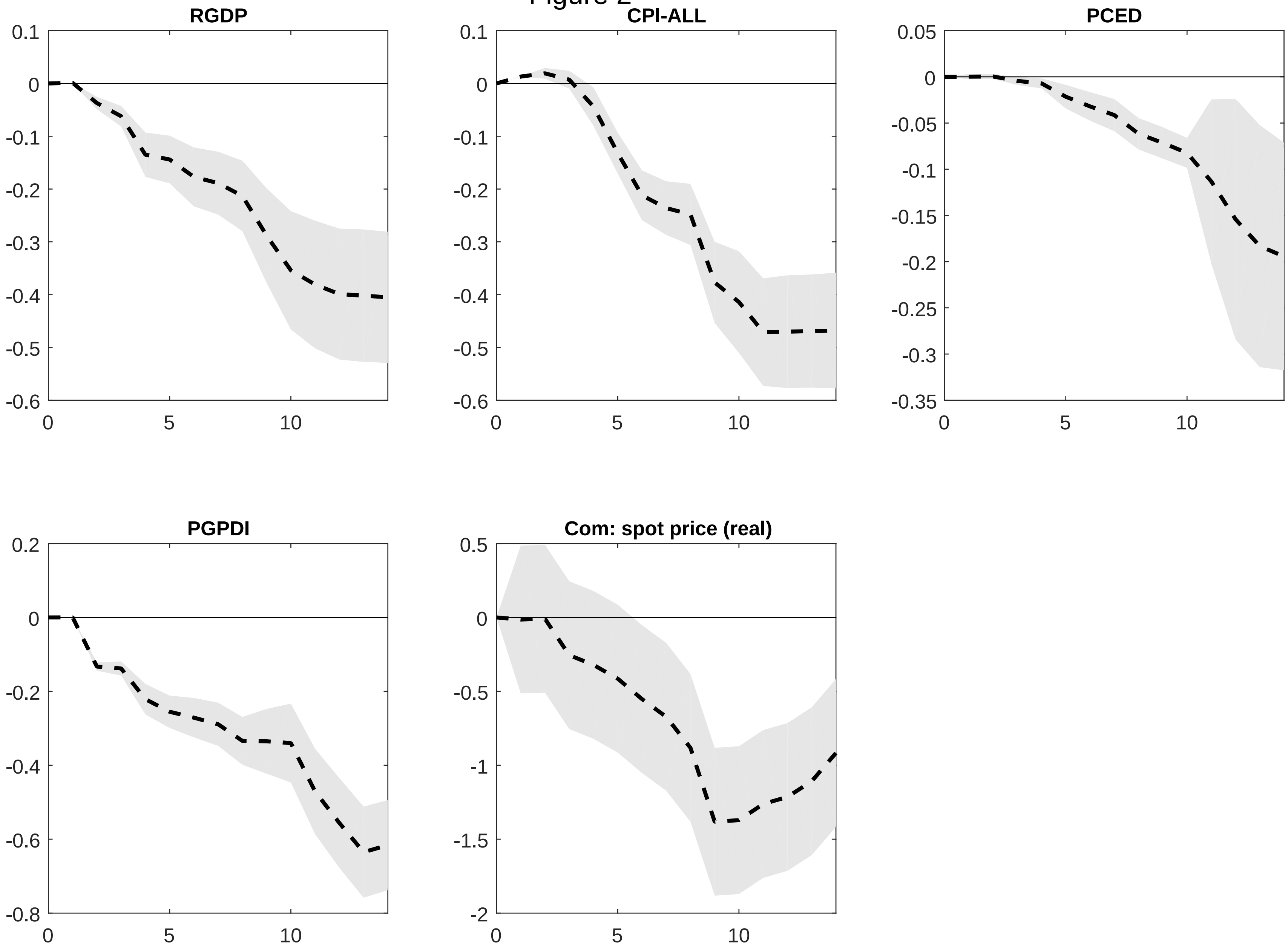
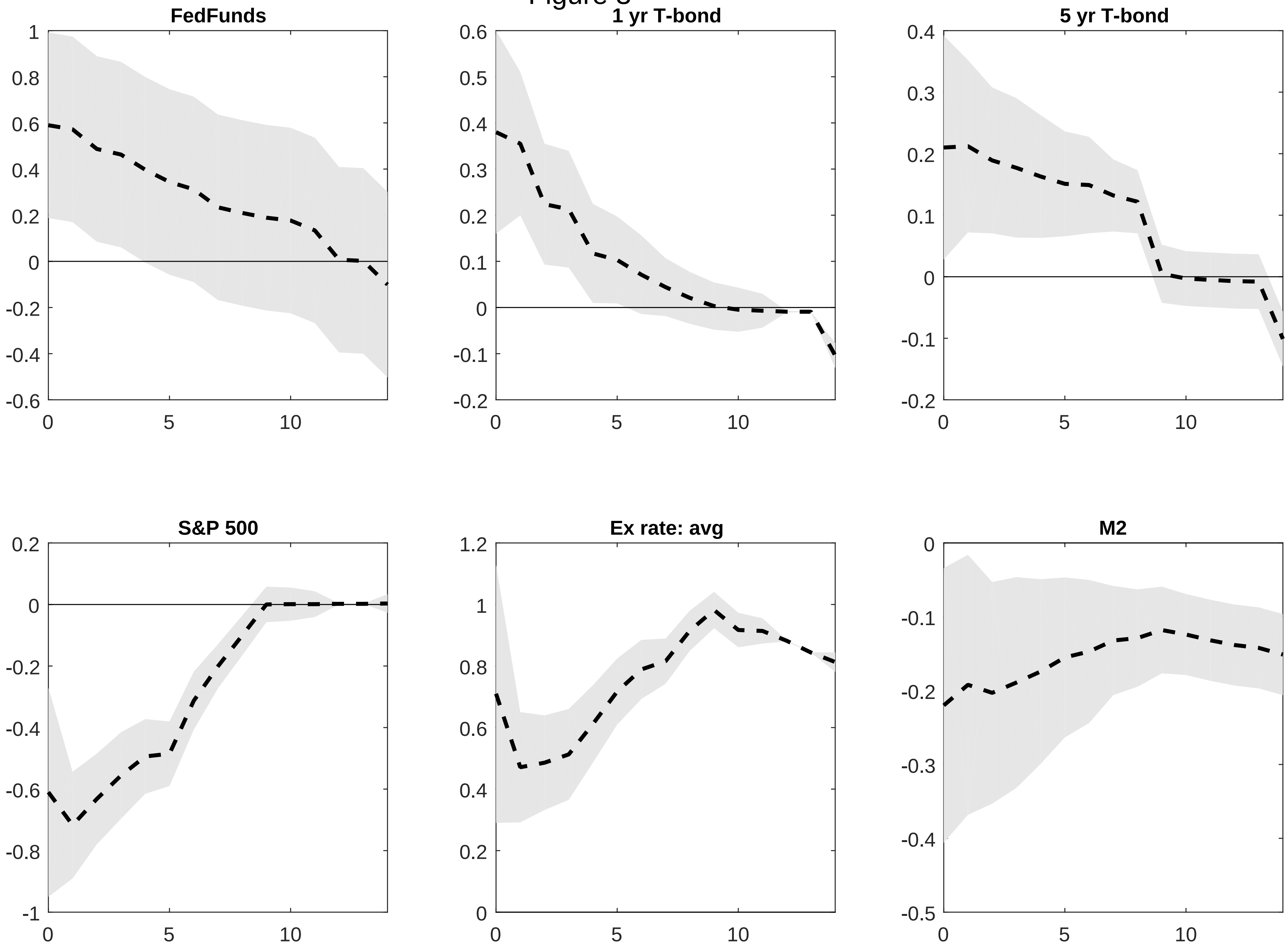

