DEPARTMENT OF ECONOMICS

UNIVERSITY OF MILAN - BICOCCA

WORKING PAPER SERIES

Comparing Hybrid DSGE Models

Alessia Paccagnini

No. 228 – November 2012

Dipartimento di Economia Politica
Università degli Studi di Milano - Bicocca
http://dipeco.economia.unimib.it

Comparing Hybrid DSGE Models

Alessia Paccagnini*

November 19, 2012

Abstract

This paper discusses the estimation of Dynamic Stochastic General Equilibrium (DSGE) models using hybrid models. These econometric tools provide the combination of an atheoretical statistical representation and the theoretical features of the DSGE model. A review of hybrid models presents the main aspects of these tools and why they are needed in the recent macroeconometric literature. Some of these models are compared to classical econometrics models (such as Vector Autoregressive, Factor Augmented VAR and Bayesian VAR) in a marginal data density analysis.

JEL CODES: C11, C15, C32

KEYWORDS: Model Estimation, Bayesian Analysis, DSGE Models, Vector Autoregressions

"Dynamic equilibrium theory made a quantum leap between the early 1970s and the late 1990s. In the comparatively brief space of 30 years, macroeconomists went from writing prototype models of rational expectations (think of Lucas, 1972) to handling complex constructions like the economy in Christiano, Eichenbaum, and Evans (2005). It was similar to jumping from the Wright brothers to an Airbus 380 in one generation".

Jesus Fernández-Villaverde in "The Econometrics of DSGE Models"(2009, pag.2)

"A well-defined statistical model is one whose underlying assumptions are valid for the data chosen"

Aris Spanos in "The Simultaneous-Equations Model Revisited: Statistical Adequacy and Identification" (1990, pag.89)

*Università degli Studi Milano - Bicocca Department of Economics E-mail : alessia.paccagnini@unimib.it. I thank Laurie Anderson, David Barnes, and participants at the Time Series Reading Group at the EUI, especially, Stelios Bekiros, Helmut Lutkepohl, Massimiliano Marcellino, and Tomasz Wozniak. Alessia Paccagnini acknowledges financial support from "Dote ricercatori": FSE, Regione Lombardia.

1 Introduction

The new micro-founded dynamic stochastic general equilibrium DSGE 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). Model validation using DSGE models allows the econometrician to establish a link between structural features of the economy and reduced form parameters, which was not always possible with the usual large-scale macroeconomic models. For more than two decades, vector autoregressive models (VARs) (Sims, 1980) have been used for macroeconomic modelling and forecasting.

On the side of the statistical representation, the VAR suffers from the 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. The use of "Minnesota" priors (Doan *et al.*, 1984) has been proposed to shrink the parameters space and thus overcome the curse of dimensionality.

On the other side, the calibrated 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 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. One of the main reasons of the extensive use of Bayesian methods is that they afford researchers the chance to estimate and evaluate a wide variety of macro models that frequentist econometrics often find challenging. Bayesian times series methods can be extremely useful in DSGE estimation and forecasting. 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. The combination of rich structural models, novel solution algorithms and powerful simulation techniques has allowed researchers to develop the so-called "New Macro-econometrics" (Fernández-Villaverde, 2009 and Fernández-Villaverde *et al.*, 2010). In this framework, the hybrid 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). The hybrid models improve the state-space representation of a DSGE, providing a complete analysis of the data law of motion and better capture the dynamic properties of the DSGE models. 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-VAR à l'Ireland). A different approach called DSGE-VAR à la Del Negro and Schorfheide (2004)[1] is based on the works of DeJong *et al.* (1996) and Ingram and Whiteman (1994). The DSGE-VAR à la Del Negro and Schorfheide combines the VAR representation as an econometric tool for empirical validation with prior information derived from the DSGE model in estimation. The DSGE-FAVAR (Consolo *et al.*, 2009) and the Augmented (B) VAR model (Fernández-de-Córdoba and Torres, 2010) have been introduced to use latent variables in estimating the DSGE. Other models add DSGE model with equations for non-modelled variables, such the DSGE-DFM (Dynamic Factors Model) (Boivin and Giannoni, 2006) and the approach proposed by Schorfheide *et al.* (2010). Instead, in Canova (2012), a hybrid model is proposed to correct the inability to capture the long-run features of the data.

In the literature, there are several works[2] which compares the forecasting performance of DSGE, considering some of the above mentioned hybrid models, against a simple VAR approach. Contrary to them, in this paper, we discuss about the main aspects of the hybrid models. We evidence how they evolved and how they have been used in several comparison exercise. An empirical exercise on US economy shows how it is possible to compare some of these hybrid models to non-DSGE based approaches (VAR, BVAR, Factor Augmented VAR) using the marginal data density[3] (MDD).

The paper is set out as follows. Section 2 discusses the concept of model validation applied to DSGE models. In Section 3, the hybrid models are analyzed in details. Section 4 compares some hybrid models to a classical VAR, a Bayesian VAR, and a Factor-Augmented VAR. Section 5 concludes.

2 Model Validation

The starting point to understand the reason and the evolution of using hybrid models is the concept of model validation. Model validation using DSGE models allows the econometrician to establish a link between structural features of the economy and reduced form parameters, something that was not always possible with the usual large-scale macroeconomic models. Improvements in computational power and the development of new econometric methods are crucial to the popularity of the use of DSGE models. Moreover, very few papers

[1] Fernández-de-Córdoba and Torres (2010) refer to the Ireland methodology as the DSGE-VAR, while the Del Negro and Schorfheide (2004) is called the VAR-DSGE. Some other authors use the term DSGE-VAR for the Del Negro and Schorfheide. To avoid any misunderstanding, in this paper we use DSGE-VAR à l'Ireland and DSGE-VAR à la Del Negro and Schorfheide to refer to these methodologies.

[2] These include 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), Liu *et al.* (2009), Fernández-de-Córdoba and Torres (2010), Lees *et al.* (2011), and Bekiros and Paccagnini (2012) among others.

