Estimating point and density forecasts for the US Economy with a Factor-Augmented Vector Autoregressive DSGE model *

Stelios Bekiros [a,b†]

[a]European University Institute (EUI)

[b]Rimini Centre for Economic Analysis (RCEA)

Alessia Paccagnini[‡]

Università degli Studi di Milano-Bicocca, Department of Economics,
Piazza Ateneo Nuovo 1, 20126 Milano, Italy

Abstract

Although policymakers and practitioners are particularly interested in DSGE models, these are typically too stylized to be applied directly to the data and often yield weak prediction results. Very recently, hybrid DSGE models have become popular for dealing with some of the model misspecifications. Major advances in estimation methodology could allow these models to outperform well-known time series models and effectively deal with more complex real-world problems as richer sources of data become available. In this study we introduce a Bayesian approach to estimate a novel Factor Augmented DSGE model that extends the model of Consolo *et al.* [Consolo, A., Favero, C.A., and Paccagnini, A., 2009. On the Statistical Identification of DSGE Models. Journal of Econometrics, 150, 99-115]. We perform a comparative predictive evaluation of point and density forecasts for many different specifications of estimated DSGE models and various classes of VAR models, using datasets from the US economy including real-time data. Simple and hybrid DSGE models are implemented, such as DSGE-VAR and tested against standard, Bayesian and Factor Augmented VARs. The results can be useful for macro-forecasting and monetary policy analysis.

JEL Classification: C32, C11, C15, C53, D58

Keywords: Density Forecasting, Marginal data density, DSGE-FAVAR, Real-Time data

*The first author is supported by the Marie Curie Fellowship (FP7-PEOPLE-2011-CIG, N° 303854) under the 7th European Community Framework Programme. The second author acknowledges financial support from "Dote ricercatori": FSE, Regione Lombardia. The usual disclaimers apply.

†[a] Department of Economics, Via della Piazzuola; 43, I-50133, Florence, Italy; Tel.: +39 055 4685 916; Fax: +39 055 4685 902; Tel.: +39 055 4685916; Fax: +39 055 4685 902; E-mail address: stelios.bekiros@eui.eu ; [b]RCEA, Via Patara, 3, 47900, Rimini, Italy;Tel.: +39 0541 434 142; Fax: +39 0541 55 431.

‡Tel.: +39-026-4483-046; Fax: +39-026-4483-085; email: alessia.paccagnini@unimib.it

1 Introduction

The dynamic stochastic general equilibrium models appear to be particularly suited for evaluating the consequences of alternative macroeconomic policies, as shown in the works of Smets and Wouters (2003, 2004), Del Negro and Schorfheide (2004), Adolfson *et al.* (2008) and Christiano *et al.* (2005). However, the calibrated dynamic stochastic general equilibrium (DSGE) models face many important challenges such as the fragility of parameter estimates, statistical fit and the weak reliability of policy predictions as reported in Stock and Watson (2001), Ireland (2004) and Schorfheide (2010). In recent years Bayesian estimation of DSGE models has become popular for many reasons, mainly because it is a system-based estimation approach that offers the advantage of incorporating assumptions about the parameters based on economic theory. These assumptions can reduce weak identification issues. The popularity of the Bayesian approach is also explained by the increasing computational power available to estimate and evaluate medium- to large-scale DSGE models using Markov chain Monte Carlo (MCMC) simulators.

Recently, increasing efforts have been undertaken to use DSGE models for forecasting. DSGE models were not considered as forecasting tools until the works of Smets and Wouters (2003, 2004) on the predictability of DSGE models compared to alternative non-structural models. In the macro-econometric literature, hybrid or mixture DSGE models have become popular for dealing with some of the model misspecifications as well as the trade-off between theoretical coherence and empirical fit (Schorfheide, 2010). They are categorized in additive hybrid models and hierarchical hybrid models. The hybrid models provide a complete analysis of the data law of motion and better capture the dynamic properties of the DSGE models. In the recent literature, different attempts of hybrid models have been introduced for solving, estimating and forecasting with DSGEs. Sargent (1989) and Altug (1989) proposed augmenting a DSGE model with measurement error terms that follow a first order autoregressive process, known as the DSGE-AR approach. Ireland (2004) proposed a method that is similar to the DSGE-AR, but imposing no restriction on the measurement errors, assuming that residuals follow a first-order vector autoregression (DSGE-AR à l'Ireland). A different approach called DSGE-VAR was proposed by Del Negro and Schorfheide (2004) and was based on the works DeJong *et al.* (1996) and Ingram and Whiteman (1994). The main idea behind the DSGE-VAR is the use of the VAR representation as an econometric tool for empirical validation, combining prior information derived from the DSGE model in estimation. However, it has several problems. One of the main problems in finding a statistical representation for the data by using a VAR, is overfitting due to the inclusion of too many lags and too many variables, some of which may be insignificant. The problem of overfitting results in multicollinearity and loss of degrees of freedom, leading to inefficient estimates and large out-of-sample forecasting errors. It is possible to overcome this problem by using the well-known Minnesota or Inverse-Wishart priors. The use of these priors has been proposed to shrink the parameters space and thus overcome the curse of dimensionality. Following this idea in combining the DSGE model information and the VAR representation, two alternative econometric tools have been also introduced: the DSGE-FAVAR (Consolo *et al.*, 2009) and the Augmented VAR-DSGE model (Fernández-de-Córdoba and Torres, 2010). The main idea behind the Factor Augmented DSGE (DSGE-FAVAR) is the use of factors to improve the statistical identification in validating the models. Consequently, the VAR representation is replaced by a FAVAR model as the statistical benchmark.

In this paper, we conduct an exhaustive empirical exercise that includes the comparison of the out-of-sample predictive performance of estimated DSGE models with that of standard VARs, Bayesian VARs and Factor Augmented VARs estimated on the same data set for the US economy. We use point and density forecasts to evaluate predictability. In particular, the Root Mean Squared Forecast Error and the Modified Diebold - Mariano test are applied for the optimal lag specifications of the models, and more importantly the Marginal Data Density as a measure of model fit, which arises naturally in the computation of posterior model odds. In addition, to evaluate the quality of density forecasts we use the Probability Integral Transform (PIT), which was recently popularized in forecasting exercises involving DSGE-based models by Herbst and Schorfheide (2012) and Del Negro and Schorfheide (2012). It is employed in order to check whether forecasts provide a realistic description of actual uncertainty as inferred by the distribution of observed data. We focus on many different specifications of the DSGE models, i.e., the simple DSGE, the DSGE-VAR and specifically on the Factor Augmented DSGE (DSGE-

FAVAR) model with emphasis on Bayesian estimation. We use time series data from 1960:Q4 to 2010:Q4 for the real GDP, the harmonized Consumer Price Index and the nominal short-term federal funds interest rate and we produce their forecasts for the out-of-sample testing period 1997:Q1-2010:Q4. The motivation comes from a group of recent papers that compares the forecasting performance of DSGE against VAR models, e.g., Smets and Wouters (2004), Ireland (2004), Del Negro and Schorfheide (2004), Del Negro *et al.* (2007), Adolfson *et al.* (2008), Christoffel *et al.* (2008), Rubaszek and Skrzypczynski (2008), Ghent (2009), Kolosa *et al.* (2009), Consolo *et al.* (2009), Fernandez-de-Cordoba and Torres (2010) and Edge *et al.* (2010) amongst others. Finally, we repeat the forecasting comparison using Real-Time data to further test and evaluate the comparative predictability of the "winner" models from the conventional dataset.

The remainder of this paper is organized as follows. Section 2 describes the standard and Bayesian VARs as well as the Factor Augmented VAR model. In section 3 the simple DSGE model is analyzed, and the hybrid DSGE-VAR and DSGE-FAVAR models are described in detail. In section 4 the data are described and the comparative empirical results of the point and density forecasting evaluation are illustrated and analyzed. A specific section presents the results for a real-time dataset. Finally, section 5 concludes.

2 VAR Models

2.1 Classical VAR

As suggested by Sims (1980), the standard unrestricted VAR, has the following compact format

$$\mathbf{Y}_t = \mathbf{X}_t \mathbf{\Phi} + \mathbf{U} \tag{1}$$

where $\mathbf{Y}_t$ is a $(T \times n)$ matrix with rows Y_t', and $\mathbf{X}$ is a $(T \times k)$ matrix ($k = 1 + np, p =$number of lags) with rows $X_t' = [1, Y_{t-1}', ..., Y_{t-p}']$. $\mathbf{U}$ is a $(T \times n)$ matrix with rows u_t', $\mathbf{\Phi}$ is a $(k \times n) = [\Phi_0, \Phi_1, ..., \Phi_p]'$, while the one-step ahead forecast errors u_t have a multivariate $N(0, \Sigma_u)$ conditional on past observations of Y.

2.2 Bayesian VAR

The curse of dimensionality is one of the main problem is using a VAR model when we need to estimate many parameters. Instead of eliminating longer lags, the BVAR imposes restrictions on these coefficients by assuming that they are more likely to be near zero than the coefficients on shorter lags (Litterman, 1981; Doan *et al.,* 1984; Todd, 1984; Litterman, 1986; Spencer, 1993). Usually, the restrictions are imposed by specifying normal prior distributions with zero means and small standard deviations for all coefficients, with a decreasing standard deviation as the lags increase. The only exception is the coefficient on a variable's first lag that has a mean of unity. Litterman (1981) used a diffuse prior for the constant. Formally speaking, these prior means can be written as follows

$$\Phi_i \sim N(1, \sigma^2_{\Phi_i}) \text{ and } \Phi_j \sim N(0, \sigma^2_{\Phi_j}), \tag{2}$$

where Φ_i denotes the coefficients associated with the lagged dependent variables in each equation of the VAR, while Φ_j represents any other coefficient. The prior variances $\sigma^2_{\Phi_i}$ and $\sigma^2_{\Phi_j}$ specify the uncertainty of the prior means, $\Phi_i = 1$ and $\Phi_j = 0$, respectively. In this study, we impose their prior mean on the first own lag for variables in growth rate, such as a white noise setting (Del Negro and Schorfheide 2004; Adolfson *et al.* 2007; Banbura *et al.* 2010). Instead, for level variables, we use the classical Minnesota prior (Del Negro and Schorfheide, 2004). The specification of the standard deviation of the distribution of the prior imposed on variable j in equation i at lag m, for all i, j and m, denoted by $S(i, j, m)$, is specified as follows

$$S(i, j, m) = [w \times g(m) \times F(i, j)] \frac{\hat{\sigma}_i}{\hat{\sigma}_j}, \tag{3}$$

where

$$F(i,j) = \begin{cases} 1 \text{ if } i = j \\ k_{ij} \text{ otherwise, } 0 \leq k_{ij} \leq 1 \end{cases} \tag{4}$$

is the tightness of variable j in equation i relative to variable i and by increasing the interaction, i.e. it is possible for the value of k_{ij} to loosen the prior (Dua and Ray, 1995). The ratio $\frac{\hat{\sigma}_i}{\hat{\sigma}_j}$ consists of estimated standard errors of the univariate autoregression, for variables i and j. This ratio scales the variables to account for differences in the units of measurement, without taking into account the magnitudes of the variables. The term w measures the standard deviation on the first lag, and also indicates the overall tightness. A decrease in the value of w results in a tighter prior. The function $g(m) = m^{-d}$, $d > 0$ is the measurement of the tightness on lag m relative to lag 1, and is assumed to have a harmonic shape with a decay of d, which tightens the prior on increasing lags. Following the standard Minnesota prior settings, we choose the overall tightness (w) to be equal to 0.3, while the lag decay (d) is 1 and the interaction parameter (k_{ij}) is set equal to 0.5.

Furthermore, as defined in Sims and Zha (1998) under the Normal-Wishart prior the mean of the posterior distribution, i.e., the coefficient estimator (given Σ) is Normal under the form

$$\widehat{\Phi} = (\overline{H}^{-1} + \mathbf{X}'\mathbf{X})^{-1}(\overline{H}^{-1}\overline{\Phi} + \mathbf{X}'\mathbf{Y})$$

where $\overline{\Phi}$ is the prior mean of the coefficient matrix and $\overline{H}$ is the diagonal positive-definite matrix. The corresponding estimator of the error covariance is

$$\widehat{\Sigma} = T^{-1}(\mathbf{Y}'\mathbf{Y} - \widehat{\Phi}'(\overline{H}^{-1} + \mathbf{X}'\mathbf{X})^{-1}\widehat{\Phi} + \overline{\Phi}'\overline{H}^{-1}\overline{\Phi} + \overline{S})$$

where $\overline{S}$ is the diagonal scale matrix in the prior Inverse-Wishart distribution for Σ. Both $\overline{S}$ and $\overline{H}$ have elements along the diagonal scaled by a parameter (λ_0) which controlls the overall tightness of the prior on Σ[1]. In our empirical implementation, we choose the overall tightness (w) to be 0.1, the lag decay $d = 1$ and the control parameter of the prior tightness λ_0 equal to 1.

2.3 Factor Augmented VAR

A recent strand in the econometric literature mainly by Stock and Watson (2002), Forni and Reichlin (1996, 1998) and Forni *et al.* (1999, 2000) has shown that very large macroeconomic datasets can be properly modelled using dynamic factor models, where the factors can be considered as an "exhaustive summary of the information" in the data. The rationale underlying dynamic factor models is that the behavior of several variables is driven by few common forces, the factors, plus idiosyncratic shocks. Hence, the factors-approach can be useful in alleviating the omitted variable problem in empirical analysis using traditional small-scale models. Bernanke and Boivin (2003) and Bernanke *et al.* (2005) utilized factors in the estimation of VAR to generate a more general specification. Chudik and Pesaran (2011) illustrated how a VAR augmented by factors could help in keeping the number of estimated parameters under control without loosing relevant information.