RGDP
CPI-ALL
PCED
PGPDI
Com: spot price (real)
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
0.05
-0.05
-0.1
-0.15
-0.2
-0.25
-0.3
-0.35
0.2
-0.2
-0.4
-0.6
-0.8
0.5
-0.5
-1
-1.5
-2
5
10

Figure 3


FedFunds
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
0
5
10
1 yr T-bond
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.2
0
5
10
5 yr T-bond
0.4
0.3
0.2
0.1
0
-0.1
-0.2
0
5
10
S&P 500
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
0
5
10
Ex rate: avg
1.2
1
0.8
0.6
0.4
0.2
0
0
5
10
M2
0
-0.1
-0.2
-0.3
-0.4
-0.5
0
5
10

BANK OF GREECE WORKING PAPERS

185. Adam, A., and T., Moutos, "Industry-Level Labour Demand Elasticities Across the Eurozone: Will There Be Any Gain After the Pain of Internal Devaluation?" July, 2014.

186. Tagkalakis, O.A., "Fiscal Policy, Net Exports, and the Sectoral Composition of Output in Greece", September 2014.

187. Hondroyiannis, G. and D., Papaoikonomou, "When Does it Pay To Tax? Evidence from State-Dependent Fiscal Multipliers in the Euro Area", October 2014.