[3] The marginal data density is defined as the integral of the likelihood function with respect to the prior density of the parameters.

discuss the main aspects of validating the DSGE model, despite its widespread use for forecasting (Edge and Gürkaynak, 2011). The model validation involves selecting a loss function which measures the distance between the set of economic statistics and the set of statistics obtained from the simulated data. Canova (2005) explains that there are essentially four approaches to apply the concept of model validation to DSGE. The first approach is based on an R^2-type measure. In Watson (1993), the economic model is viewed as an approximation of the stochastic data generating process, considering that in the statistical sense the model is not true. The goodness-of-fit (R^2-type) measure is introduced to provide an evaluation of the approximation. The key ingredient of the measure is the size of the error needed to be added to the data generated by the model for the autocovariance implied by the model plus the error to match the autocovariance of the observed data. The second approach measures the distance using the sampling variability of the actual data. Some examples of this approach are the GMM-based of Christiano and Eichenbaum (1992) and of Fève and Langot (1994), the indirect approach of Cecchetti *et al.* (1993) and the frequency domain approach of Diebold *et al.* (1998).[4] The third approach measures the distance by using the sampling variability of the simulated data, such as testing calibration, provide a simple way to judge the distance between population moments and the statistics from a simulated macroeconomic model, as in Gregory and Smith (1991). This method has been used by Soderlind (1994) and Cogley and Nason (1994) to evaluate their DSGE models. The fourth approach measures the distance by using the sampling variability of both actual and simulated data, as discussed in DeJong et al. (1996, 2000), Geweke (1999) and Schorfheide (2000).

The main difference in the model validation, as pointed out by Canova (1994), is between the estimation and the calibration[5] approaches.

Essentially, the estimation approach tries to answer the question: "Given that the model is true, how false is it?", while the calibration approach tries to answer: "Given that the model is false, how true is it?". In the model testing process, an econometrician takes the model seriously as a DGP (data-generating process) and analyzes whether the features of the specification are at variance with the data. A calibrationist takes the opposite view: the model, as a DGP for the data, is false. As the sample size grows, the data-generated by the model will have more variation from the observed data. Statistical models rely on economic theory so loosely that VAR can fail to uncover parameters that are truly structural. This disadvantage may be crucial in policy evaluation exercises, since VAR can exhibit instability across periods when monetary and fiscal policies change. However, the VAR suffers from the 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 the loss of degrees of freedom, leading to inefficient estimates and large out-of-sample forecasting errors.

[4] This approach is an extension of Watson (1993), proposing a spectral analysis.

[5] Kydland and Prescott (1982, 1991) identify calibration as embodying the approach to econometrics articulated and practiced by Firsch (1933a,1933b).

It is possible to overcome this problem by using what have become well-known as "Minnesota" priors (see Doan *et al.*, 1984). The use of "Minnesota" priors has been proposed to shrink the parameters space. The basic principle behind this procedure is that all equations are centered around a random walk with drift. This idea has been modified by Kadiyala and Karlsson (1997) and Sims and Zha (1998). In Ingram and Whiteman (1994), a RBC model is used to generate a prior for a reduced form VAR, as a development of the "Minnesota" priors procedure. The key element is that the dimension of the observable vector exceeds the dimension of the state vector. In Ingram and Whiteman (1996, 2000), a prior is placed on the parameters of a simple linearized DSGE, which is then compared with a Bayesian VAR (BVAR) in a forecasting exercise. Moreover, the calibrated DSGE models are typically too stylized to be taken directly to the data and often yield fragile results, when traditional econometric methods are used for estimation (hypothesis testing, forecasting evaluation) (Smets and Wouters, 2003, and Ireland, 2004).

Due to all these issues, macroeconometrics face the trade-off between theoretical coherence and empirical fit. To solve this dilemma, different attempts at hybrid models have been introduced to bridge the theory of the DSGE model imposed by the restrictions and the estimation on the data. The hybrid models have been introduced for dealing with some of the model misspecifications as well as the trade-off between theoretical coherence and empirical fit (Schorfheide, 2010).

3 Hybrid Models

The main of idea of hybrid models is to improve the state-space representation of a DSGE, providing a complete analysis of the data law of motion and better capture the dynamic properties of the DSGE models.

Consider the following linear state-space representation of a DSGE model with no time-varying parameters (θ):

$$\begin{aligned} y_t &= \Psi_0(\theta) + \Psi_1(\theta)t + \Psi_s(\theta)s_t \qquad (1) \\ s_t &= \Phi_1(\theta)s_{t-1} + \Phi_\epsilon(\theta)\epsilon_t \ , \end{aligned}$$

where y_t is a vector of $(k \times 1)$ observables, such as aggregate output, inflation, and interest rates. This vector represents the measurement equation. Instead, the vector s_t $(n \times 1)$ contains the unobserved exogenous shock processes and the potentially unobserved endogenous state variables of the model. The model specification is completed by setting the initial state vector s_0 and making distributional assumptions for the vector of innovations ϵ_t $(E[\epsilon_t] = 0, E\left[\epsilon_t\epsilon_t^{'}\right] = I$ and $E[\epsilon_t\epsilon_{t-j}] = 0$ for $j \neq 0)$.

There exist essentially two approaches (see Schorfheide, 2010) to building empirical models that combine the restrictions of a DSGE model with a pure statistical model: the additive hybrid models and hierarchical hybrid models.

3.1 Additive Hybrid Models

The additive hybrid model augments the state-space model (1) with a latent process z_t that bridges the gap between data and theory:

$$\begin{aligned} y_t &= \Psi_0(\theta) + \Psi_1(\theta)t + \Psi_s(\theta)s_t + \Lambda_0 + \Lambda_1 t + \Lambda_z z_t \\ s_t &= \Phi_1(\theta)s_{t-1} + \Phi_\epsilon(\theta)\epsilon_t \\ z_t &= \Gamma_1 z_{t-1} + \Gamma_\eta \eta_t. \end{aligned} \tag{2}$$

The process z_t is called measurement error, and blames the collection of data rather than the DSGE model construction for the gap between the theory and the data.

The main examples of additive hybrid models are: the DSGE-AR (Sargent, 1989, Altug, 1989), the DSGE-VAR à l' Ireland (2004), the DSGE-DFM (Boivin and Giannoni, 2006), the Augmented (B)VAR (Fernández-de-Córdoba and Torres, 2010), the DSGE with non-modelled variables (Schorfheide, Sill, and Kryshko, 2010) and the Augmented DSGE for Trends (Canova, 2012).

3.1.1 The DSGE-AR method

Sargent (1989) and Altug (1989) proposed an approach to solving DSGE models, by augmenting the model with unobservable errors as described in equation (2).