Let $\mathbf{X}_t$ denote an $N \times 1$ vector of economic time series and $\mathbf{Y}_t$ a vector of $M \times 1$ observable macroeconomic variables which are a subset of $\mathbf{X}_t$. In this context, most of the information contained in $\mathbf{X}_t$ is captured by $\mathbf{F}_t$, a $k \times 1$ vector of unobserved factors. The factors are interpreted as an addition to the observed variables, as common forces driving the dynamics of the economy. The relation between the "informational" time series $\mathbf{X}_t$, the observed variables $\mathbf{Y}_t$ and the factors $\mathbf{F}_t$ is represented by the following dynamic factor model:

$$\mathbf{X}_t = \mathbf{\Lambda}^f\mathbf{F} + \mathbf{\Lambda}^y\mathbf{Y}_t + e_t \tag{5}$$

where $\mathbf{\Lambda}^f$ is a $N \times k$ matrix of factor loadings, $\mathbf{\Lambda}^y$ is a $N \times M$ matrix of coefficients that bridge the observable $\mathbf{Y}_t$ and the macroeconomic dataset, and e_t is the vector of $N \times 1$ error terms. These terms are mean zero, normal distributed, and uncorrelated with a small cross-correlation. In fact, the estimator

[1] More details can be found in Sims and Zha (1998) and Robertson and Tallman (1999).

allows for some cross-correlation in e_t that must vanish as N goes to infinity. This representation nests also models where $\mathbf{X}_t$ depends on lagged values of the factors (Stock and Watson, 2002).

For the estimation of the FAVAR model equation (5), we follow the two-step principal components approach proposed by Bernanke *et al.* (2005). In the first step factors are obtained from the observation equation by imposing the orthogonality restriction $\mathbf{F}'\mathbf{F}/T = \mathbf{I}$.This implies that $\hat{\mathbf{F}} = \sqrt{T}\hat{\mathbf{G}}$, where $\hat{\mathbf{G}}$ are the eigenvectors corresponding to the K largest eigenvalues of $\mathbf{X}\mathbf{X}^{'}$, sorted in descending order. Stock and Watson (2002) showed that the factors can be consistently estimated by the first r principal components of $\mathbf{X}$, even in the presence of moderate changes in the loading matrix $\mathbf{\Lambda}$. For this result to hold it is important that the estimated number of factors, k, is larger or equal than the true number r. Bai and Ng (2000) proposed a set of selection criteria to choose k that are generalizations of the BIC and AIC criteria. In the second step, we estimate the FAVAR equation replacing $\mathbf{F}_t$ by $\hat{\mathbf{F}}_t$. Following Bernanke *et al.* (2005), firstly we remove $\mathbf{Y}_t$ from the space covered by the principal components, and then we partition the matrix $\mathbf{X}_t$ in two categories of information variables: slow-moving (which do not respond contemporaneously to unanticipated changes in monetary policy) and fast-moving (which respond contemporaneously to monetary shocks) [2]. We proceed to extracting two factors from slow variables and one factor from fast variables and we call them respectively "slow factors" and "fast factor"[3]. It is worth noting that the factors are not uniquely identified, but this is not a problem in our context because we will not attempt a structural interpretation of the estimated factors. Finally, having determined the number of factors, we specify a Factor Augmented VAR by considering only one-lag of the factors according to BIC criterion. The potential identification of the macroeconomic shocks can be performed according to Bernanke *et al.* (2005) using the Cholesky decomposition.

3 DSGE Models

Simple DSGE models with forward-looking features are usually referred to as a benchmarks in the literature. The model we estimate is the small-scale DSGE implemented by Del Negro and Schorfheide (2004, 2012), i.e., without capital accumulation, wage stickiness, price indexation and habit formation. The latter are considered typical features of medium-scale models such as introduced by Smets and Wouters (2007). Following the idea of combining the DSGE model information and the VAR representation, among other models that have been proposed in the literature, in this study we use the DSGE-VAR and DSGE-FAVAR hybrid models.

In this DSGE setup the economy is made up of four components. The first component is the representative household with habit persistent preferences. This household maximizes an additively separable utility function which is separable into consumption, real money balances and hours worked over an infinite lifetime. The household gains utility from consumption relative to the level of technology, real balances of money, and disutility from hours worked. The household earns interest from holding government bonds and earns real profits from the firms. Moreover, the representative household pays lump-sum taxes to the government. The second component is a perfectly competitive, representative final goods producer which is assumed to use a continuum of intermediate goods as inputs, and the prices for these inputs are given. The producers of these intermediate goods are monopolistic firms which use labour as the only input. The production technology is the same for all the monopolistic firms. Nominal rigidities are introduced in terms of price adjustment costs for the monopolistic firms. Each firm maximizes its profits over an infinite lifetime by choosing its labour input and its price. The third component is the government which spends in each period a fraction of the total output, which fluctuates exogenously. The government issues bonds and levies lump-sum taxes, which are the main part of its budget constraint. The last component is the monetary authority, which follows a Taylor rule regarding the inflation target

[2] In a recent paper, Boivin *et al.* (2009) impose the constraint that $\mathbf{Y}_t$ is one of the common components in the first step, guaranteeing that the estimated latent factors $\hat{\mathbf{F}}_t$ recover the common dynamics which are not captured by $\mathbf{Y}_t$. The authors, comparing the two methodologies, concluded that the results are similar.

[3] As suggested by Bai and Ng (2000) information criteria can be used to determine the number of factors but, as they are not so decisive, one can limit the number of factors to three (two slows and one fast) to strike a balance between the variance of the original series explained by the principal components and the difference in the parameterization of the VAR and the FAVAR.

and the output gap. There are three economic shocks: an exogenous monetary policy shock (in the monetary policy rule), and two autoregressive processes, AR(1), which model government spending and technology shocks. To solve the model, optimality conditions are derived for the maximization problems. After linearization around the steady-state, the economy is described by the following system of equations

$$\tilde{x}_t = E_t[\tilde{x}_{t+1}] - \frac{1}{\tau}(\tilde{R}_t - E_t[\tilde{\pi}_{t+1}]) + (1-\rho_g)\tilde{g}_t + \rho_Z \frac{1}{\tau}\tilde{z}_t \tag{6}$$

$$\tilde{\pi}_t = \beta E_t[\tilde{\pi}_{t+1}] + \kappa[\tilde{x}_t - \tilde{g}_t] \tag{7}$$

$$\tilde{R}_t = \rho_R \tilde{R}_{t-1} + (1-\rho_R)(\psi_1 \tilde{\pi}_t + \psi_2 \tilde{x}_t) + \epsilon_{R,t} \tag{8}$$

$$\tilde{g}_t = \rho_g \tilde{g}_{t-1} + \epsilon_{g,t} \tag{9}$$

$$\tilde{z}_t = \rho_z \tilde{z}_{t-1} + \epsilon_{z,t}, \tag{10}$$

where x is the detrended output (divided by the non-stationary technology process), π is the gross inflation rate, and R is the gross nominal interest rate. The tilde denotes percentage deviations from a steady state or, in the case of output, from a trend path (King, 2000; Woodford, 2003). The model can be solved by applying the algorithm proposed by Sims (2002). Define the vector of variables $\tilde{Z}_t = \left(\tilde{x}_t, \tilde{\pi}_t, \tilde{R}_t, \tilde{g}_t, \tilde{z}_t, E_t\tilde{x}_{t+1}, E_t\tilde{\pi}_{t+1}\right)$ and the vector of shocks as $\epsilon_t = (\epsilon_{R,t}, \epsilon_{g,t}, \epsilon_{z,t})$. Therefore the previous set of equations, (6) - (10), can be recasted into a set of matrices $(\Gamma_0, \Gamma_1, C, \Psi, \Pi)$ accordingly to the definition of the vectors $\tilde{Z}_t$ and ϵ_t

$$\Gamma_0 \tilde{Z}_t = C + \Gamma_1 \tilde{Z}_{t-1} + \Psi \epsilon_t + \Pi \eta_t \tag{11}$$

where η_{t+1}, such that $E_t\eta_{t+1} \equiv E_t\left(y_{t+1} - E_t y_{t+1}\right) = 0$, is the expectations error. As a solution to (11), we obtain the following transition equation as a policy function

$$\tilde{Z}_t = T\left(\theta\right)\tilde{Z}_{t-1} + R\left(\theta\right)\epsilon_t \tag{12}$$

and in order to provide the mapping between the observable data and those computed as deviations from the steady state of the model we set the following measurement equations as in Del Negro and Schorfheide (2004):

$$\begin{aligned} \Delta \ln x_t &= \ln\gamma + \Delta\tilde{x}_t + \tilde{z}_t \\ \Delta \ln P_t &= \ln\pi^* + \tilde{\pi}_t \\ \ln R_t^a &= 4\left[(\ln r^* + \ln\pi^*) + \tilde{R}_t\right] \end{aligned} \tag{13}$$

which can be also casted into matrices as

$$Y_t = \Lambda_0\left(\theta\right) + \Lambda_1\left(\theta\right)\tilde{Z}_t + v_t \tag{14}$$

where $Y_t = \left(\Delta\ln x_t, \Delta\ln P_t, \ln R_t\right)'$, $v_t = 0$ and Λ_0 and Λ_1 are defined accordingly. For completeness, we write the matrices T, R, Λ_0 and Λ_1 as a function of the structural parameters in the model, $\theta = \left(\ln\gamma, \ln\pi^*, \ln r^*, \kappa, \tau, \psi_1, \psi_2, \rho_R, \rho_g, \rho_Z, \sigma_R, \sigma_g, \sigma_Z\right)'$. Such a formulation derives from the rational expectations solution. The evolution of the variables of interest, Y_t, is therefore determined by (12) and (14) which impose a set of restrictions across the parameters on the moving average (MA) representation. Given that the MA representation can be very closely approximated by a finite order VAR representation, Del Negro and Schorfheide (2004) propose to evaluate the DSGE model by assessing the validity of the restrictions imposed by such a model with respect to an unrestricted VAR representation. The choice of the variables to be included in the VAR is however completely driven by those entering in the DSGE model regardless of the statistical goodness of the unrestricted VAR.

3.1 Estimating linearized DSGE Models

Several econometric procedures have been proposed to parameterize and evaluate DSGE models. Kydland and Prescott (1982) use calibration, Christiano and Eichenbaum (1992) consider the generalized method of moments (GMM) estimation of equilibrium relationships, while Rotemberg and Woodford (1997) and Christiano *et al.* (2005) use the minimum distance estimation based on the discrepancy among VAR and DSGE model impulse response functions. Moreover the full-information likelihood-based estimation is considered by Altug (1989), McGrattan (1994), Leeper and Sims (1994) and Kim (2000). In recent years, Bayesian estimation became very popular. According to An and Schorfheide (2007) there are essentially three main characteristics. Firstly, the Bayesian estimation is system-based and fits the solved DSGE model to a vector of aggregate time series, as opposed to the GMM which is based on equilibrium relationships, such as the Euler equation for the consumption or the monetary policy rule. Secondly, it is based on the likelihood function generated by the DSGE model rather than the discrepancy between DSGE responses and VAR impulse responses. Thirdly, prior distributions can be used to incorporate additional information into the parameter estimation.

Priors distributions are important to estimate DSGE models. According to An and Schorfheide (2007) priors might downweigh regions of the parameter space that are at odds with observations which are not contained in the estimation sample. Priors could add curvature to a likelihood function that is (nearly) flat for some parameters, given a strong influence to the shape of the posterior distribution. Table 1 lists the prior distributions for the structural parameters of the DSGE model which are adopted from Del Negro and Schorfheide (2004).

Table 1: Prior Distributions for the DSGE model parameters

Name	Range	Density	Starting value	Mean	Standard deviation
$\ln\gamma$	$\mathbb{R}$	Normal	0.500	0.500	0.250
$\ln\pi^*$	$\mathbb{R}$	Normal	1.000	1.000	0.500
$\ln r^*$	$\mathbb{R}^+$	Gamma	0.500	0.500	0.250
κ	$\mathbb{R}^+$	Gamma	0.040	0.030	0.150
τ	$\mathbb{R}^+$	Gamma	3.000	3.000	0.500
ψ_1	$\mathbb{R}^+$	Gamma	1.500	1.500	0.250
ψ_2	$\mathbb{R}^+$	Gamma	0.300	0.125	0.100
ρ_R	$[0,1)$	Beta	0.400	0.500	0.200
ρ_G	$[0,1)$	Beta	0.800	0.800	0.100
ρ_Z	$[0,1)$	Beta	0.200	0.200	0.100
σ_R	$\mathbb{R}^+$	Inv.Gamma	0.100	0.100	0.139
σ_G	$\mathbb{R}^+$	Inv.Gamma	0.300	0.350	0.323
σ_Z	$\mathbb{R}^+$	Inv.Gamma	0.400	0.875	0.430

Note: The model parameters $\ln\gamma$, $ln\pi^*$, $\ln r^*$, σ_R, σ_g, and σ_z are scaled by 100 to convert them into percentages. The Inverse Gamma priors are of the form $p(\sigma|\nu, s) \propto \sigma^{-\nu-1}e^{-\nu s^2/2\sigma^2}$, where v=4 and s equals 0.2, 0.5, and 0.7, respectively. Approximately 1.5% of the prior mass lies in the indeterminacy region of the parameter space. The prior is truncated to restrict it to the determinacy region of the DSGE model, to avoid multiple equilibria typical in rational expectations models .