188. Charalambakis, C. E., "On Corporate Financial Distress Prediction: What Can we Learn From Private Firms in a Small Open Economy?, November 2014.

189. Pagratis, S., E., Karakatsani and E. Louri, "Bank Leverage and Return on Equity Targeting: Intrinsic Procyclicality of Short-Term Choices", November 2014.

190. Evgenidis, A. and C., Siriopoulos, "What are the International Channels Through Which a US Policy Shock is Transmitted to the World Economies? Evidence from a Time Varying Favar", January 2015.

191. Louzis, D. P., and A.T., Vouldis, "Profitability in the Greek Banking System: a Dual Investigation of Net Interest and Non-Interest Income", February 2015.

192. Papaspyrou, S.T, "EMU 2.0 - Drawing Lessons From the Crisis - a New Framework For Stability and Growth", March 2014.

193. Litina, A and T, Palivos, "Corruption and Tax Evasion: Reflections on Greek Tragedy", June 2015.

194. Balfoussia, H. and H.D. Gibson, "Financial Conditions and Economic Activity: The Potential Impact of the Targeted Longer-Term Refinancing Operations (TLTROS)", July 2015.

195. Louzis, P. D., "Steady-State Priors and Bayesian Variable Selection in VAR Forecasting", July 2015.

196. Zografakis, S. and A., Sarris, "The Distributional Consequences of the Stabilization and Adjustment Policies in Greece During the Crisis, with the Use of A Multisectoral Computable General Equilibrium Model", August 2015.

197. Papageorgiou, D. and E. Vourvachaki, "The Macroeconomic Impact of Structural Reforms in Product and Labour Markets: Trade-Offs and Complementarities", October 2015.

198. Louri, H., and P. M. Migiakis, "Determinants of Euro-Area Bank Lending Margins: Financial Fragmentation and ECB Policies", October 2015.

199. Gibson, D. H, S.G. Hall, and G. S. Tavlas, "The effectiveness of the ECB's asset purchase programs of 2009 to 2012", November 2015.

200. Balfoussia, H and D. Malliaropulos, "Credit-less recoveries: the role of investment-savings imbalances", November 2015.

201. Kalyvitis, S., "Who Exports High-Quality Products? Some Empirical Regularities from Greek Exporting Firms", December 2015.

202. Papadopoulos, S., P. Stavroulias and T. Sager, "Systemic Early Warning Systems for EU15 Based on the 2008 Crisis", January 2016.

203. Papadopoulos, G., S. Papadopoulos and T. Sager, "Credit Risk Stress Testing for EU15 Banks: a Model Combination Approach", January 2016.

204. Louzis, P. D., "Macroeconomic Forecasting and Structural Changes in Steady States", March 2016.

205. Christodoulakis, N. and C. Axioglou, "Underinvestment and Unemployment: the Double Hazard in the Euro Area", April 2016.

206. Bardaka, C. I., “Structural and Cyclical Factors of Greece’s Current Account Balances: A Note”, May 2016.

207. Lothian, J. R., and G. S. Tavlas, “How Friedman and Schwartz Became Monetarists”, May 2016.

208. Balfoussia, H. and H. D. Gibson, “Firm Investment and Financial Conditions in the Euro Area: Evidence from Firm-Level Data”, June 2016.

209. Samantas, G.I., “On the Optimality of Bank Competition Policy”, July 2016.

210. Malliaropulos, D., and P.M. Migiakis, “The Re-Pricing of Sovereign Risks Following the Global Financial Crisis”, July 2016.

211. Asimakopoulos, I., P.K. Avramidis, D. Malliaropulos, N. G. Travlos, “Moral Hazard and Strategic Default: Evidence from Greek Corporate Loans”, July 2016.

212. Tavlas, S.G., “New Perspectives on the Great Depression: a Review Essay”, September 2016.

213. Papapetrou, E. and P. Tsalaporta, “Inter-Industry Wage Differentials in Greece: Rent-Sharing and Unobserved Heterogeneity Hypotheses”, November 2016.

214. Gibson, D.H, S.G. Hall, and G.S. Tavlas, “Self-Fulfilling Dynamics: the Interactions of Sovereign Spreads, Sovereign Ratings and Bank Ratings during the Euro Financial Crisis”, November 2016.

215. Baumann, U. and M. Vasardani, “The Slowdown in US Productivity Growth - What Explains it and Will it Persist?”, November 2016.