A matrix Γ_1 governs the persistence of the residuals; the covariance matrix, $E_t\eta_t\eta_t' = V$, is uncorrelated. In this specification the ϵ_t 's generate the comovements between the observables, whereas the elements of z_t pick up idionsyncratic dynamics which are not explained by the structural part of the hybrid model. However, if we set Ψ_0, Ψ_1 and Λ_z to zero, the DSGE model components can be used to describe the fluctuations of y_t around a deterministic trend path, ignoring the common trend restrictions of the structural model. For instance, Smets and Wouters (2003) estimated their model using this pattern with a two-step procedure. In the first step, the deterministic trends are extracted from the data; in the second step, the DSGE model is estimated using linear detrended observations.

Sargent (1989) and Altug (1989) assume that the measurement errors are uncorrelated with the data generated by the model, hence the matrices Γ_1 and V are diagonal and the residuals are uncorrelated across variables:

$$
\begin{aligned}
\Gamma_1 &= \begin{bmatrix} \gamma_y & 0 & 0 \\ 0 & \gamma_c & 0 \\ 0 & 0 & \gamma_l \end{bmatrix} \\
V &= \begin{bmatrix} v_y^2 & 0 & 0 \\ 0 & v_c^2 & 0 \\ 0 & 0 & v_l^2 \end{bmatrix}.
\end{aligned}
$$

Essentially, this methodology combines the DSGE model with an AR model for the measurement residuals.

3.1.2 The DSGE-VAR à l' Ireland

Ireland (2004) proposed a more general framework for measurement errors, allowing the residuals to follow an unconstrained, first-order vector autoregression. This multivariate approach has the main advantage of imposing no restrictions on the cross-correlation of the measurement errors, allowing it to capture all the movements and co-movements in the data not explained by the DSGE model. The matrices Γ_1 and V are given by:

$$
\begin{aligned}
\Gamma_1 &= \begin{bmatrix} \gamma_y & \gamma_{yc} & \gamma_{yl} \\ \gamma_{cy} & \gamma_c & \gamma_{cl} \\ \gamma_{ly} & \gamma_{lc} & \gamma_l \end{bmatrix} \\
V &= \begin{bmatrix} v_y^2 & v_{yc} & v_{yl} \\ v_{cy} & v_c^2 & v_{cl} \\ v_{ly} & v_{lc} & v_l^2 \end{bmatrix}.
\end{aligned}
$$

This multivariate approach is more flexible and general in the treatment of measurement errors, but some empirical evidence (such as Fernández-de-Córdoba and Torres, 2010) shows the forecast performance of the traditional DSGE-AR outperforms the DSGE-VAR à l'Ireland. Malley and Woitek (2010) propose an extension, allowing for a vector autoregressive moving average (VARMA) process to describe the movements and co-movements of the model's errors not explained by the basic RBC model.

3.1.3 The DSGE-DFM

Macroeconomists have access to large cross-sections of aggregate variables that include measures of sectorial economic activities and prices as well as numerous financial variables. Hybrid models can also be used to link DSGE models with aggregate variables that are not explicitly modelled. Using these additional variables in the estimation potentially sharpens inference about latent state variables:

$$y_t = \Psi_0(\theta) + \Psi_1(\theta)t + \Psi_s(\theta)s_t + z_{y,t} \quad (3)$$

$$s_t = \Phi_1(\theta)s_{t-1} + \Phi_\varepsilon(\theta)\varepsilon_t \quad (4)$$

$$x_t = \Lambda_0 + \Lambda_1 t + \Lambda_s s_t + z_{x,t}, \quad (5)$$

where y_t is the vector of the observable variables that are described by the DSGE model and x_t is a large vector of non-modelled variables.

Since the structure of this model resembles that of a dynamic factor model (DFM), e.g. Sargent and Sims (1977), Geweke (1977), and Stock and Watson (1989), Schorfheide (2010) refers to the system (3) to (5) as an example of a combination of DSGE and DFM (Boivin and Giannoni, 2006). Roughly speaking, the vector of factors is given by the state variables associated with the DGSE model. The processes $z_{y,t}$ and $z_{x,t}$ are uncorrelated across series and model idiosyncratic but potentially serially correlated movements (or measurement errors) in the observables. Moreover, equation (4) links the variables x_t to the DSGE model. This relation generates comovements between the y_t' s and the x_t' s and allows the computation of impulse responses to the structural shocks ϵ_t.

3.1.4 The Augmented (B)VAR

The Augmented (B)VAR (Fernández-de-Córdoba and Torres, 2010) is a combination of the unrestricted VAR with the DSGE model and is conducted by increasing the size of the VAR representation. In this methodology, x_t is a vector of observable economic variables assumed to drive the dynamics of the economy. The structural approach assumes that DSGE models contain additional economic information, not fully captured by x_t. The additional information is summarized by using a vector of unobserved variables z_t. Fernández-de-Córdoba and Torres (2010) explain that these non-observed variables can be total factor productivity, marginal productivity, or any other information given by the economic model, but they do not belong to the observed variable set.

The joint dynamics of (x_t, z_t) are given by the following transition equation:

$$\begin{bmatrix} x_t \\ z_t \end{bmatrix} = \Phi(L) \begin{bmatrix} x_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_t^x \\ \varepsilon_t^z \end{bmatrix}$$

This system cannot be estimated directly since z_t are non-observed, but z_t can be obtained using the DSGE model to create a new variable Z_t, which is used to expand the size of the VAR. It is possible to construct a VAR with the following specification:

$$\begin{bmatrix} x_t \\ Z_t \end{bmatrix} = \begin{bmatrix} \phi_{11}(L) & \phi_{12}(L) \\ \phi_{21}(L) & \phi_{22}(L) \end{bmatrix} \begin{bmatrix} x_{t-1} \\ Z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_t^x \\ \varepsilon_t^z \end{bmatrix}$$

where x_t are the macroeconomic data that the DSGE model seeks to explain and Z_t is a vector derived from the DSGE model. If the model specification is correct, the relation between x_t and Z_t should then capture additional economic information relevant to modelling the dynamics of x_t. A standard unrestricted VAR implies that $\phi_{12}(L) = 0$.

3.1.5 DSGE with non-modelled variables

Schorfheide *et al.* (2010) develop a method of generating a DSGE model-based forecast for variables that do not explicitly appear in the model (non-core variables). They consider the following representation:

$$\begin{aligned} y_t &= \Psi_0(\theta) + \Psi_s(\theta)\varsigma_t \qquad (6) \\ s_t &= \Phi_1(\theta)s_{t-1} + \Phi_\varepsilon(\theta)\varepsilon_t \end{aligned}$$

where eq (6) is the measurement equation, where $\varsigma_t = [s_t', s_{t-1}'M_s'(\theta)]'$ includes the state variables of the model (s_t), the lagged variables for the growth rates, $s_{t-1}'M_s'(\theta)$[6]. To this state space representation, we add an auxiliary regression:

$$z_t = \Lambda_0 + \widehat{s}_{t|t}'\Lambda_s + \xi_t$$

where the $\widehat{s}_{t|t}'$ is derived by the Kalman Filter to obtain estimates of the latent state variables, based on the DSGE model parameter estimates. ξ_t is a variable-specific noise process, $\xi_t = \rho\xi_{t-1} + \eta_t$ and $\eta_t \sim N(0, \sigma_\eta^2)$.