In the Bayesian framework, the likelihood function is reweighted by a prior density. The prior is useful to add information which is contained in the estimation sample. Since priors are always subject to revisions, the shift from prior to posterior distribution can be considered as an indicator of the different sources of information. If the likelihood function peaks at a value that is at odds with the information that has been used to construct the prior distribution, then the marginal data density (MDD) of the DSGE model is defined as:

$$p(Y) = \int L(\theta|Y)p(\theta)d\theta$$

The marginal data density is the integral of the likelihood ($L(\theta|Y)$) taken according to the prior distribution ($p(\theta)$), that is the weighted average of likelihood where the weights are given by priors. The MDD can be used to compare different models M_i, $p(Y|M_i)$.We can rewrite the log-marginal data density as:

$$\begin{aligned} \ln(p(Y|M) &= \sum_{t=1}^{T} \ln p(y_t|Y^{t-1}, M) = \\ &= \sum_{t=1}^{T} \ln \left[\int p(y_t|Y^{t-1}, \theta, M) p(\theta|Y^{t-1}, M) d\theta \right] \end{aligned}$$

where $\ln(p(Y|M)$ can be interpreted as a predictive score (Good, 1952) and the model comparison based on posterior odds captures the relative one-step-ahead predictive performance. To compute the MDD, we consider the Geweke (1999) modified harmonic mean estimator. Harmonic mean estimators are based on the identity:

$$\frac{1}{p(Y)} = \int \frac{f(\theta)}{L(\theta|Y)p(\theta)} p(\theta|Y) d\theta$$

where $f(\theta)$ has the property that $\int f(\theta)d\theta = 1$ (Gelfand and Dey, 1994). Conditional on the choice of $f(\theta)$, an estimator is[4]:

$$\widehat{p}_G(Y) = \left[\frac{1}{n_{sim}} \sum_{s=1}^{n_{sim}} \frac{f(\theta^{(s)})}{L(\theta^{(s)}|Y)p(\theta^{(s)})} \right]^{-1} \tag{15}$$

where $\theta^{(s)}$ is drawn from the posterior $p(\theta|Y)$. For a numerical approximation efficient, $f(\theta)$ should be chosen so that the summands are of equal magnitude. Geweke (1999) proposed to use the density of a truncated multivariate normal distribution:

$$\begin{aligned} f(\theta) &= \tau^{-1}(2\pi)^{-\frac{d}{2}} |V_\theta|^{-\frac{1}{2}} \exp\left[-0.5(\theta - \overline{\theta})' V_\theta^{-1} (\theta - \overline{\theta})\right] \\ &\times I \left\{ (\theta - \overline{\theta})' V_\theta^{-1} (\theta - \overline{\theta}) \leq F_{\chi_d^2}^{-1}(\tau\tau) \right\} \end{aligned}$$

In the above $\overline{\theta}$ and V_θ are the posterior mean and covariance matrix computed from the output of the posterior simulator, d is the dimension of the parameter vector, $F_{\chi_d^2}$ is the cumulative density function of a χ^2 random variable with d degrees of freedom, and $\tau \in (0, 1)$. We set $\tau = 0.90$ which provides the most accurate computation as in Geweke (1999). Moreover, we compute the log-marginal data density with $\tau = 0.95$ and $\tau = 0.99$ but we found that the marginal likelihood is not sensitive to the value of τ. As shown in Schorfheide (2000) the estimated marginal likelihood does not so serially depend on τ, whilst the discrepancy of eq (15) across truncation levels τ and simulation runs, is less than 0.60. If the posterior of θ is in fact normal then the summands in eq. (15) are approximately constant.

3.2 DSGE-VAR

Building on the work by Ingram and Whiteman (1994), the DSGE-VAR approach of Del Negro and Schorfheide (2004) was designed to improve forecasting and monetary policy analysis with VARs. Del Negro-Schorfheide's (2004) approach is to use the DSGE model to build prior distributions for the VAR. This approach is employed both for the small and the medium scale DSGE model. Basically, the estimation initializes with an unrestricted VAR of order p

$$\mathbf{Y}_t = \mathbf{\Phi}_0 + \mathbf{\Phi}_1 \mathbf{Y}_{t-1} + ... + \mathbf{\Phi}_p \mathbf{Y}_{t-p} + \mathbf{u}_t \tag{16}$$

[4]As reported in An and Schorfheide (2007) researchers tend to restrict the parameter space to the subspace in which the linearized DSGE model has a unique rational expectations solutions. We follow the adjustment proposed by Del Negro and Schorfheide (2004) and An and Schorfheide (2007), with a small percentage (around 1.5%-2%) to the indeterminacy region.

In compact format:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\Phi} + \mathbf{U} \tag{17}$$

$\mathbf{Y}$ is a $(T \times n)$ matrix with rows Y_t', $\mathbf{X}$ is a $(T \times k)$ matrix ($k = 1 + np, p$ =number of lags) with rows $X_t' = [1, Y_{t-1}', ..., Y_{t-p}']$, $\mathbf{U}$ is a $(T \times n)$ matrix with rows u_t' and Φ is a $(k \times n) = [\Phi_0, \Phi_1, ..., \Phi_p]'$.The one-step-ahead forecast errors u_t have a multivariate normal distribution $N(0, \Sigma_u)$ conditional on past observations of Y. The log-likelihood function of the data is a function of Φ and Σ_u

$$L(\mathbf{Y}|\Phi, \Sigma_u) \propto |\Sigma_u|^{-\frac{T}{2}} \exp\left\{-\frac{1}{2} tr\left[\Sigma_u^{-1}\left(\mathbf{Y}'\mathbf{Y} - \boldsymbol{\Phi}'\mathbf{X}'\mathbf{Y} - \mathbf{Y}'\mathbf{X}\boldsymbol{\Phi} + \boldsymbol{\Phi}'\mathbf{X}'\mathbf{X}\boldsymbol{\Phi}\right)\right]\right\} \tag{18}$$

The prior distribution for the VAR parameters proposed by Del Negro and Schorfheide (2004) is based on the statistical representation of the DSGE model given by a VAR approximation. Let Γ_{xx}^*, Γ_{yy}^*, Γ_{xy}^* and Γ_{yx}^* be the theoretical second-order moments of the variables Y and X implied by the DSGE model, where

$$\begin{gathered} \Phi^*(\theta) = \Gamma_{xx}^{*-1}(\theta)\,\Gamma_{xy}^*(\theta) \\ \Sigma^*(\theta) = \Gamma_{yy}^*(\theta) - \Gamma_{yx}^*(\theta)\,\Gamma_{xx}^{*-1}(\theta)\,\Gamma_{xy}^*(\theta) \end{gathered} \tag{19}$$

The moments are the dummy observation priors used in the mixture model. These vectors can be interpreted as the probability limits of the coefficients in a VAR estimated on the artificial observations generated by the DSGE model. Conditional on the vector of structural parameters in the DSGE model θ, the prior distributions for the VAR parameters $p(\Phi, \Sigma_u|\theta)$ are of the Inverted-Wishart (IW) and Normal forms

$$\begin{gathered} \Sigma_u\,|\theta \sim IW\left((\lambda T \Sigma_u^*(\theta), \lambda T - k, n\right) \\ \boldsymbol{\Phi}\,|\Sigma_u, \theta \sim N\left(\boldsymbol{\Phi}^*(\theta), \Sigma_u \otimes (\lambda T \Gamma_{XX}(\theta))^{-1}\right) \end{gathered} \tag{20}$$

where the parameter λ controls the degree of model misspecification with respect to the VAR: for small values of λ the discrepancy between the VAR and the DSGE-VAR is large and a sizeable distance is generated between the unrestricted VAR and DSGE estimators. Large values of λ correspond to small model misspecification and for $\lambda = \infty$ beliefs about DSGE misspecification degenerate to a point mass at zero. Bayesian estimation could be interpreted as estimation based on a sample in which data are augmented by a hypothetical sample in which observations are generated by the DSGE model, the so-called dummy prior observations (Theil and Goldberg, 1961; Ingram and Whiteman, 1994). Within this framework λ determines the length of the hypothetical sample. The posterior distributions of the VAR parameters are also of the Inverted-Wishart and Normal forms. Given the prior distribution, posterior distributions are derived by the Bayes theorem

$$\Sigma_u\,|\theta, \mathbf{Y} \sim IW\left((\lambda+1)\,T\hat{\Sigma}_{u,b}(\theta), (\lambda+1)\,T - k, n\right) \tag{21}$$

$$\boldsymbol{\Phi}\,|\Sigma_u, \theta, \mathbf{Y} \sim N\left(\hat{\boldsymbol{\Phi}}_b(\theta), \Sigma_u \otimes [\lambda T \Gamma_{XX}(\theta) + \mathbf{X}'\mathbf{X}]^{-1}\right) \tag{22}$$

$$\hat{\boldsymbol{\Phi}}_b(\theta) = (\lambda T \Gamma_{XX}(\theta) + \mathbf{X}'\mathbf{X})^{-1}(\lambda T \Gamma_{XY}(\theta) + \mathbf{X}'\mathbf{Y}) \tag{23}$$

$$\hat{\Sigma}_{u,b}(\theta) = \frac{1}{(\lambda+1)\,T}\left[(\lambda T \Gamma_{YY}(\theta) + \mathbf{Y}'\mathbf{Y}) - (\lambda T \Gamma_{XY}(\theta) + \mathbf{X}'\mathbf{Y})\,\hat{\boldsymbol{\Phi}}_b(\theta)\right] \tag{24}$$

where the matrices $\hat{\boldsymbol{\Phi}}_b(\theta)$ and $\hat{\Sigma}_{u,b}(\theta)$ have the interpretation of maximum likelihood estimates of the VAR parameters based on the combined sample of actual observations and artificial observations generated by the DSGE. Equations (21) and (22) show that the smaller λ is, the closer the estimates are to the OLS estimates of an unrestricted VAR. Instead, the higher λ is, the closer the VAR estimates will be tilted towards the parameters in the VAR approximation of the DSGE model ($\hat{\boldsymbol{\Phi}}_b(\theta)$ and $\hat{\Sigma}_{u,b}(\theta)$). In order to obtain a non-degenerate prior density (20), which is a necessary condition for the existence of a well-defined Inverse-Wishart distribution and for computing meaningful marginal likelihoods, λ has to be greater than λ_{MIN}

$$\begin{aligned}
\lambda_{MIN} &\geq \frac{n+k}{T}; k = 1 + p \times n \\
p &= \text{lags} \\
n &= \text{endogenous variables.}
\end{aligned}$$

Hence, the optimal lambda must be greater than or equal to the minimum lambda $\left(\hat{\lambda} \geq \lambda_{MIN}\right)$.

Essentially, the DSGE-VAR tool allows the econometrician to draw posterior inferences about the DSGE model parameters θ. Del Negro and Schorfheide (2004) explain that the posterior estimate of θ has the interpretation of a minimum-distance estimator, where the discrepancy between the OLS estimates of the unrestricted VAR parameters and the VAR representation of the DSGE model is a sort of distance function. The estimated posterior of parameter vector θ depends on the hyperparameter λ. When $\lambda \rightarrow 0$, in the posterior the parameters are not informative, so the DSGE model is of no use in explaining the data. Unfortunately, the posteriors (22) and (21) do not have a closed form and we need a numerical method to solve the problem. The posterior simulator implemented by Del Negro and Schorfheide (2004) is the Markov Chain Monte Carlo Method and the algorithm used is the Random Walk - Metropolis Hastings (RW-MH) acceptance method. The optimal λ is given by maximizing the log of the marginal data density

$$\hat{\lambda} = \underset{\lambda \geqslant \lambda_{MIN}}{\arg\max} \ln p(\mathbf{Y}|\lambda)$$

According to the optimal lambda $\left(\hat{\lambda}\right)$, a corresponding optimal mixture model is chosen. This hybrid model is called DSGE-VAR$\left(\hat{\lambda}\right)$ and $\hat{\lambda}$ is the weight of the priors. It can also be interpreted as the restriction of the theoretical model on the actual data.

3.3 DSGE-FAVAR

Based on Bernanke *et al.* (2005), a FAVAR benchmark for the evaluation of a DSGE model will include a vector of observable variables and a small vector of unobserved factors extracted from a large data-set of macroeconomic time series, that capture additional economic information relevant to model the dynamics of the observables. In this study we implement the DSGE-FAVAR model of Consolo *et al.* (2009). The statistical representation has the following specification:

$$\begin{aligned}
\begin{pmatrix} \mathbf{Y}_t \\ \mathbf{F}_t \end{pmatrix} &= \begin{bmatrix} \mathbf{\Phi}_{11}(L) & \mathbf{\Phi}_{12}(L) \\ \mathbf{\Phi}_{21}(L) & \mathbf{\Phi}_{22}(L) \end{bmatrix} \begin{pmatrix} \mathbf{Y}_{t-1} \\ \mathbf{F}_{t-1} \end{pmatrix} + \begin{pmatrix} \mathbf{u}_t^Y \\ \mathbf{u}_t^F \end{pmatrix} \qquad (25) \\
\mathbf{Y}_t &= (\Delta \ln x_t, \Delta \ln P_t, \ln R_t) \\
\mathbf{F}_t &= \left(F_{1t}^s, F_{2t}^s, F_{3t}^f\right)
\end{aligned}$$

where $\mathbf{Y}_t$ are the observable variables included in the simple DSGE model and $\mathbf{F}_t$ is a small vector of unobserved factors relevant to modelling the dynamics of $\mathbf{Y}_t$ (F_{1t}^s, F_{2t}^s are the two slow factors and F_{3t}^f is the fast factor). The system reduces to the standard VAR when $\Phi_{12}(L) = 0$. Importantly, and differently from Boivin and Giannoni (2006), this FAVAR is not interpreted as the reduced form of a DSGE model at hand. In fact, in this case the restrictions implied by the DSGE model on a general FAVAR are very difficult to trace and model evaluation becomes even more difficult to implement. A very tightly parameterized theory model can have a very highly parameterized reduced form if one is prepared to accept that the relevant theoretical concepts in the model are a combination of many macroeconomic and financial variables. The DSGE-FAVAR is implemented in the same way as the DSGE-VAR.

4 Empirical results

We use quarterly data of the US economy from 1960:Q4 to 2010:Q4. The out-of sample period spans 1997:Q1 to 2010:Q4. The data for real output growth comes from the Bureau of Economic Analysis as Gross Domestic Product (GDP), while Consumer price index (CPI) data (seasonally adjusted, 1982-1984=100) are derived from the Bureau of Labor Statistics. Both series are taken in first difference logarithmic transformation. The interest rate series (FR rate) are constructed as in Clarida, Galì and Gertler (2000), namely for each quarter the interest rate is computed as the average federal funds rate during the first month of the quarter, including business days only. These three time series also represent the three equations of the DSGE model[5]. The complete dataset is used to extract factors for FAVAR and DSGE-FAVAR models. In order to construct the FAVAR we extract factors from a balanced panel of 109 monthly and quarterly macroeconomic and financial time series, following the dataset built by Stock and Watson (2002). The dataset involves several measures of industrial production, interest rates, various price indices, employment and other important macroeconomic and also financial variables. In this set-up, the number of informational time series N is large (larger than time period T) and must be greater than the number of factors and observed variables in the FAVAR system ($k + M \ll N$). In the panel data used, there are some variables in monthly format, which are transformed into a quarterly data using the first month in the quarter observation to be consistent with the transformed variables in the estimated DSGE model. All series have been transformed to induce stationarity. The series are taken as levels or transformed into logarithms, first or second difference (in level or logarithms) according to series characteristics [6]. As suggested by Bai and Ng (2000), we use information criteria to determine the number of factors but, as they are not so decisive, we limit the number of factors to three to strike a balance between the variation in the original series explained by the principal components and the difference in the parameterization of the VAR and the FAVAR. It is also worth noting that the factors are not uniquely identified, but this is not a problem in our context because we will not attempt a structural interpretation of the estimated factors.

We compare the out-of-sample forecasting performance of VAR models including BVAR (with Minnesota and Inverse-Wishart priors) and FAVAR, and of the DSGE class including DSGE-VAR, DSGE-FAVAR. We use the Root Mean Squared Forecast Error (RMSE) for the optimal lag specifications (one to four) selected by the Schwartz Bayesian information criterion (SIC). More importantly, we compare the log of the marginal data densities (MDD). Based on the MDD a forecasting exercise is provided using a rolling procedure for h-steps-ahead. The GDP, CPI and FR rate forecasts are estimated for the out-of-sample testing period 1997:Q1 - 2010:Q4. The forecasting investigation for the quarterly US data is performed over the one-, two-, three- and four-quarter-ahead horizon with a rolling estimation sample, based on the works of Marcellino (2004) and Brüggemann *et al.* (2008) for datasets of quarterly frequency. In particular, the models are re-estimated each quarter over the forecast horizon to update the estimate of the coefficients, before producing the quarter-ahead forecasts. Finally, in order to evaluate the models' forecast accuracy, we use the Modified Diebold and Mariano test (MDM) proposed by Harvey, Leybourne and Newbold (1997) as well as the Marginal Data Density (MDD). The use of the MDM is required as the simple Diebold-Mariano could be seriously over-sized when the prediction horizon increases. We provide point and density forecasts for the out-of-sample period. The latter became recently popular in DSGE forecasting analysis by Herbst and Schorfheide (2012) and Del Negro and Schorfheide (2012). We evaluate the quality of density forecasts with respect to the Probability Integral Transform (PIT) measure developed by Dawid (1984) and introduced in economic applications by Diebold *et al.* (1998).

[5] The specification of the DSGE model is characterized by the absence of habit formation in investment and consumption and of price indexation. In our model, we focus on three key variables i.e., GDP, CPI and FF rate. It is a small-scale New Keynesian model as described by Del Negro and Schorfheide (2012) without capital accumulation, wage stickiness, price indexation and habit formation. Also the deviation of output is not from the flexible economy output. These features are typical of medium-scale models as introduced by Smets and Wouters (2007). In spite of its simplicity in similar studies it is used as benchmark, particularly in evaluating the comparative forecasting performance of various models. It is considered a good compromise vis-a-vis the adequate description of the economy and attainable predictability. Moreover, features as habit formations and price indexation could be more important in order to "match" investment and consumption data, rather than GDP growth and the CPI.

[6] The Appendix contains a detailed description of all series and their corresponding transformations.

Firstly, we report estimation results for the log of Marginal Data Density (MDD). In particular, following Del Negro and Schorfheide (2006) we adopt the MDD as a measure of model fit, which arises naturally in the computation of posterior model odds. The prior distribution for the DSGE model parameters (θ), which are similar to the priors used by Del Negro and Schorfheide (2004), were already illustrated in Table 1. This MDD measure has two dimensions: goodness of in-sample fit on the one hand and a penalty for model complexity or degrees of freedom on the other hand. The DSGE-VAR and the DSGE-FAVAR are estimated with a different number of lags on the sample 1960:Q4 -1996:Q4. From 1997:Q1, we start our forecasting evaluation as implemented in Herbst and Schorfheide (2012). The parameter λ is chosen from a grid which is unbounded from above. In our empirical exercise, the log of the MDD is computed over a discrete interval, $\ln p(Y|\lambda, M)$. The minimum value, $\lambda_{\min} = \frac{n+k}{T}$, is model dependent and is related to the existence of a well-defined Inverse-Wishart distribution. For completeness, it is worth mentioning that $\lambda = 0$ refers to the VAR and the FAVAR model with no prior and it is not possible to compute the marginal likelihood in this particular case. Therefore, we can show the log of MDD for any value of λ larger than $\lambda_{\min}$. Importantly, $\lambda_{\min}$ depends on the degrees of freedom in the VAR or FAVAR and therefore, given estimation on the same number of available observations, $\lambda_{\min}$ for a DSGE-FAVAR will always be larger than $\lambda_{\min}$ for a DSGE-VAR[7].

Table 2: Optimal lambda for the DSGE-VAR and DGSE-FAVAR calculated with Markov Chain Monte Carlo and Metropolis Hastings method

	λ_{MIN}	$\hat{\lambda}$	$\hat{\lambda} - \lambda_{MIN}$	$\frac{\hat{\lambda}-\lambda_{MIN}}{\lambda_{MIN}}$	$\ln p(Y\|\hat{\lambda}, M)$	Bayes Factor vs M_1
DSGE-VAR(1)	0.05	0.12	0.07	1.40	-563.440	$exp[37.26]$
DSGE-VAR(2)	0.08	0.15	0.07	0.88	-536.929	$exp[10.75]$
DSGE-VAR(3) (M_1)	0.1	0.2	0.1	1	-526.180	1
DSGE-VAR(4)	0.12	0.25	0.13	1.08	-530.421	$exp[4.24]$
	λ_{MIN}	$\hat{\lambda}$	$\hat{\lambda} - \lambda_{MIN}$	$\frac{\hat{\lambda}-\lambda_{MIN}}{\lambda_{MIN}}$	$\ln p(Y\|\hat{\lambda}, M)$	Bayes Factor vs M_2
DSGE-FAVAR(1)	0.08	0.14	0.06	0.75	-558.021	$exp[27.37]$
DSGE-FAVAR(2)	0.1	0.20	0.1	1	-538.052	$exp[7.40]$
DSGE-FAVAR(3) (M_2)	0.12	0.25	0.13	1.08	-530.649	1
DSGE-FAVAR(4)	0.14	0.35	0.21	1.5	-534.145	$exp[3.50]$

Table 2 shows the main results related to the DSGE-VAR implemented using a different number of lags (from 1 up to 4). Each minimum λ (λ_{MIN}) is given by the features of the model (number of observations, number of endogenous variables, number of lags), and the optimal lambda ($\widehat{\lambda}$) is calculated using the Markov Chain Monte Carlo with Metropolis Hastings acceptance method (with 110,000 replications, we discard the first 10,000 ones). $\ln p(Y|M)$ is the log-MDD of the DSGE model specifications computed based on Geweke's (1999) modified harmonic mean estimator. The Bayes factor (ratio of posterior odds to prior odds), as in An and Schorfheide (2007) helps us to understand the improvement of the log-MDD of a specific model. We compare different models against the benchmark model (M) maximizing the MDD. According to Table 2, we select the DSGE-VAR with 3 lags for the full sample 1960-1996. We repeat our exercise for the DSGE-FAVAR. We select one lag for the factors and we implement - as in case of the DSGE-VAR, - the DSGE-FAVAR with a different number lags from 1 to 4. As Table 2 shows, the DSGE-FAVAR with 3 lags is chosen. In Table 3, we compare the logarithm of the MDD of the hybrid models, DSGE-VAR and DSGE-FAVAR against the DSGE, the Bayesian VAR, the VAR and the Factor Augmented VAR. In each model category, the MDD is maximized for 3 lags.

[7] For the DSGE-VAR over the sample 1960:Q4-1996:Q4, the lambda grid is given by $\Lambda = \left\{ \begin{array}{c} 0,\ 0.05,\ 0.08,\ 0.10,\ 0.12,\ 0.15,\ 0.20,\ 0.25, \\ 0.30,\ 0.35,\ 0.40\ 0.5,\ 0.6,\ 0.7,\ 0.8,\ 0.9,\ 1,\ 10,\ 100 \end{array} \right\}$.
For the DSGE-FAVAR over the sample 1960:Q4-1996:Q4, the lambda grid is given by $\Lambda = \left\{ \begin{array}{c} 0,\ 0.08,\ 0.1,\ 0.12,\ 0.14{,}0.15,\ 0.2,\ 0.25, \\ 0.3,\ 0{,}35,\ 0.4,\ 0.5,\ 0.6,\ 0.7,\ 0.8,\ 0.9,\ 1,\ 10,\ 100 \end{array} \right\}$.
In both lambda intervals, we consider the λ_{MIN} across lags from 1 to 4.

Table 3: Log of the Marginal Data Density and Bayes Factor for the sample 1960:Q4-1996:Q4

	$lnp(Y\|M)$
DSGE	*-588.435*
DSGE-VAR(3)	*-526.180*
DSGE-FAVAR(3)	*-530.649*
MP-BVAR(1)	-570.244
MP-BVAR(2)	-534.305
MP-BVAR(3)	*-524.858*
MP-BVAR(4)	-535.723
IW-BVAR(1)	-546.132
IW-BVAR(2)	-544.219
IW-BVAR(3)	*-535.522*
IW-BVAR(4)	-574.969
VAR(1)	-568.056
VAR(2)	-541.083
VAR(3)	*-534.063*
VAR(4)	-534.063
FAVAR(1)	-560.592
FAVAR(2)	-542.964
FAVAR(3)	*-540.622*
FAVAR(4)	-548.641

4.1 Point Forecast evaluation

Table 4 reports the RMSE for all models and variables. An exhaustive exercise was conducted with one to four lags based on the Schwartz Bayesian information criterion (SIC). The results provide evidence that in general three lags is the optimal number for all models[8]. In general, the results from RMSE are in accordance with those from the MDD estimation. In particular, for the GDP series the simple DSGE model provides the lowest RMSE in all forecasting horizons except for the two-steps-ahead where the FAVAR outperforms all other models. The DSGE-FAVAR is the next best performer whilst the VAR, MP-BVAR and DSGE-VAR models present similar predictive performance and on average they generate the highest forecast errors. The IW-BVAR is the worst forecaster. Next, in case of the CPI variable, the DSGE-FAVAR model outperforms the other models for almost all steps-ahead. However, the simple DGSE model and the FAVAR both present a similar predictive performance in terms of RMSE. In particular the DSGE outranks both in two- and three-quarters-ahead. The VAR model seems slightly better than MP-BVAR, whilst the IW-BVAR provides with relatively high scores for the RMSE especially for one-, two- and four-quarters-ahead. The results for the FR rate series provide further evidence of the FAVAR and DSGE-FAVAR superiority. Specifically, when comparing the RMSE scores of all model classes, DSGE-FAVAR is the best performer for one- and three-steps-ahead and FAVAR for the other two forecasting horizons. However, in case of two-quarters-ahead the difference is clearly marginal. The next lowest error is produced by the DSGE model except for the longest horizon where the DGSE-VAR provides with a better score. Overall, the BVAR and the VAR models produce similar scores and they both underperform relatively to the other models.

[8] In case of BVAR the optimal number of lags was two, but the SIC score was very close to the one corresponding to three lags.

Table 4: Root Mean Square Forecast Error (RMSE) for GDP, CPI and FF rate

	VAR	MP-BVAR	IW-BVAR	FAVAR	DSGE	DSGE-VAR	DSGE-FAVAR
GDP							
1	0.721	0.724	1.032	0.696	0.690	0.720	0.711
2	0.734	0.731	1.119	0.697	0.733	0.722	0.714
3	0.742	0.737	1.334	0.691	0.685	0.728	0.701
4	0.748	0.737	1.940	0.691	0.688	0.732	0.705
CPI							
1	0.953	0.940	1.101	0.835	0.841	0.955	0.833
2	0.929	0.927	1.222	0.863	0.849	0.920	0.863
3	0.897	0.906	0.943	0.858	0.790	0.890	0.853
4	0.890	0.908	1.853	0.841	0.954	0.886	0.840
FF rate							
1	4.176	4.122	5.026	2.771	3.397	4.135	2.738
2	4.096	4.115	4.913	2.828	3.570	3.985	2.829
3	4.055	4.155	3.820	2.903	3.108	3.951	2.491
4	4.150	4.268	5.412	2.505	4.314	4.043	2.946

Next, the MDM pairwise test is employed in order to evaluate the comparative point-forecast accuracy. The MDM test employes the Newey-West estimator (1987, 1994) of the asymptotic variance matrix, to correct for the autocorrelation of the forecast errors. The plain Diebold-Mariano assumes for h-steps-ahead forecasts all autocorrelations of order equal or greater to h for the squared forecast errors difference are zero, and consequently the empirical autocorrelation of the errors tends to be of higher order than *h-1*. The results are reported in Tables 5, 6 and 7. The MDM test is based on the squared prediction errors. The MDM test statistics for GDP indicate that aside from the IW-BVAR pairs, almost none of the models consistently outperforms any of the other for all quarter-ahead forecasts, namely their pairwise forecast comparison shows no statistically significant difference at the 5% and 1% level. Only in case of all IW-BVAR pairs vs. the other models in all horizons, and of the VAR vs. DSGE-VAR and MP-BVAR vs. DSGE-VAR (for three-steps ahead), differential predictability is significant at the 1% or 5% level. Interestingly the IW-BVAR model shows a distinctive differential predictability albeit as the worst performer when the MDM results are combined with the RMSE scores. On the contrary for the CPI series, the DSGE-FAVAR model in any pair shows a distinctively significant predictability at 1% in all step-ahead forecasts. In fact, most models for all forecast horizons appear to have a significant pairwise predictability. The FAVAR and the simple DSGE also generate similar MDM results, yet not in all forecasting horizons and not vis-a-vis all pairs as the DSGE-FAVAR. Some exceptions include the pairs VAR vs. MP-BVAR, VAR vs. DSGE-VAR and MP-BVAR vs. DSGE for two- and four-steps-ahead. In accordance with the MDD and RMSE results, it is evident that in case of CPI the DSGE-FAVAR set-up outperforms the other models, whilst overall the simple DSGE model and the FAVAR both present an equally consistent predictive performance. Finally, in case of FR rate the MDM results lead to a variant assessment of differential predictability, albeit the majority of cases produce a statistically significant MDM score. The two longest forecasting horizons provide with insignificant results for the IW-BVAR in many evaluations. While it appears that no particular model consistently and comparatively outperforms any of the other, yet the DSGE-FAVAR and FAVAR present significant scores at the 1% level in almost all pairwise comparisons, except when compared against IW-BVAR for three- and four-quarters-ahead. While in terms of RMSE, the DSGE-FAVAR, FAVAR and the small-scale DSGE provide with comparable results, the examination of differential predictability favors the first two. The DSGE-FAVAR is superior when examined against the VAR, MP-BVAR, IW-BVAR and DSGE-VAR models and only when compared to FAVAR and DSGE in most horizons, it shows weak or no differential predictability. Overall, the combined investigation of the MDD, RMSE and MDM results is indicative of a consistent outranking classification amongst the investigated models in favor of the DSGE-FAVAR and FAVAR, followed by the simple DSGE.