This augmented state-space can be interpreted as a factor model. The factors are given by the state variables of the DSGE model, while the measurement equation associated with the DSGE model describes

[6] In Schorfheide *et al.* (2010), they assume the lagged values of output, consumption, investment, and real wages. These variables are part of the set of the endogenous state variables, in which we have capital and interest rate.

the way in which the core macroeconomic variables load on factors, and the auxiliary regression describes the way in which additional (non-core) macroeconomic variables load on the factors. This representation is a simplified version of the DSGE-DFM since the DSGE with non-modelled variables do not attempt to estimate the DSGE model and the auxiliary regression simultaneously.

3.1.6 The Augmented DSGE for Trends

One of the most discussed problem in using a DSGE model for estimation is its inability to capture the long-run features of the data. Canova (2012) proposes a way to correct these problems using a hybrid model:

$$
\begin{aligned}
y_t &= \Psi_0(\theta) + \Psi_s(\theta) s_t + \Lambda_0 + \Lambda_z z_t \\
s_t &= \Phi_1(\theta) s_{t-1} + \Phi_\varepsilon(\theta)\varepsilon_t \\
z_t &= \Gamma_1 z_{t-1} + \Gamma_2 \overline{z}_{t-1} + \Gamma_\eta \eta_t \\
\overline{z}_t &= \overline{z}_{t-1} + v_t.
\end{aligned} \tag{7}
$$

Depending on the restrictions imposed on the variances of η_t and ν_t, the process z_t is integrated of order one or two and can generate a variety of stochastic trend dynamics.

3.2 Hierarchical Hybrid Models

The second class of hybrid models used for estimating the DSGE model is the hierarchical hybrid.

Consider the following modification of the additive hybrid model:

$$
\begin{aligned}
y_t &= \Lambda_0 + \Lambda_1 t + \Lambda_s s_t \\
s_t &= \Gamma_1 z_{t-1} + \Gamma_\eta \eta_t,
\end{aligned} \tag{8}
$$

where

$$
\begin{aligned}
\Lambda_i &= \Psi_i(\theta) + \eta_i^\Psi, \quad i = 0, 1, s \\
\Gamma_i &= \Phi_i(\theta) + \eta_i^\Phi, \quad i = 1, \epsilon.
\end{aligned} \tag{9}
$$

In this setup, $\Psi_i(\theta)$ and $\Phi_i(\theta)$ are interpreted as restrictions on the unrestricted state-space matrices Λ_i

and Γ_i; instead, the disturbances, η_i^{Ψ} and η_i^{Φ} can capture deviations from the restriction functions $\Psi_i(\theta)$ and $\Phi_i(\theta)$. This kind of hybrid model is related to Bayesian econometrics, since the stochastic restrictions (9) correspond to a prior distribution of the unrestricted state-space matrices conditional on the DSGE model parameters θ.

In the literature, there are essentially two examples of hierarchical hybrid models: the DSGE-VAR (Del Negro and Schorfheide, 2004) and the DSGE-FAVAR (Consolo *et al.*, 2009).

3.2.1 The DSGE-VAR

The basic idea of the DSGE-VAR (Del Negro and Schorfheide, 2004) approach is to use the DSGE model to build prior distributions for the VAR. The starting point for the estimation is an unrestricted VAR of order p:

$$\mathbf{Y}_t = \mathbf{\Phi}_0 + \mathbf{\Phi}_1 \mathbf{Y}_{t-1} + ... + \mathbf{\Phi}_p \mathbf{Y}_{t-p} + \mathbf{u}_t. \tag{10}$$

In compact format:

$$Y = X\Phi + U \tag{11}$$

Y is a $(T \times n)$ matrix with rows Y_t', X is a $(T \times k)$ matrix ($k = 1 + np, p =$number of lags) with rows $X_t' = [1, Y_{t-1}', ..., Y_{t-p}']$, 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:

$$L(Y|\Phi, \Sigma_u) \propto |\Sigma_u|^{-\frac{T}{2}} \exp\left\{-\frac{1}{2} tr\left[\Sigma_u^{-1}\left(Y'Y - \Phi'X'Y - Y'X\Phi + \Phi'X'X\Phi\right)\right]\right\}. \tag{12}$$

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:

$$\begin{aligned}
\Phi^*(\theta) &= \Gamma_{xx}^{*-1}(\theta)\,\Gamma_{xy}^*(\theta) \qquad (13)\\
\Sigma^*(\theta) &= \Gamma_{yy}^*(\theta) - \Gamma_{yx}^*(\theta)\,\Gamma_{xx}^{*-1}(\theta)\,\Gamma_{xy}^*(\theta).
\end{aligned}$$

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:

$$\begin{aligned} \Sigma_u\,|\theta &\sim IW\left((\lambda T \Sigma_u^*(\theta), \lambda T - k, n\right) \qquad (14)\\ \mathbf{\Phi}\,|\Sigma_u, \theta &\sim N\left(\mathbf{\Phi}^*(\theta), \Sigma_u \otimes (\lambda T \Gamma_{XX}(\theta))^{-1}\right), \end{aligned}$$

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, and 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:

$$\begin{aligned} \Sigma_u\,|\theta, Y &\sim IW\left((\lambda+1)\,T\hat{\Sigma}_{u,b}(\theta), (\lambda+1)\,T - k, n\right) \qquad (15)\\ \mathbf{\Phi}\,|\Sigma_u, \theta, Y &\sim N\left(\hat{\mathbf{\Phi}}_b(\theta), \Sigma_u \otimes \left[\lambda T \Gamma_{XX}(\theta) + \mathbf{X}'\mathbf{X}\right]^{-1}\right) \qquad (16)\\ \hat{\mathbf{\Phi}}_b(\theta) &= \left(\lambda T \Gamma_{XX}(\theta) + \mathbf{X}'\mathbf{X}\right)^{-1}\left(\lambda T \Gamma_{XY}(\theta) + \mathbf{X}'\mathbf{Y}\right)\\ \hat{\Sigma}_{u,b}(\theta) &= \frac{1}{(\lambda+1)\,T}\left[\left(\lambda T \Gamma_{YY}(\theta) + \mathbf{Y}'\mathbf{Y}\right) - \left(\lambda T \Gamma_{XY}(\theta) + \mathbf{X}'\mathbf{Y}\right)\hat{\mathbf{\Phi}}_b(\theta)\right], \end{aligned}$$

where the matrices $\hat{\mathbf{\Phi}}_b(\theta)$ 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 (15) and (16) 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 ($\hat{\mathbf{\Phi}}_b(\theta)$ and $\hat{\Sigma}_{u,b}(\theta)$).

In order to obtain a non-degenerate prior density (14), 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}, such that: $\lambda_{MIN} \geq \frac{n+k}{T}; k = 1 + p \times n$, where p =lags and n =endogenous variables.

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

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 $\lambda \to 0$, in the posterior the parameters are not informative, so the DSGE model is of no use in explaining the data. Unfortunately, the posteriors (16) and (15) do not have a closed form and we need a numerical method to solve the problem. The posterior simulator used by Del Negro and Schorfheide (2004) is the Markov Chain Monte Carlo Method and the algorithm used is the Metropolis-Hastings acceptance method. This procedure generates a Markov Chain from the posterior distribution of θ and this Markov Chain is used for Monte Carlo simulations. See Del Negro and Schorfheide (2004) for more details.

The optimal λ is given by maximizing the log of the marginal data density:

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

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

Unfortunately, Del Negro and Schorfheide (2004) do not propose any statistical tool to verify the power of their procedure. Moreover, Del Negro, Schorfheide, Smets and Wouters (2007b) explain "*...the goal of our article is not to develop a classical test of the hypothesis that the DSGE model restrictions are satisfied; instead, we stress the Bayesian interpretation of the marginal likelihood function of* $p(\lambda|Y)$*, which does not require any cutoff or critical values. ...* ".

3.2.2 The DSGE-FAVAR

In the DSGE-FAVAR (Consolo *et al.*, 2009), the statistical representation is a Factor Augmented VAR instead of a VAR model. A FAVAR benchmark for the evaluation of the previous DSGE model will take the following specification:

$$\begin{pmatrix} \mathbf{Y}_t \\ \mathbf{F}_t \end{pmatrix} = \begin{bmatrix} \mathbf{\Phi}_{11}(L) & \mathbf{\Phi}_{12}(L) \\ \mathbf{\Phi}_{21}(L) & \mathbf{\Phi}_{22}(L) \end{bmatrix} \begin{pmatrix} \mathbf{Y}_{t-1} \\ \mathbf{F}_{t-1} \end{pmatrix} + \begin{pmatrix} \mathbf{u}_t^Z \\ \mathbf{u}_t^F \end{pmatrix} \tag{17}$$

where Y_t are the observable variables included in the DSGE model and F_t is a small vector of unobserved factors extracted from a large data-set of macroeconomic time series, which capture additional economic information relevant to modelling the dynamics of Y_t. The system reduces to the standard VAR used to evaluate DSGE models if $\Phi_{12}(L) = 0$.

Importantly, and differently from Boivin and Giannoni (2006), the 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 remaining part of the procedure is implemented in the same way as the DSGE-VAR.

3.3 Hybrid Models Comparison

Hybrid models were introduced to estimate and to perform forecasting analysis. Table 1 compares the different hybrid models, considering the way the DSGE model is estimated. In additive hybrid models, we note the use of the marginal likelihood estimation for the DSGE-AR and for the DSGE-VAR à l'Ireland. Instead, for the other additive, the contribution are performed using Bayesian estimation. In the DSGE-trend, we find estimation both in Bayesian and marginal likelihood terms. The hierarchical hybrid models are estimated using only Bayesian tools. Even if the main issue in using DSGE is the forecasting, not all of the above mentioned hybrid models are presented in a forecasting exercise. The DSGE-AR, the DSGE-DFM, and the DSGE-trend are not shown in a forecasting horse-race. Instead, in the others, a forecasting comparison is implemented to assess the power of the new hybrid models introduced. In all these forecasting exercises, the proposed hybrid model outperforms the other models in the comparison. Moreover, a lot of paper show some of the hybrid models in forecasting application for different countries, using different approaches. For example, Adolfson *et al.* (2008), Ali *et al.* (2008), Ghent (2009), Kolosa *et al.* (2009), Consolo *et al.* (2009), Malley and Woitek (2010), Lees *et al.* (2011), and Bekiros and Paccagnini (2012) among others. An interesting comparison is proposed by Warne, Coenen, and Christoffel (2012) propose the issue of forecasting with DSGE and DSGE-VAR models, focusing on Bayesian estimation of the predictive distribution and its mean and covariance. They introduce the predictive likelihood as a natural model selection in Bayesian literature and a tool for point and density forecasts.

In this paper, we compare the two hierarchical hybrid models, the DSGE-VAR and the DSGE-FAVAR, using the marginal data density, i.e. the integral of the likelihood function with respect to the prior density of the parameters. We use this analysis, in a comparison with a classical VAR, a Bayesian VAR, and a Factor-Augmented VAR, estimating a simple DSGE model.

Table 1: Comparison

	ESTIMATION	MAIN CONTRIBUTION for DSGE	FORECASTING
ADDITIVE			
DSGE-AR Altug (1989), Sargent (1989)	Maximum Likelihood	introduce a univariate measurement (AR)	NO
DSGE-VAR à l'Ireland Ireland (2004)	Maximum Likelihood	introduce a multivariate measurement (VAR)	YES
DSGE-DFM Boivin and Giannoni (2006)	Bayesian	latent variables added to state-space	NO
DSGE with non-modelled variables Schorfheide, Sill, and Kryshko (2010)	Bayesian	auxiliary regressions like measurement equations in a DFM linking non-core variables to state-space of DSGE	YES
Augmented (B) VAR Fernandéz-de-Cordoba and Torres (2010)	Bayesian	non-observed variables from DSGE added to state space	YES
DSGE for trends Canova (2012)	Maximum Likelihood Bayesian	de-trend equation for variables added to state space	NO
HIERARCHICAL			
DSGE-VAR Del Negro and Schorfheide (2004)	Bayesian	VAR representation added by artificial data from DSGE	YES
DSGE-FAVAR Consolo, Favero, and Paccagnini (2009)	Bayesian	FAVAR representation added by artificial data from DSGE	YES