Table 5: Pairwise forecast comparison for the GDP with the Modified Diebold-Mariano test

GDP	PERIODS			
	1	2	3	4
VAR vs MP-BVAR	0.111	0.095	0.108	0.033
VAR vs IW-BVAR	0.000	0.000	0.000	0.004
VAR vs FAVAR	0.085	0.096	0.102	0.078
VAR vs DSGE	0.133	0.460	0.074	0.090
VAR vs DSGE-VAR	0.303	0.036	0.022	0.030
VAR vs DSGE-FAVAR	0.248	0.167	0.126	0.078
MP-BVAR vs FAVAR	0.077	0.101	0.117	0.098
MP-BVAR vs DSGE	0.124	0.417	0.084	0.111
MP-BVAR vs DSGE-VAR	0.118	0.039	0.015	0.099
MP-BVAR vs DSGE-FAVAR	0.213	0.190	0.142	0.104
IW-BVAR vs FAVAR	0.001	0.000	0.000	0.004
IW-BVAR vs DSGE	0.000	0.002	0.004	0.004
IW-BVAR vs DSGE-VAR	0.003	0.000	0.000	0.004
IW-BVAR vs DSGE-FAVAR	0.000	0.000	0.001	0.004
IW-BVAR vs MP-BVAR	0.000	0.000	0.000	0.004
FAVAR vs DSGE	0.273	0.049	0.052	0.283
FAVAR vs DSGE-VAR	0.077	0.129	0.144	0.102
FAVAR vs DSGE-FAVAR	0.013	0.056	0.118	0.141
DSGE vs DSGE-VAR	0.129	0.120	0.102	0.116
DSGE vs DSGE-FAVAR	0.081	0.052	0.045	0.165
DSGE-VAR vs DSGE-FAVAR	0.251	0.298	0.185	0.117

Notes: The Modified Diebold-Mariano (1995) test is based on squared prediction errors (p-values).

Table 6: Pairwise forecast comparison for the CPI with the Modified Diebold-Mariano test

CPI	PERIODS			
	1	2	3	4
VAR vs MP-BVAR	0.000	0.257	0.000	0.001
VAR vs IW-BVAR	0.082	0.000	0.373	0.000
VAR vs FAVAR	0.002	0.001	0.035	0.024
VAR vs DSGE	0.000	0.006	0.001	0.029
VAR vs DSGE-VAR	0.024	0.009	0.003	0.179
VAR vs DSGE-FAVAR	0.000	0.001	0.026	0.036
MP-BVAR vs FAVAR	0.000	0.001	0.012	0.005
MP-BVAR vs DSGE	0.005	0.000	0.001	0.079
MP-BVAR vs DSGE-VAR	0.001	0.003	0.000	0.000
MP-BVAR vs DSGE-FAVAR	0.000	0.001	0.010	0.008
IW-BVAR vs FAVAR	0.007	0.000	0.265	0.000
IW-BVAR vs DSGE	0.010	0.001	0.112	0.000
IW-BVAR vs DSGE-VAR	0.085	0.001	0.355	0.000
IW-BVAR vs DSGE-FAVAR	0.008	0.000	0.252	0.000
IW-BVAR vs MP-BVAR	0.067	0.000	0.397	0.000
FAVAR vs DSGE	0.010	0.363	0.001	0.017
FAVAR vs DSGE-VAR	0.000	0.002	0.059	0.030
FAVAR vs DSGE-FAVAR	0.001	0.001	0.004	0.099
DSGE vs DSGE-VAR	0.003	0.000	0.002	0.018
DSGE vs DSGE-FAVAR	0.036	0.179	0.002	0.021
DSGE-VAR vs DSGE-FAVAR	0.000	0.004	0.043	0.046

Notes: As in Table 5

Table 7: Pairwise forecast comparison for the FF rate with the Modified Diebold-Mariano test

FF Rate	PERIODS			
	1	2	3	4
VAR vs MP-BVAR	0.000	0.077	0.002	0.004
VAR vs IW-BVAR	0.001	0.000	0.124	0.195
VAR vs FAVAR	0.000	0.001	0.001	0.003
VAR vs DSGE	0.000	0.000	0.001	0.070
VAR vs DSGE-VAR	0.001	0.000	0.000	0.001
VAR vs DSGE-FAVAR	0.001	0.002	0.000	0.002
MP-BVAR vs FAVAR	0.000	0.000	0.001	0.003
MP-BVAR vs DSGE	0.000	0.004	0.000	0.350
MP-BVAR vs DSGE-VAR	0.000	0.000	0.000	0.001
MP-BVAR vs DSGE-FAVAR	0.002	0.000	0.000	0.001
IW-BVAR vs FAVAR	0.000	0.002	0.365	0.032
IW-BVAR vs DSGE	0.001	0.000	0.386	0.229
IW-BVAR vs DSGE-VAR	0.000	0.000	0.143	0.177
IW-BVAR vs DSGE-FAVAR	0.001	0.000	0.467	0.055
IW-BVAR vs MP-BVAR	0.001	0.001	0.107	0.218
FAVAR vs DSGE	0.000	0.000	0.004	0.005
FAVAR vs DSGE-VAR	0.002	0.000	0.001	0.004
FAVAR vs DSGE-FAVAR	0.003	0.001	0.008	0.015
DSGE vs DSGE-VAR	0.001	0.000	0.001	0.015
DSGE vs DSGE-FAVAR	0.001	0.002	0.001	0.003
DSGE-VAR vs DSGE-FAVAR	0.000	0.000	0.001	0.002

Notes: As in Table 5

4.2 Density Forecast evaluation

The forecast evaluation is completed with the assessment of density forecasts. The goal is to infer on whether the analyzed forecasts provide a realistic description of actual uncertainty. As opposed to point forecasts where their accuracy can be compared by computing RMSEs (or any other symmetric loss function), the evaluation of density forecasts is less straightforward as the true density is never observed. Moreover, the question of how accurate density forecasts are is posed. For the Bayesian estimation of simple and hybrid DSGE models predictive distributions are subjective and provide a measure of the consistency of predicted probabilities versus their observed frequencies. In order to assess whether model density forecasts are consistent and well-calibrated Del Negro and Schorfheide (2012) generate histograms for probability integral transformations (PITs). This kind of evaluation became recently popular in many studies, e.g., in Herbst and Schorfheide, (2012), Kolasa *et al.* (2012) and Wolters (2013). Hence, we use predictive densities implied by all investigated DSGE models as well as by the VAR, BVARs and the FAVAR.

The Probability Integral Transform (PIT) was developed by Rosenblatt (1952) and introduced in the economic literature by Diebold, Gunther and Tay (1998). More formally, it is based on the relationship between the data generating process and the sequence of density forecasts via probability integral transforms of the observed data with respect to the density forecasts. The PIT of an observation $y^0_{i,T+h}$ based on its time T predictive distribution is

$$z_{i,h,T} = \int_{-\infty}^{y^0_{i,T+h}} p(y_{i,T+h}|Y_{1:T})dy_{i,T+h}, \tag{26}$$

thereby it represents the cumulative density of the random variable $y_{i,T+h}$ evaluated at the observation $y^0_{i,T+h}$. As shown in Del Negro and Schorfheide (2012) the PIT can be approximated by

$$z_{i,h,T} \approx \sum_{j=1}^{n_{sim}} I\left\{y^{(j)}_{i,T+h} \leq y_{i,T+h}\right\} \tag{27}$$

where $y_{i,T+h}$ is the value of y_i observed in period $T+h$ and $I\{x \geq \alpha\}$ denotes the indicator function. If the sequence of density forecasts is an accurate description of actual uncertainty, the sequence of PITs $z_{i,h,T}$ should be distributed uniformly between zero and one, thus the density forecast is well calibrated. Diebold *et al.* (1998) show that for $h = 1$, the $z_{i,h,T}$'s are uniformly distributed and independent across time $T : z_{i,h,T} \sim iidU[0,1]$ and they propose a number of approaches to forecast evaluation. We follow the one implemented in Herbst and Schorfheide (2012) and Del Negro and Schorfheide (2012). We divide the unit interval in K subintervals [9] and check if the fraction of PITs in each is close to K^{-1}. We set $K = 10$ so the density forecast is split into probability bands (equally sized histogram bins) that each covers 10% of the probability mass. In this way, the observed realizations and the corresponding out-of-sample density forecasts should be contained in each of the probability bands, otherwise the density forecasts are not a good characterization of the actual data distribution. The histograms of PITs are presented in Figures 1 and 2 for all models. The bars represent the fraction of realized observations falling into the corresponding deciles of the density forecasts, therefore the theoretical value of 10% for a well-calibrated model is represented by a solid line. We focus on one- and four-quarters-ahead forecasts[10].

As it is depicted in Figure 1 for the first-quarter-ahead (shorter horizon) the density forecasts are too diffuse especially for the GDP and CPI as many PITs fall into the middle bins with the exception of the FF rate. For output growth, an overly large fraction of PITs fall into the 0.3-0.6 bins which indicates that the density forecasts are not very well-calibrated. The peak in the middle of the histograms of

[9] In general, if the density forecast is divided into probability bands of equal coverage, data realisations should be uniformly distributed across all probability bands.

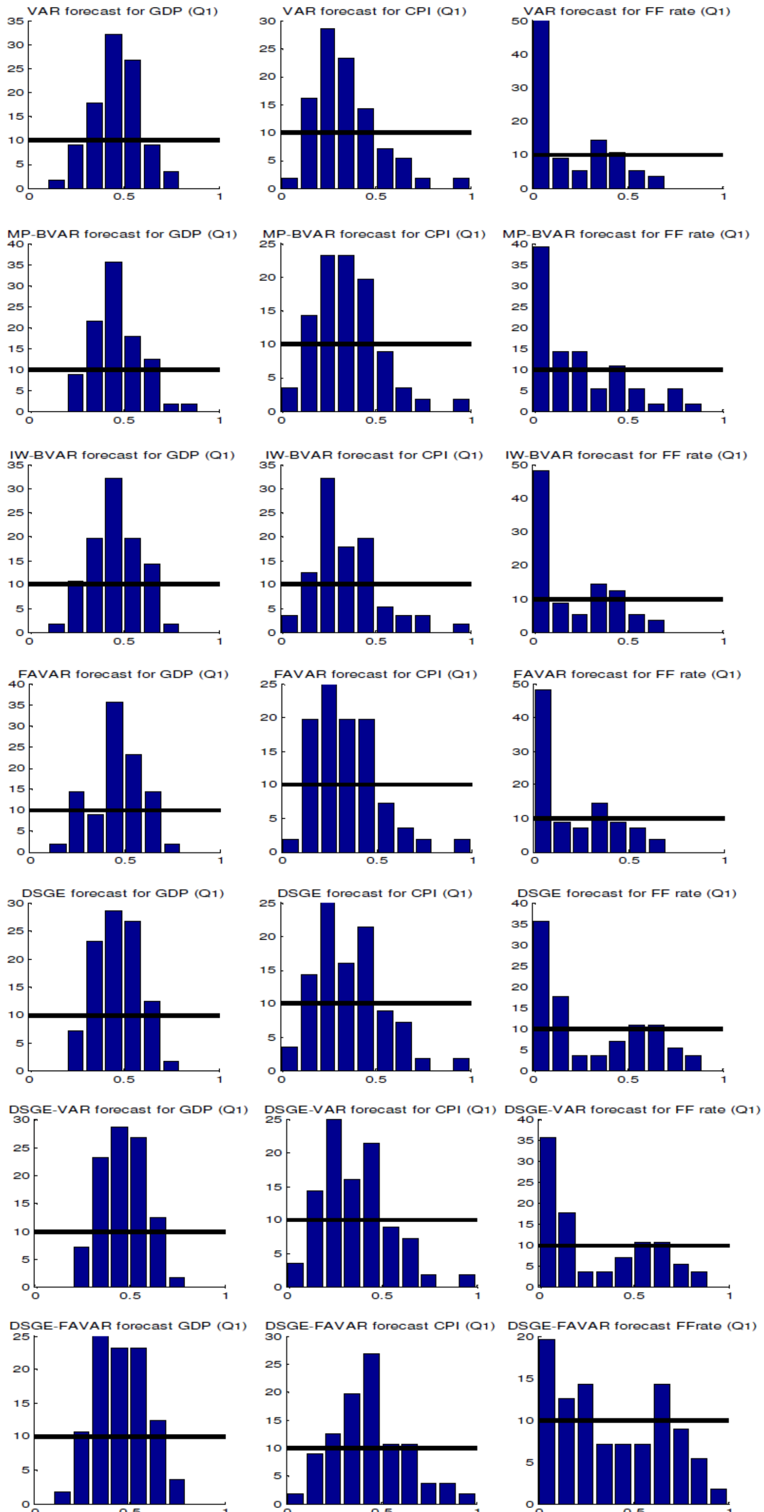
[10] Kolasa *et al.* (2012) report only the longest horizon i.e., four-quarters-ahead, whilst Del Negro and Schorfheide (2012) present three forecasting horizons. In this study we comparatively display the short-run (one-step-ahead) and the long-run (four-quarters-ahead) density forecast depiction in an attempt to capture the (potential) difference between these two predictive periods. The results for the medium-term horizons are almost identical and qualitatively similar and are available upon request.

the GDP and CPI forecasts shows that these overestimate actual uncertainty. The PITs for CPI are somewhat closer to a uniform distribution, yet there is still a peak in the middle of the distributions and the histograms for almost all models do not cover the entire distribution including the tails. Regarding the FF rate as in Kolasa *et al.* (2012) and Wolters (2013) an unproportional number of PITs fall in the lowest bin and the right tail is not covered, reflecting the fact that the models tend to overpredict the interest rate level and hence uncertainty is overestimated by the density forecasts. However, the comparative evaluation of the different models demonstrates that the DSGE-FAVAR outranks the other models in terms of a well-calibrated predictive density for all variables examined. Indeed, a peak is observed (for GDP and CPI) yet the PIT covers the entire distribution including the tails bringing it closer to a uniform shape. Especially for the FF rate the DSGE-FAVAR density appears to be better calibrated with a reduction in the left tail. The next best models seem to be the DSGE, the DSGE-VAR and in a less extent the FAVAR. For VAR, MP-BVAR and IW-BVAR the empirical distribution looks very different from a uniform distribution and the discrepancy is evident. The tails are not covered and uncertainty is overestimated [11]. The results are in full accordance with the works by Del Negro and Schorfheide (2012), Kolasa *et al.* (2012) and Wolters (2013). Overall, the density forecasts from the hybrid DSGE models and especially from DSGE-FAVAR appear better calibrated. Possibly, this is an after-effect of the imposition of tight restrictions on the data by the simple DSGEs. If the data rejects these restrictions, large stochastic shocks are needed to fit the model to the dataset which results in high shock uncertainty. As mentioned in Wolters (2013), DSGEs provide better densities during normal times. Under average exogenous shocks the models return back to a steady state, albeit they could not predict recessions and booms as significantly larger exogenous shocks are required to capture these. As hybrid models DSGE-FAVAR and DSGE-VAR relax these restrictions, misspecification can be absorbed by stochastic shocks (Del Negro and Schorfheide, 2009) and the estimated variance of shocks is lower, which results in somewhat tighter predictive distributions. Nevertheless, even though DSGEs are poor density forecasters in crisis times, they tend to display a better PIT than VARs and BVARs in our study. As sample observations for the recent financial crisis are included in our dataset, probably all models are not able to forecast the intensity of such a deep recession. Consequently, the left tail of density forecasts is pronounced in particular for the CPI and mainly for the FF rate during 2007-2010. All forecasts have a skewed distribution, i.e., they assign substantial probability to low rates that never materialized. In addition, a positive probability to relatively high FF rates and CPI is assigned even by the best performers (DSGE-FAVAR, DSGE and DSGE-VAR) that are not really observed during the forecast period. One potential explanation is that the forecasts are generated after the Great Moderation period whilst the estimation sample covers this period or before. According to Del Negro and Schorfheide (2012) the shock standard deviations are estimated to capture the average of the pre- and post-moderation volatility a fact that leads to overprediction of the volatility during the forecast period.