4 An Empirical Exercise

4.1 The Simple DSGE Model

Simple DSGE models with forward-looking features are usually referred to as a benchmarks in the literature. In a 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 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:

$$\tilde{x}_t = E_t[\tilde{x}_{t+1}] - \frac{1}{\tau}(\tilde{R}_t - E_t[\tilde{\pi}_{t+1}]) + (1-\rho_g)\tilde{g}_t + \rho_Z \frac{1}{\tau}\tilde{z}_t \tag{18}$$

$$\tilde{\pi}_t = \beta E_t[\tilde{\pi}_{t+1}] + \kappa[\tilde{x}_t - \tilde{g}_t] \tag{19}$$

$$\tilde{R}_t = \rho_R \tilde{R}_{t-1} + (1-\rho_R)(\psi_1 \tilde{\pi}_t + \psi_2 \tilde{x}_t) + \epsilon_{R,t} \tag{20}$$

$$\tilde{g}_t = \rho_g \tilde{g}_{t-1} + \epsilon_{g,t} \tag{21}$$

$$\tilde{z}_t = \rho_z \tilde{z}_{t-1} + \epsilon_{z,t}, \tag{22}$$

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 $\tilde{\mathbf{Z}}_t = \left(\tilde{x}_t, \tilde{\pi}_t, \tilde{R}_t, \tilde{g}_t, \tilde{z}_t, E_t\tilde{x}_{t+1}, E_t\tilde{\pi}_{t+1}\right)$ and the vector of shocks as $\epsilon_t = (\epsilon_{R,t}, \epsilon_{g,t}, \epsilon_{z,t})$. Therefore the previous set of equations, (18) - (22), can be recasted into a set of matrices $(\mathbf{\Gamma}_0, \mathbf{\Gamma}_1, \mathbf{C}, \mathbf{\Psi}, \mathbf{\Pi})$ accordingly to the definition of the vectors $\tilde{\mathbf{Z}}_t$ and ϵ_t

$$\mathbf{\Gamma}_0 \tilde{\mathbf{Z}}_t = \mathbf{C} + \mathbf{\Gamma}_1 \tilde{\mathbf{Z}}_{t-1} + \mathbf{\Psi} \epsilon_t + \mathbf{\Pi} \eta_t \tag{23}$$

where C is a vector of constants, ϵ_t is an exogenously evolving random disturbance and η_t is a vector of expectations errors, $\left(E_t\left(\eta_{t+1}\right) = \mathbf{0}\right)$, not given exogenously but to be treated as part of the model solution. 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)

$$\begin{aligned} \Delta \ln x_t &= \ln \gamma + \Delta \tilde{x}_t + \tilde{z}_t \\ \Delta \ln P_t &= \ln \pi^* + \tilde{\pi}_t \\ \ln R_t^a &= 4\left[(\ln r^* + \ln \pi^*) + \tilde{R}_t\right] \end{aligned} \tag{24}$$

which can be also casted into matrices as

$$\mathbf{Y}_t = \mathbf{\Lambda}_0(\theta) + \mathbf{\Lambda}_1(\theta)\tilde{\mathbf{Z}}_t + v_t \tag{25}$$

where $\mathbf{Y}_t = (\Delta \ln x_t, \Delta \ln P_t, \ln R_t)'$, $v_t = 0$ and $\mathbf{\Lambda}_0$ and $\mathbf{\Lambda}_1$ are defined accordingly. For completeness, we write the matrices **T, R,** $\mathbf{\Lambda}_0$ and $\mathbf{\Lambda}_1$ as a function of the structural parameters in the model, $\theta = \begin{pmatrix} \ln\gamma, \ln\pi^*, \ln r^*, \kappa, \tau, \psi_1, \psi_2, \\ \rho_R, \rho_g, \rho_Z, \sigma_R, \sigma_g, \sigma_Z \end{pmatrix}'$. Such a formulation derives from the rational expectations solution. The evolution of the variables of interest, $\mathbf{Y}_t$, is therefore determined by (23) and (25) which impose a set of restrictions across the parameters on the moving average (MA)representation (Fernández-Villaverde *et al.,* 2007; Ravenna, 2007) . 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. Policy variables set by optimization - typically included $\widetilde{\mathbf{Z}}_t$ - are naturally endogenous as optimal policy requires some response to current and expected developments of the economy. Expectations at time t for some of the variables of the systems at time $t+1$ are also included in the vector $\mathbf{Z}_t$,whenever the model is forward-looking. Models like (23) can be solved using standard numerical techniques as in Sims (2002) and the solution can be expressed as follows

$$\widetilde{\mathbf{Z}}_t = \mathbf{A}_0 + \mathbf{A}_1 \widetilde{\mathbf{Z}}_{\mathbf{t-1}} + \mathbf{R}\epsilon_t. \quad (26)$$

4.2 Comparison Strategy

In this paper, we compare the marginal data density of the two hierarchical hybrid models, the DSGE-VAR à la Del Negro and Schorfheide (2004) and the DSGE-FAVAR, with the marginal data density of three econometrics models: the Classical VAR, the Bayesian VAR, and the Factor Augmented VAR.

4.2.1 Marginal Data Density

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:

$$\widehat{p}_G(Y) = \left[\frac{1}{n_{sim}} \sum_{s=1}^{n_{sim}} \frac{f(\theta^{(s)})}{L(\theta^{(s)}|Y)p(\theta^{(s)})} \right]^{-1} \tag{27}$$

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:

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

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 $\tau \in (0, 1)$. If the posterior of θ is in fact normal then the summands in eq. (27) are approximately constant.

4.2.2 Classical VAR

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

$$\mathbf{Y}_t = \mathbf{X}_t \mathbf{\Phi} + \mathbf{U} \tag{28}$$

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.

4.2.3 Bayesian VAR

The BayesianVAR, as described in Litterman (1981), Doan *et al.* (1984), Todd (1984), Litterman (1986) and Spencer (1993) has become a widely popular approach to dealing with overparameterization. One of main problems in using VAR models is that many parameters need to be estimated, although some of them may be insignificant. Instead of eliminating longer lags, the Bayesian VAR imposes restrictions on these coefficients by assuming that they are more likely to be near zero than the coefficients on shorter lags. Obviously, if there are strong effects from less important variables, the data can counter this assumption.

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[7].