Similarly, the density forecasts look quite different from a uniform distribution and the discrepancy increases with the forecast horizon as it can be shown in Figure 2 for four-steps-ahead (longest horizon). Most PITs are too diffuse for the GDP and CPI falling into the middle bins including to some extent the FF rate as well. In this case, the FAVAR does not seem to present an equally descent performance as it was shown for the shorter horizon, while the DSGE as well fails to present a uniform-type distribution especially with respect to GDP and CPI. The DSGE-FAVAR is not performing equally well for the FF rate as in the first-quarter-ahead, although it is quite better for GDP and CPI predictions amongst the other models. Indeed, as the forecasting horizon increases the peak in the middle of the histograms is sharper for all models and it is expected that all models would overestimate actual uncertainty. Overall, the comparative investigation of the PITs could be informative of an indicative - though somewhat vague - outranking classification on the part of the DSGE-FAVAR, DSGE-VAR and DSGE, amongst the analyzed models.

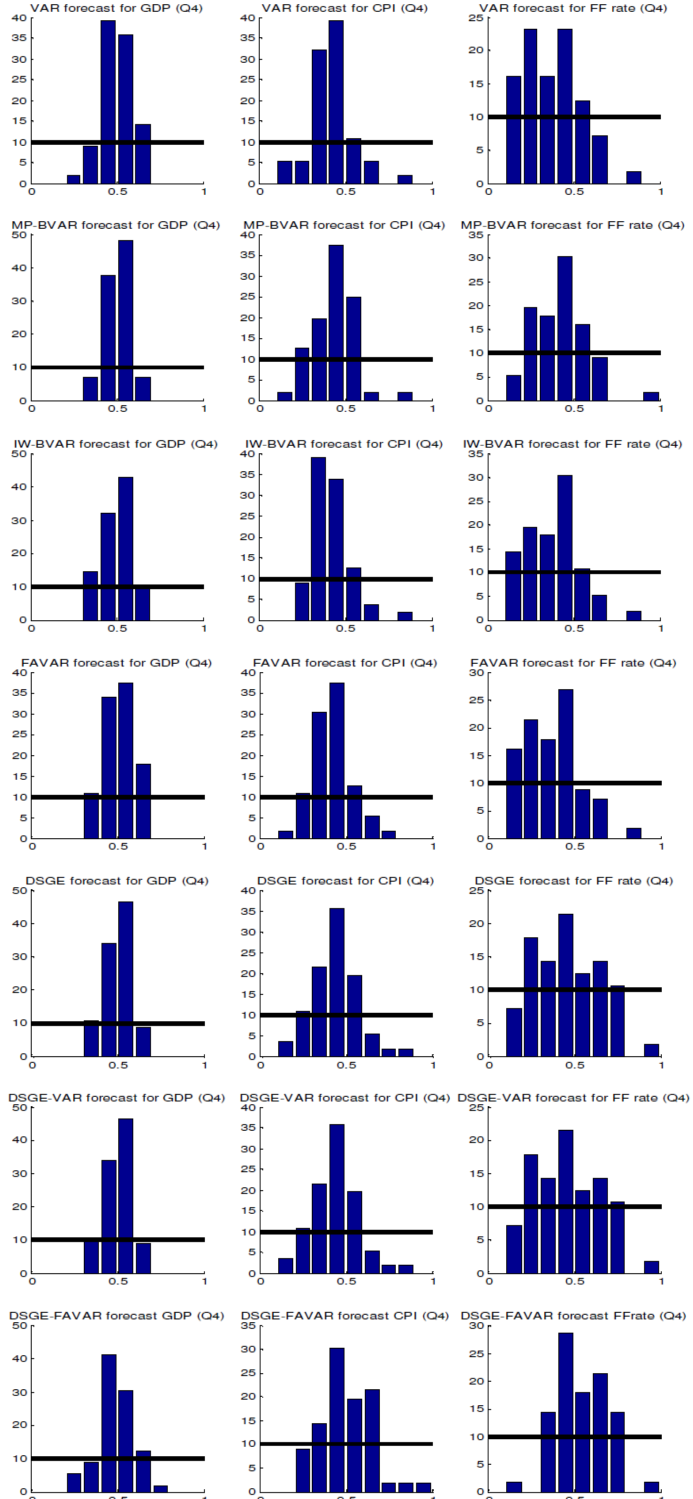
[11] There are formal tests to check for a uniform distribution (Berkowitz, 2001), even though as reported by Elder *et al.*, 2005 and Gerard and Nimark (2008) they have to be treated with high caution. In this study we do not add formal tests in accordance with Del Negro and Schorfheide (2012) and Kolasa *et al.* (2012), as the histograms already show a crystal clear evidence against a uniform distribution of the PITs.

Figure 1: PIT histograms for One-Quarter-Ahead Forecasts



Note: The density forecasts are displayed as probability bands each covering 10% of the density, whilst bars depict the fraction of realized observations falling into the deciles of density forecasts. For a well-calibrated model, the theoretical value (10%) is illustrated by the solid line.

Figure 2: PIT histograms for Four-Quarter-Ahead Forecasts



Note: As in Figure 1.

4.3 Real-Time Data analysis

Point and density forecasting analysis is repeated using Real-Time data. We utilize the Philadelphia Fed "Real-Time Data Set for Macroeconomists" described in detail by Croushore and Stark (2001). This ensures the comparability of the forecasting errors as all predictions are formulated using a similar data set (Edge *et al.*, 2010; Kolasa *et al.*, 2012; Del Negro and Schorfheide, 2012). We report RMSE, MDM and PIT results for the "winner" models as those emerged from our previous exhaustive analysis[12].

4.3.1 Point forecasts

Table 8 displays the RMSE scores for the winner models. For the GDP series the simple DSGE model outperforms the other models for all steps-ahead, as opposed to the results from our initial dataset where the FAVAR model was the best for the two-steps-ahead. The DSGE-VAR is the next best performer, whilst the DSGE-FAVAR shows a similar predictive performance with FAVAR and on average they generate the highest forecast errors. With regard to CPI, the DSGE provides once more with the relatively lowest RMSE score amongst the investigated models for all quarters-ahead, except for the two-steps-ahead where the DSGE-FAVAR outperforms the other models. The DSGE-VAR is marginally better than the FAVAR in accordance with the PIT results provided above. When comparing the RMSEs for the FF rate the results are diverse, in that the best model for two- and three-steps-ahead is the DSGE-FAVAR while the DSGE produces the lowest RMSE for the shortest and longest horizon. To sum-up, the overall best performance in terms of the RMSE using Real-Time data amongst the winner models is DSGE followed by the two other hybrids, i.e., DSGE-FAVAR and DSGE-VAR.

Table 8: Root Mean Square Forecast Error (RMSE) for GDP, CPI and FF rate using Real-Time data

	FAVAR	DSGE	DSGE-VAR	DSGE-FAVAR
GDP				
1	0.770	0.685	0.719	0.789
2	0.782	0.708	0.717	0.792
3	0.777	0.676	0.718	0.782
4	0.767	0.677	0.720	0.783
CPI				
1	0.965	0.806	0.936	0.926
2	0.984	0.953	0.951	0.948
3	0.977	0.840	0.959	0.965
4	0.989	0.816	0.943	0.943
FF rate				
1	3.885	2.920	4.045	3.636
2	3.993	4.066	4.088	3.766
3	4.026	3.969	4.276	3.396
4	4.185	3.376	4.294	3.953

In the original dataset the pairwise differential predictability was tested via the MDM with the Newey-West estimator to correct for the autocorrelation of the forecast errors. The results in Tables 9, 10 and 11 for the Real-Time data showed that for the GDP the DSGE-FAVAR displayed evidence of weak or strong differential forecastability (1% , 5% or 10%) in almost all step-ahead forecasts, as opposed to the initial results where the pairwise forecast comparison demonstrated a lack of statistical significance. For the pairs DSGE vs. DSGE-VAR and FAVAR vs. DSGE the MDM p-values are high and the comparison

[12] The MDD has already been utilized to select the best lag structure for all models including the winners.

proved mostly insignificant. In general, the values were different from those of the raw data towards the increase of statistical significance. For the CPI variable, most models for all forecast horizons appear to have a significant pairwise predictability, except for the DSGE pairs against the DSGE-VAR and DSGE-FAVAR for the two-periods-ahead and the DSGE-VAR vs. DSGE-FAVAR for all steps-ahead. If we compare this outcome to the RMSE results, the DSGE seems to corroborate its enhanced predictability when examined versus the other "winner" models. Interestingly, the hybrids show a distinctive differential predictability in their comparative evaluation. In case of the FF rate the majority of pairwise comparisons produce a statistically significant MDM score at 1% level. Only the pair DSGE vs DSGE-VAR for the two-quarters-ahead shows insignificance. In combination with the RMSE results the DSGE-FAVAR and the DSGE confirm their outperformance. The outcome for nowcasts is slightly different in that the FAVAR for Real-Time data does not emerge as the best performer in two- and four-steps-ahead as was the case for the initial dataset. Overall, the point forecast evaluation via the RMSE and MDM revealed the superiority of the DSGE simple and hybrid models.

Table 9: Pairwise forecast comparison for the GDP with the Modified Diebold-Mariano test using Real-Time data

GDP	PERIODS			
	1	2	3	4
FAVAR vs DSGE	0.039	0.032	0.079	0.087
FAVAR vs DSGE-VAR	0.002	0.013	0.022	0.038
FAVAR vs DSGE-FAVAR	0.000	0.011	0.128	0.026
DSGE vs DSGE-VAR	0.137	0.217	0.164	0.143
DSGE vs DSGE-FAVAR	0.022	0.025	0.079	0.074
DSGE-VAR vs DSGE-FAVAR	0.001	0.010	0.027	0.031

Notes: As in Table 5

Table 10: Pairwise forecast comparison for the CPI with the Modified Diebold-Mariano test using Real-Time data

CPI	PERIODS			
	1	2	3	4
FAVAR vs DSGE	0.000	0.003	0.000	0.000
FAVAR vs DSGE-VAR	0.003	0.002	0.044	0.003
FAVAR vs DSGE-FAVAR	0.000	0.000	0.001	0.000
DSGE vs DSGE-VAR	0.000	0.395	0.000	0.000
DSGE vs DSGE-FAVAR	0.000	0.427	0.000	0.000
DSGE-VAR vs DSGE-FAVAR	0.029	0.344	0.260	0.427

Notes: As in Table 5

Table 11: Pairwise forecast comparison for the FF rate with the Modified Diebold-Mariano test using Real-Time data

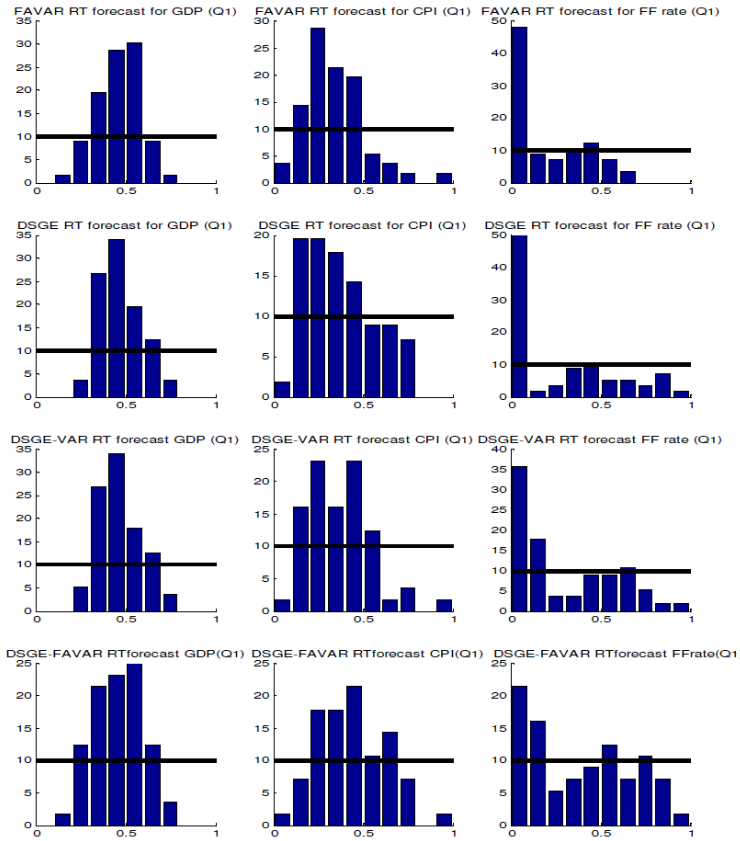
FF Rate	PERIODS			
	1	2	3	4
FAVAR vs DSGE	0.000	0.002	0.001	0.001
FAVAR vs DSGE-VAR	0.000	0.003	0.000	0.042
FAVAR vs DSGE-FAVAR	0.000	0.000	0.043	0.001
DSGE vs DSGE-VAR	0.000	0.307	0.000	0.000
DSGE vs DSGE-FAVAR	0.000	0.000	0.000	0.002
DSGE-VAR vs DSGE-FAVAR	0.000	0.000	0.000	0.000

Notes: As in Table 5

4.3.2 Density forecasts

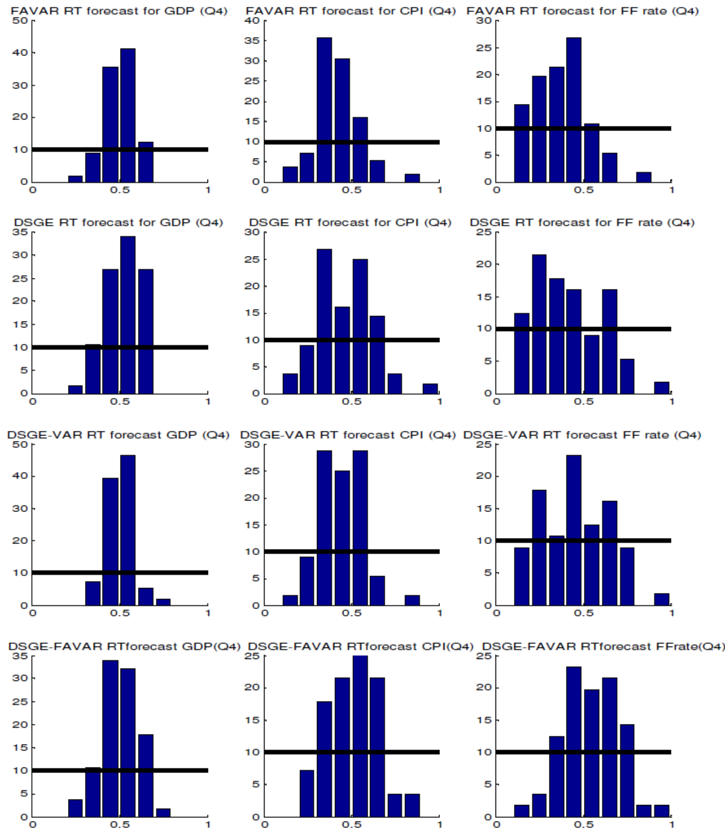
The histograms of PITs for the short and long horizon are presented in Figures 3 and 4 respectively for the winner models. The picture that emerges is similar to the conventional dataset. Namely, the PITs are diffuse especially for the GDP and CPI with a large fraction falling into the 0.2-0.6 bins, a sign of actual uncertainty overestimation. The density forecasts do not include the tails with the exception of the FF rate as in Kolasa *et al.* (2012) and Wolters (2013). This signifies that the models tend to overpredict the interest rate level. From the comparative evaluation of the different models the DSGE-FAVAR appears to marginally outrank the FAVAR, simple DSGE and the DSGE-VAR providing a more well-calibrated predictive density for all variables. Again, this might be due to the imposition of tighter restrictions on the data by the simple DSGE. Hybrids relax these restrictions, hence misspecification can be absorbed by stochastic shocks as reported in Del Negro and Schorfheide (2009). For the one-step-ahead Real-Time density forecasts the FF rate PITs have a skewed distribution, i.e., they assign a large probability to low rates that are never realized exactly as for the initial data, yet now even more pronounced. The rolling sample generation of predictions as well as the Great Moderation in-sample period could explain this fact according to Del Negro and Schorfheide (2012). As the average of the pre- and post- Great Moderation volatility is captured by the sample, an overprediction of the volatility within the forecast period occurs. Indicatively, the PITs for the DSGE-FAVAR look better for Real-Time data in case of the CPI and FF rate amongst the other models compared to the initial dataset, whilst the same outcome is observed for the FAVAR and DSGE. Finally, for the longest horizon investigation it is revealed that even though the discrepancy increases with the forecast horizon, the PITs for the FF rate are much better calibrated for the DSGE models as opposed to the FAVAR. The rest of the density forecasts are similar to the four-quarter-ahead results for the conventional dataset. The skewness in the FF rate PITs is vanished. Indeed the overall picture leads to the conclusion that the PITs are diffuse for the GDP and CPI, yet the FF rate ones seem now better-calibrated. Amongst the winners, the DSGE-FAVAR presents a marginally better outcome than DSGE-VAR and DSGE, though not particularly different compared to the original dataset density examination.

Figure 3: PIT histograms for One-Quarter-Ahead Forecasts of Real-Time data



Note: As in Figure 1.

Figure 4: PIT histograms for Four-Quarter-Ahead Forecasts of Real-Time data



Note: As in Figure 1.

5 Conclusions

This study investigated and employed advanced Bayesian methods for estimating dynamic stochastic general equilibrium (DSGE) models. These models appear to be particularly suited for conducting policy evaluation, as shown in the works of Smets and Wouters (2003, 2004), Del Negro and Schorfheide (2004), Adolfson *et al.* (2008) and Christiano *et al.* (2005). However, calibrated DSGE models face many important challenges such as the fragility of parameter estimates, statistical fit and the weak reliability of policy forecasts as reported in Stock and Watson (2001), Ireland (2004) and Schorfheide (2010). In recent years Bayesian estimation has become popular mainly due to the increasing computational power available to estimate large-scale DSGE models using Markov chain Monte Carlo simulations. DSGE models can pose identification problems for frequentist estimation that no amount of data or computing power can overcome. New macroeconomic research is drawn to the application of Bayesian statistics because DSGE models are often seen as abstractions of actual economies.

We included a comparative evaluation of the out-of-sample predictive performance of many different specifications of estimated DSGE models and various classes of VAR models, using point and density forecasts to evaluate their predictability for the US economy. Simple and hybrid DSGE models were implemented, such as DSGE-VARs and Factor Augmented DGSEs (DSGE-FAVAR), and tested against standard VARs, Bayesian VARs and Factor Augmented VARs (FAVAR). The investigated period spanned 1960:Q4 to 2010:Q4 for the real GDP, the harmonized CPI and the nominal short-term interest rate. We produced their forecasts for the out-of-sample testing period 1997:Q1-2010:Q4 using in addition Real-Time data. The results were evaluated with the use of Bayesian method of the marginal data density as well as the root mean squared forecast error. The modified Diebold-Mariano (1995) pairwise test was also employed to measure comparatively the differential forecastability. The best forecasting performance for the GDP was consistently produced by the plain DSGE model also in case of Real-Time data. For the CPI and FF rate macroeconomic variables different models provided with the most accurate forecasts depending on the forecast horizon, the statistical measure of predictability and the use of Real-Time dataset. In the majority of cases the DSGE-FAVAR was the best performer whilst the FAVAR or the DSGE were also valid alternatives, the latter in particular when using Real-Time forecasts. Interestingly, the FAVAR model that provided with a consistent outranking behavior for the conventional dataset, was not included in the list of winners when Real-Tine data was used. VAR and BVAR specifications provided with less satisfying forecasting results. Point forecasts of DSGE and hybrid DSGEs were quite precise. A comparison against the statistical VAR and BVAR forecasts showed that the structural form of the DSGEs that is helpful in estimation inference and interpretability, does not worsen their forecasting performance. More importantly, conditioning on accurate nowcasts the forecasting performance was increased for the CPI and FF rate persistent variables. However, whilst the point forecasts were quite accurate, density forecasts of DSGE models - even for the outperformer DSGE-FAVAR - were not really well-calibrated in all horizons and in general they overestimated uncertainty. A possible explanation based on the same results by Del Negro and Schorfheide (2012) and Kolasa *et al.* (2012) is attributed to the inclusion of the Great Moderation evaluation period in estimation as opposed to the transition in a highly volatile out-of-sample period especially after 2007.

As hybrid DSGE models have become popular for dealing with some of the DSGE model misspecifications basically due to major advances in Bayesian estimation methodology, policymakers and practitioners can rely on these models for forecasting applications. As shown in this study these models can compete and outperform well-known statistical models (e.g., VARs and BVARs) and effectively deal with more complex real-world problems as richer data sources and Real-Time datasets become available. The present comparative model validation can be useful to monetary policy analysis and macro-forecasting.

References

Adolfson, M., Andersson, M, Linde, J, Villani, M, and A. Vredin (2007) Modern Forecasting Models in Action: Improving Macroeconomic Analyses at Central Banks, International Journal of Central Banking 3(4), 111-144.

Adolfson M, Stefan L, Jesper L, Mattias V (2008) Evaluating an Estimated New Keynesian Small Open Economy Model, Journal of Economic Dynamics and Control, 32(8), 2690-2721

Altug S (1989) Time-to-Build and Aggregate Fluctuations: Some New Evidence, International Economic Review, 30(4), 889-920

An S, Schorfheide F (2007) Bayesian Analysis of DSGE Models, Econometric Reviews, 26(2-4) 113-172.

Bai J., Ng S. (2000) Determining the Number of Factors in Approximate Factor Models, Econometrica, 70.

Banbura M, Giannone D, Reichlin L (2010) Large Bayesian vector autoregressions, Journal of Applied Econometrics, 25(1), 71-92

Berkowitz, J., (2001) Testing density forecasts, with applications to risk management. Journal of Business and Economic Statistics 19(4), 465–474

Bernanke B S., Boivin J. (2003) Monetary Policy in a Data-Rich Environment, Journal of Monetary Economics 50 (3), 525-546.

Bernanke B S., Boivin J. and Eliasz P. (2005) Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach, The Quarterly Journal of Economics, MIT Press, vol. 120(1), pages 387-422, January.

Boivin J, Giannoni MP (2006) DSGE Models in a Data-Rich Environment, NBER Working Papers 12772

Boivin J, Giannoni MP and Mihov I. (2009) Sticky Prices and Monetary Policy: Evidence from Disaggregated US Data, American Economic Review, American Economic Association, vol. 99(1), 350-84

Brüggemann R, H. Lütkepohl and M. Marcellino (2008) Forecasting euro area variables with German pre-EMU data, Journal of Forecasting, 27(6), 465–481

Christiano, L. J., M. Eichenbaum, C. Evans (2005) Nominal Rigidities and the Dynamic effects of a Shock to Monetary Policy, Journal of Political Economy 113, 1-45.

Christiano, L. J., Eichenbaum M, (1992) Current Real Business Cycle Theories and Aggregate labor Market Fluctuation, American Economic Review, 82, 430-450.

Christoffel K, Günter C, Anders W (2008) The New Area-Wide Model of the Euro Area - A Micro-Founded Open-Economy Model for Forecasting and Policy Analysis, European Central Bank Working Paper Series n. 944

Chudik A, Pesaran M. H, (2011) Infinite-dimensional VARs and factor models, Journal of Econometrics 163(1), 4-22

Clarida R, Galí J, Gertler M (2000) Monetary Policy Rules And Macroeconomic Stability: Evidence And Some Theory, The Quarterly Journal of Economics, 115(1), 147-180

Consolo A, Favero CA and Paccagnini A (2009) On the Statistical Identification of DSGE Models, Journal of Econometrics, 150, 99-115.

Croushore D. and Stark T, (2001) A Real-Time Data Set for Macroeconomists, Journal of Econometrics, 105, 111–30.

Dawid, A. (1984) Statistical Theory: The Prequential Approach, Journal of the Royal Statistical Society Series A 147(2), 278-292.

DeJong D, Ingram B, Whiteman C (1996) A Bayesian Approach to Calibration, Journal of Business Economics and Statistics, 14, 1-9

Del Negro M, Schorfheide F (2004) Priors from General equilibrium Models for VARs, International Economic Review, 45, 643-673

Del Negro M, Schorfheide F (2006) How Good Is What You've Got? DSGE-VAR as a Toolkit for Evaluating DSGE Models, Federal Reserve Bank of Atlanta Economic Review 91, Second Quarter.

Del Negro M, Schorfheide F, Smets F, Wouters R (2007) On the Fit of New-Keynesisan Models, Journal of Business, Economics and Statistics, 25(2), 124-162

Del Negro M, Schorfheide F (2012) DSGE Model-Based Forecasting, Handbook of Economic Forecasting, Volume 2, July 1

Del Negro M, Schorfheide F (2009) Monetary Policy Analysis with Potentially Misspecified Models, American Economic Review, 99, 1415–50.

Diebold FX, Mariano RS (1995) Comparing predictive accuracy, Journal of Business and Economic Statistics, 13(3), 253-263

Diebold, F. X., Todd A. G, and Tay A. S, (1998) Evaluating Density Forecasts with Applications to Financial Risk Management, International Economic Review 39, 863–83.

Doan T, Litterman R, Sims C (1984) Forecasting and Conditional Projections Using Realistic Prior Distributions, Econometric Reviews, 3, 1-100

Dua P, Ray SC (1995) A BVAR model for the Connecticut Economy, Journal of Forecasting, 14(3), 167-180

Edge, R., M. Kiley, and J.-P. Laforte (2010) A Comparison of Forecast Performance Between Federal Reserve Sta Forecasts, Simple Reduced-Form Models, and a DSGE Model, Journal of Applied Econometrics 25(4), 720-754

Elder R., Kapetanios G., Taylor T., Yates T., (2005) Assessing the MPC's fan charts. Bank of England Quarterly Bulletin Autumn, 326–345.

Fernández-de-Córdoba G, Torres JL (2010) Forecasting the Spanish Economy with a DSGE Model: An Augmented VAR Approach, Journal of the Spanish Economic Association. Forthcoming

Forni M, Reichlin L. (1996): Dynamic Common Factors in Large Cross-Sections, Empirical Economics, 21, 27-42.

Forni M, Reichlin L. (1998): Let's Get Real: A Dynamic Factor Analytical Approach to Disaggregated Business Cycle, Review of Economic Studies, 65, 453-474.