Formally speaking, these prior means can be written as follows

$$\Phi_i \sim N(1, \sigma^2_{\Phi_i}) \text{ and } \Phi_j \sim N(0, \sigma^2_{\Phi_j}), \tag{29}$$

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

$$S(i, j, m) = [w \times g(m) \times F(i, j)] \frac{\hat{\sigma}_i}{\hat{\sigma}_j}, \tag{30}$$

where

$$F(i, j) = \begin{cases} 1 \text{ if } i = j \\ k_{ij} \text{ otherwise, } 0 \leq k_{ij} \leq 1 \end{cases} \tag{31}$$

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.

[7] Litterman (1981) used a diffuse prior for the constant. The means of the prior are popularly called the "Minnesota Priors" due to the development of the idea at the University of Minnesota and the Federal Reserve Bank at Minneapolis.

4.2.4 Factor Augmented VAR

Bernanke *et al.* (2005) estimate a VAR augmented by factors where the factors can be considered as an "exhaustive summary of the information " of a huge macroeconomics dataset.

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:

$$\mathbf{X}_t = \mathbf{\Lambda}^f \mathbf{F} + \mathbf{\Lambda}^y \mathbf{Y}_t + e_t \tag{32}$$

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 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 (32), we follow the two-step principal components approach proposed by Bernanke *et al.* (2005). More details are reported in Bernanke *et al.* (2005) and Consolo *et al.* (2009). As in Bernanke *et al.* (2005), we partition the matrix $\mathbf{X}_t$ in two categories of information variables: slow-moving and fast-moving. Slow-moving variables (e.g., real variables such as wages or spending) do not respond contemporaneously to unanticipated changes in monetary policy, while fast-moving (e.g., interest rates) respond contemporaneously to monetary shocks. 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". After having determined the number of factors, as suggested by Bai and Ng (2000), 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.

4.3 Data

We use US economy data. Three quarterly time series from 1970:1 to 2010:4 are implemented in estimation. The data for real output growth come from the Bureau of Economic Analysis (Gross Domestic Product-SAAR, Billions Chained 2005$). The data for inflation come from the Bureau of Labor Statistics (CPI-U: All Items, seasonally adjusted, 1982-1984=100). GDP and CPI are taken in first difference logarithmic transformation. The interest rate series are constructed as in, Galì and Gertler (2000); for each quarter the interest rate is computed as the average federal funds rate (source: Haver Analytics) during the first month of the quarter, including business days only. These three time series represent the three equations of the DSGE model. In order to construct the FAVAR we proceed to extract factors from a balanced panel of 109 monthly and quarterly macroeconomic and financial time series, following the dataset built by Stock and Watson (2002). The dataset involves several measures of industrial production, interest rates, various price indices, employment and other important macroeconomic and also financial variables. In this set-up, the number of informational time series N is large (larger than time period T) and must be greater than the number of factors and observed variables in the FAVAR system ($k + M \ll N$). In the panel data used, there are some variables in monthly format, which are transformed into a quarterly data using end-of-period observations. All series have been transformed to induce stationarity. The series are taken the level or transformed into logarithms, first or second difference (in level or logarithms) according to series characteristics (see the Appendix for a description of all series and details of the transformations).

4.4 Estimation of linearized DSGE Models

The DSGE-VAR and the DSGE-FAVAR are implemented following the Del Negro and Schorfheide algorithm, as reported in Del Negro and Schorfheide (2004) and explained in Lees *et al.* (2011). The first step is to specify the prior for the DSGE model parameters. This involves determining the prior distributions of the DSGE parameters and the key parameters of those distributions (such as measures of location and dispersion). Table 2 lists the prior distribution for the structural parameters of the DSGE model which are adopted from Del Negro and Schorfheide (2004)[8]. Second, the DSGE priors have been specified, the model is transformed into state space form, thus linking the theoretical model to the observation equations. Restrictions on the admissible parameter space for the estimation also need to be specified. Using the csminwel procedure of Sims, one estimates the DSGE parameters with the highest posterior probability. The rational expectations solution from csminwel provides the (DSGE restricted) reduced form for the rational expectations model (Sims, 2002). Third, the posterior mode for the DSGE parameters is available, the Metropolis-Hastings algorithm can be used to explore the posterior distribution of θ. Since the VAR parameters – conditional on both θ and λ – are conjugate, it is straightforward to determine the posterior distribution of the VAR parameters. Fourth, the VAR parameters that maximize the posterior distribution are a weighted function of the expected moments from the DSGE model and the moments of the unrestricted VAR. The VAR

[8]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.

parameters at the posterior mode are given from these DSGE and unrestricted VAR moments. Searching over a grid of λ values, one can find the optimal λ value that maximizes the marginal data density.

Table 2: 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

4.5 Estimation results: log of Marginal Data Density

The DSGE-VAR and the DSGE-FAVAR are estimated with a different number of lags on the sample spanning from 1970:1 to 2010:4. The key elements of the priors are estimated using a VAR from a training sample of 10 years of data (1960:1-1969:4).

First, 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 2. 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 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. 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[9]. As suggested by An and Schorfheide (2007) and Adolfson *et al.* (2008), we check that the DSGE-VAR(∞) model is a good approximation of the DSGE model[10].

Table 3 shows the main results related to the DSGE-VAR and to the DSGE-FAVAR implemented using a different number of lags (from 1 up to 4)[11]. 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 marginal data density for the DSGE model specifications computed based on Geweke' s (1999) modified harmonic mean estimator. The Bayes factor (ratio of posterior odds to prior odds) (An and Schorfheide, 2007) helps us to understand the improvement of the log marginal data density of a specific model compared to a benchmark model (M), which in this case are the DSGE-VAR (4) and DSGE-FAVAR(4), since we select four lags as maximum.

[9] For the DSGE-VAR, the lambda grid is given by $\Lambda = \left\{ \begin{array}{c} 0, 0.07, 0.08, 0.10, 0.12, 0.14, 0.16, 0.18, 0.20, 0.25, \\ 0.30, 0{,}35, 0.40, 0.50, 0.60, 0.70, 0.80, 0.9, 1, 2, 5, 10 \end{array} \right\}$. For the DSGE-FAVAR, the lambda grid is given by $\Lambda = \left\{ \begin{array}{c} 0, 0.08, 0.10, 0.12, 0.15, 0.20, 0.25, 0.30, 0{,}35, 0.40, \\ 0.50, 0.60, 0.70, 0.80, 0.9, 1, 2, 5, 10 \end{array} \right\}$. In both lambda intervals, we consider the λ_{MIN} across lags from 1 to 4.

[10] We compare the marginal data density of a DSGE-VAR(∞) with the marginal data density of a DSGE model calculated using the state space representation. For details, see An and Schorfheide (2007).