Forni M, Hallin M, Lippi M and Reichlin L. (1999): Reference Cycles: The NBER Methodology Revisited, mimeo.

Forni M, Hallin M, Lippi M and Reichlin L. (2000): The Generalized Dynamic-Factor Model: Identification And Estimation, The Review of Economic and Statistics, MIT Press, vol. 82(4), pages 540-554, November.

Gerard H., Nimark K., (2008) Combing multivariate density forecasts using predictive criteria, Research Discussion Paper 2008-2, Reserve Bank of Australia.

Geweke J (1999) Using Simulation Methods for Bayesian Econometric Models: Inference, Development and Communication, Econometric Reviews 18(1), 1-126

Ghent A (2009) Comparing DSGE-VAR Forecasting Models: How Big are the Differences?, Journal of Economic Dynamics and Control 33(4), 864-882

Herbst E, Schorfheide F (2012) Evaluating DSGE model forecasts of comovements, Journal of Econometrics 171(2), 152-166

Ingram B, Whiteman C (1994) Supplanting the Minnesota Prior - Forecasting Macroeconomics Time Series using Real Business Cycle Model Priors, Journal of Monetary Economics, 34, 497-510

Ireland P (2004) A Method for Taking Models to the Data, Journal of Economic Dynamics and Control, 28, 1205-1226

King RG (2000) The New IS-LM Model: Language, Logic, and Limits, Federal Reserve Bank of Richmond Economic Quarterly, 86, 45-103

Kolasa M, Rubaszek M, Skrzypczynski P (2009) Putting the New Keynesian DSGE Model to the Real-Time Forecasting Test, European Central Bank Working Paper Series n. 1110

Kolasa M, Rubaszek M, Skrzypczynski P (2012) Putting the New Keynesian DSGE Model to the Real-Time Forecasting Test, Journal of Money, Credit and Banking, 2012, 44 (7), 1301–1324.

Kydland FE, Prescott EC (1982) Time to build and aggregate ‡uctuations, Econometrica 50(6), 1345-70

Litterman RB (1981) A Bayesian Procedure for Forecasting with Vector Autoregressions, Working Paper, Federal Reserve Bank of Minneapolis

Litterman RB (1986) Forecasting with Bayesian Vector Autoregressions: Five Years of Experience, Journal of Business and Statistics 4(1), 25–38

Marcellino, M. (2004) Forecasting EMU macroeconomic variables, International Journal of Forecasting, 20, 359–372

Newey, W. K., West, K. D. (1987) A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, Econometrica 55(3), 703–708.

Newey, W. K., West, K. D. (1994) Automatic lag selection in covariance matrix estimation, Review of Economic Studies 61(4), 631–654.

Robertson J. C. and Tallman E W, (1999) Vector autoregressions: forecasting and reality, Economic Review 4–18.

Rosenblatt M, (1952) Remarks on a Multivariate Transformation, Annals Mathematical Statistics, 23, 470–72.

Rubaszek M, Skrzypczynski P (2008) On the Forecasting Perfomance of a Small-Scale DSGE Model, International Journal of Forecasting, 24, 498-512

Sargent T (1989) Two Models of Measurements and the Investment Accelerator, Journal of Political Economy, 97(2), 251-287

Schorfheide F (2010) Estimation and Evaluation of DSGE Models: Progress and Challenges, University of Pennsylvania

Schorfheide F (2000) Loss function-based evaluation of DSGE models, Journal of Applied Econometrics 15 (6), 645-670.

Sims CA (1980) Macroeconomics and reality, Econometrica, 48(1), 1-48

Sims, CA (2002) Solving Linear Rational Expectations Models, Computational Economics, 20 (1-2), 1-20

Sims CA, Zha T (1998) Bayesian Methods for Dynamic Multivariate Models, International Economic Review, 39, 949-968

Smets F, Wouters R (2003) An Estimated Stochastic Dynamic General Equilibrium Model of the Euro Area, Journal of the European Economic Association, 1, 1123-1175

Smets F, Wouters R (2004) Forecasting with a Bayesian DSGE model: an application to the Euro area, Working Paper No. 389, European Central Bank, Frankfurt

Smets F, Wouters R (2007) Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach, American Economic Review 97 (3), 586–606

Spencer DE (1993) Developing a Bayesian Vector Autoregression Forecasting Model, International Journal of Forecasting 9(3), 407–421

Stock JH, Watson WM (2001) Vector autoregressions, Journal of Economic Perspectives, 15, 101-115

Stock JH, Watson WM (2002) Macroeconomic Forecasting Using Diffusion Indexe , Journal of Business Economics and Statistics, XX:II, 147-162.

Theil H, Goldberg AS (1961) On Pure and Mixed Estimation in Economics, International Economic Review, 2, 65-78

Todd RM (1984) Improving Economic Forecasting with Bayesian Vector Autoregression, Quarterly Review, Federal Reserve Bank of Minneapolis

Wolters, M. H. (2013) Evaluating point and density forecasts of DSGE models, Working Paper

Woodford M (2003) Interest and Prices, Princeton University Press

A Appendix

The source of the data is the Federal Reserve Economic Data - Federal Reserve Bank of Saint Louis (http://research.stlouisfed.org/fred2/). In order to construct the FAVAR we extract factors from a balanced panel of 109 monthly and quarterly macroeconomic and financial time series, following the dataset built by Stock and Watson (2002). The dataset involves several measures of industrial production, interest rates, various price indices, employment and other important macroeconomic and also financial variables. In the following Table, the first column has the series number, the second the series acronym, the third the series description, the fourth the transformation codes and the fifth column denotes a slow-moving variable with 1 and a fast-moving one with 0. The transformed series are tested using the Box-Jenkins procedure and the Dickey-Fuller test. Following Bernanke *et al.* (2005), the transformation codes are as follows: 1 - no transformation; 2 - first difference; 4 - logarithm; 5 - first difference of logarithm; 6 - second difference; 7 - second difference of logarithm.

Date	Long Description	Tcode	SlowCode
PAYEMS	Total Nonfarm Payrolls: All Employees	5	1
DSPIC96	Real Disposable Personal Income	5	1
NAPM	ISM Manufacturing: PMI Composite Index	1	1
UNRATE	Civilian Unemployment Rate	1	1
INDPRO	Industrial Production Index (Index 2007=100)	5	1
PCEPI	Personal Consumption Expenditures: Chain-type Price Index (Index 2005=100)	5	1
PPIACO	Producer Price Index: All Commodities (Index 1982=100)	5	1
FEDFUNDS	Effective Federal Funds Rate	1	0
IPDCONGD	Industrial Production: Durable Consumer Goods (Index 2007=100)	5	1
IPBUSEQ	Industrial Production: Business Equipment (Index 2007=100)	5	1
IPMAT	Industrial Production: Materials (Index 2007=100)	5	1
IPCONGD	Industrial Production: Consumer Goods (Index 2007=100)	5	1
IPNCONGD	Industrial Production: Nondurable Consumer Goods (Index 2007=100)	5	1
IPFINAL	Industrial Production: Final Products (Market Group) (Index 2007=100)	5	1
UNEMPLOY	Unemployed	5	1
EMRATIO	Civilian Employment-Population Ratio (%)	1	1
CE16OV	Civilian Employment	5	1
CLF16OV	Civilian Labor Force	5	1
CIVPART	Civilian Participation Rate (%)	1	1
UEMP27OV	Civilians Unemployed for 27 Weeks and Over	5	1
UEMPLT5	Civilians Unemployed - Less Than 5 Weeks	5	1
UEMP15OV	Civilians Unemployed - 15 Weeks & Over	5	1
UEMP15T26	Civilians Unemployed for 15-26 Weeks	5	1
UEMP5TO14	Civilians Unemployed for 5-14 Weeks	5	1
MANEMP	Employees on Nonfarm Payrolls: Manufacturing	5	1
USPRIV	All Employees: Total Private Industries	5	1
USCONS	All Employees: Construction	5	1
USFIRE	All Employees: Financial Activities	5	1
USTRADE	All Employees: Retail Trade	5	1
DMANEMP	All Employees: Durable Goods Manufacturing	5	1
USGOOD	All Employees: Goods-Producing Industries	5	1
USEHS	All Employees: Education & Health Services	5	1
USLAH	All Employees: Leisure & Hospitality	5	1
SRVPRD	All Employees: Service-Providing Industries	5	1
USINFO	All Employees: Information Services	5	1
USPBS	All Employees: Professional & Business Services	5	1
USTPU	All Employees: Trade, Transportation & Utilities	5	1
NDMANEMP	All Employees: Nondurable Goods Manufacturing	5	1
USMINE	All Employees: Natural Resources & Mining	5	1
USWTRADE	All Employees: Wholesale Trade	5	1
USSERV	All Employees: Other Services	5	1
AHEMAN	Average Hourly Earnings: Manufacturing	5	1
AHECONS	Average Hourly Earnings: Construction (NSA)	5	1
PPIIDC	Producer Price Index: Industrial Commodities (NSA)	5	1

PPIFGS	Producer Price Index: Finished Goods (Index 1982=100)	5	1
PPICPE	Producer Price Index: Finished Goods: Capital Equipment (Index 1982=100)	5	1
PPICRM	Producer Price Index: Crude Materials for Further Processing (Index 1982=100)	5	1
PPIITM	Producer Price Index: Intermediate Materials: Supplies & Components (Index 1982=100)	5	1
PPIENG	Producer Price Index: Fuels & Related Products & Power (Index 1982=100)	5	1
PPIFCG	Producer Price Index: Finished Consumer Goods (Index 1982=100)	5	1
PFCGEF	Producer Price Index: Finished Consumer Goods Excluding Foods (Index 1982=100)	5	1
CPIAUCSL	Consumer Price Index for All Urban Consumers: All Items (Index 1982=100)	5	1
CPIAUCNS	Consumer Price Index for All Urban Consumers: All Items (Index 1982-84=100)	5	1
CPILFESL	Consumer Price Index for All Urban Consumers: All Items Less Food & Energy (Index 1982-84=100)	5	1
CPILFENS	Consumer Price Index for All Urban Consumers: All Items Less Food & Energy (NSA Index 1982=100)	5	1
CPIUFDNS	Consumer Price Index for All Urban Consumers: Food (NSA Index 1982=100)	5	1
CPIENGNS	Consumer Price Index for All Urban Consumers: Energy (NSA Index 1982=100)	5	1
CPIENGSL	Consumer Price Index for All Urban Consumers: Energy (Index 1982-1984=100)	5	1
CPILEGSL	Consumer Price Index for All Urban Consumers: All Items Less Energy (Index 1982-1984=100)	5	1
CPIMEDSL	Consumer Price Index for All Urban Consumers: Medical Care (Index 1982-1984=100)	5	1
PPIFCF	Producer Price Index: Finished Consumer Foods (Index 1982=100)	5	1
AAA	Moody's Seasoned Aaa Corporate Bond Yield	1	0
BAA	Moody's Seasoned Baa Corporate Bond Yield	1	0
M2SL	M2 Money Stock	6	0
M2NS	M2 Money Stock (NSA)	6	0
M1NS	M1 Money Stock (NSA)	6	0
M3SL	M3 Money Stock (DISCONTINUED SERIES)	6	0
GS5	5-Year Treasury Constant Maturity Rate	1	0
GS10	10-Year Treasury Constant Maturity Rate	1	0
GS1	1-Year Treasury Constant Maturity Rate	1	0
GS3	3-Year Treasury Constant Maturity Rate	1	0
TB3MS	3-Month Treasury Bill: Secondary Market Rate	1	0
TB6MS	6-Month Treasury Bill: Secondary Market Rate	1	0
HOUST	Housing Starts: Total: New Privately Owned Housing Units Started	5	0
PERMIT	New Private Housing Units Authorized by Building Permits	5	0
HOUSTMW	Housing Starts in Midwest Census Region	5	0
HOUSTW	Housing Starts in West Census Region	5	0
HOUSTNE	Housing Starts in Northeast Census Region	5	0
HOUSTS	Housing Starts in South Census Region	5	0
PERMITS	New Private Housing Units Authorized by Building Permits - South	5	0
PERMITMW	New Private Housing Units Authorized by Building Permits - Midwest	5	0
PERMITW	New Private Housing Units Authorized by Building Permits - West	5	0
PERMITNE	New Private Housing Units Authorized by Building Permits - Northeast	5	0
PDI	Personal Dividend Income	5	0
SPREAD1	3mo-FYFF	1	0
SPREAD2	6mo-FYFF	1	0
SPREAD3	1yr-FYFF	1	0
SPREAD4	2yr-FYFF	1	0
SPREAD5	3yr-FYFF	1	0
SPREAD6	5yr-FYFF	1	0
SPREAD7	7yr-FYFF	1	0
SPREAD8	10yr-FYFF	1	0
PCECC96	Real Personal Consumption Expenditures (Billions of Chained 2005 Dollars)	5	1
UNLPNBS	Nonfarm Business Sector: Unit Nonlabor Payments (Index 2005=100)	5	1
IPDNBS	Nonfarm Business Sector: Implicit Price Deflator (Index 2005=100)	5	1
OUTNFB	Nonfarm Business Sector: Output (Index 2005=100)	5	1
HOANBS	Nonfarm Business Sector: Hours of All Persons (Index 2005=100)	5	1
COMPNFB	Nonfarm Business Sector: Compensation Per Hour (Index 2005=100)	5	1
ULCNFB	Nonfarm Business Sector: Unit Labor Cost (Index 2005=100)	5	1
COMPRNFB	Nonfarm Business Sector: Real Compensation Per Hour (Index 2005=100)	5	1
OPHNFB	Nonfarm Business Sector: Output Per Hour of All Persons (Index 2005=100)	5	1
OPHPBS	Business Sector: Output Per Hour of All Persons (Index 2005=100)	5	1
ULCBS	Business Sector: Unit Labor Cost (Index 2005=100)	5	1
RCPHBS	Business Sector: Real Compensation Per Hour (Index 2005=100)	5	1
HCOMPBS	Business Sector: Compensation Per Hour (Index 2005=100)	5	1
OUTBS	Business Sector: Output (Ineex 2005=100)	5	1
HOABS	Business Sector: Hours of All Persons (Index 2005=100)	5	1
IPDBS	Business Sector: Implicit Price Deflator (Index 2005=100)	5	1
CP	Corporate Profits After Tax	5	0
SP500	S&P 500 Index	5	0