[11] We select the maximum of four lags according to AIC and BIC on US data for the VAR representation. Instead, for factors, we select one lag according to Information Criteria.

Table 3: Optimal lambda for the DSGE-VAR and DSGE-FAVAR for the sample 1960Q4-2010Q4

	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)	0.07	0.08	0.01	0.14	-649.42	exp[15.08]
DSGE-VAR(2)	0.08	0.18	0.10	1.25	-643.10	exp[8.76]
DSGE-VAR(3)	0.10	0.25	0.15	1.5	-637.47	exp[3.13]
DSGE-VAR(4) (M_1)	0.12	0.25	0.13	1.08	-634.34	1

	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	0.08	0	0	-649.61	exp[15.42]
DSGE-FAVAR(2)	0.10	0.15	0.05	0.5	-642.84	exp[8.65]
DSGE-FAVAR(3)	0.12	0.25	0.13	1.08	-637.41	exp[3.22]
DSGE-FAVAR(4) (M_2)	0.15	0.30	0.15	1	-634.19	1

According to Table 3, the difference $\widehat{\lambda}-\lambda_{MIN}$ is the greatest in the case of a DSGE-VAR(3), and hence its corresponding ratio $\frac{\widehat{\lambda}-\lambda_{MIN}}{\lambda_{MIN}}$ is the greatest too. Looking at the log of the marginal data densities, we notice that the DSGE-VAR(4) model has the minimum value and the Bayes factor evidences how it is not possible to improve the benchmark model (the DSGE-VAR(4)).

The same analysis is repeated for the DSGE-FAVAR with lags from 1 to 4. The DSGE-FAVAR(4) exhibits the greatest difference (0.15), but the ratio $\frac{\widehat{\lambda}-\lambda_{MIN}}{\lambda_{MIN}}$ is greater in the case of the DSGE-FAVAR(3). Looking at the log of the marginal data densities, we notice that the DSGE-FAVAR(4) has the minimum value as shown by the DSGE-VAR(4) [12]. The minimum lambda is used to compute a well-defined marginal data density ($\lambda_{MIN}=0.12$ for the DSGE-VAR(4) and $\lambda_{MIN}=0.15$ for the DSGE-FAVAR(4)). Hence, the MDD analysis helps the econometrician to select the lag order in hybrid models.

Table 4 shows how the DSGE-FAVAR(4) has the minimum log of the marginal data density. The Bayes factor suggests that the benchmark model found in the data (VAR(4)), according AIC and BIC, is not the best possible representation to estimate the DSGE model. The two hybrid models, the DSGE-VAR and, especially, the DSGE-FAVAR, give better performances in terms of log marginal data density. The BVAR performs better than the benchmark model, when we include more than one lag. The same result is provided by the FAVAR.

[12] The DSGE-FAVAR nests the DSGE, VAR and FAVAR models, so using the same Matlab codes we calculated the log marginal data density for each model, changing the settings opportunely.

Table 4: Log of the Marginal Data Density and Bayes Factor for the sample 1960Q4-2010Q4

	$lnp(Y\|M)$	Bayes Factor vs M_3
DSGE	-658.92	exp[9.45]
DSGE-VAR(4)	-634.34	exp[-15.13]
DSGE-FAVAR(4)	**-634.19**	**exp[-15.28]**
BVAR(1)	-654.37	exp[4.9]
BVAR(2)	-643.72	exp[-5.75]
BVAR(3)	-638.85	exp[-10.62]
BVAR(4)	-642.70	exp[-6.77]
VAR(1)	-649.84	exp[0.37]
VAR(2)	-648.62	exp[-0.85]
VAR(3)	-647.61	exp[-1.86]
VAR(4) (M_3)	-649.47	1
FAVAR(1)	-649.61	exp[0.14]
FAVAR(2)	-644.97	exp[-4.5]
FAVAR(3)	-642.32	exp[-7.15]
FAVAR(4)	-639.84	exp[-9.63]

5 Concluding Remarks

Several papers discuss the forecasting accuracy of the hybrid models. In this paper, we illustrate main hybrid models implemented in the recent macroeconometric literature. Two hierarchical hybrid models, the DSGE-VAR and the DSGE-FAVAR, are compared to a classical VAR, a Bayesian VAR, and a Factor Augmented VAR. We compare the DSGE-based models considering the Marginal Data Density (MDD). The MDD helps to select the lag order, and hence the best hybrid model to estimate a simple DSGE model.

References

[1] Adolfson, Malin, Stefan Laséen, Jesper Lindé and Mattias Villani (2008): "Evaluating an Estimated New Keynesian Small Open Economy Model", *Journal of Economic Dynamics and Control* Elsevier,

vol. 32(8), pages 2690-2721.

[2] Ali, Dib, Mohamed Gammoudi, and Kevin Moran, 2008. "Forecasting Canadian time series with the New Keynesian model," Canadian Journal of Economics, Canadian Economics Association, vol. 41(1), pages 138-165, February.

[3] Altug, Sumro (1989): " Time-to-Build and Aggregate Fluctuations: Some New Evidence", International Economic Review, 30(4), pp. 889-920.

[4] An, Sungbae and Frank Schorfheide (2007): "Bayesian Analysis of DSGE Models," *Econometric Reviews*, Taylor and Francis Journals, vol. 26(2-4), pages 113-172.

[5] Bai, Jushan and Serena Ng (2000): "Determining the Number of Factors in Approximate Factor Models", *Econometrica*, 70.

[6] Banbura, Marta, Domenico Giannone and Lucrezia Reichlin (2010): "Large Bayesian vector auto regressions", Journal of Applied Econometrics 25(1):71–92.

[7] Bekiros, Stelios and Alessia Paccagnini (2012): "On the predictability of time-varying VAR and DSGE models", *Empirical Economics*, forthcoming.

[8] Bernanke, Ben S., Jean Boivin and Piotr Eliasz (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.

[9] Boivin, Jean and Marc P. Giannoni (2006): "DSGE Models in a Data-Rich Environment", NBER Working Papers 12772.

[10] Canova, Fabio (1994): "Statistical Inference in Calibrated Models", *Journal of Applied Econometrics*, 9, S123-S144.

[11] Canova, Fabio (1995): "Sensitivity Analysis and Model Evaluation in Simulated Dynamic General Equilibrium Economies", *International Economic Review*, 36, 477-501.

[12] Canova, Fabio and Luca Gambetti (2004): "On the Time Variations of US Monetary Policy: Who is Right?", Money Macro and Finance (MMF) Research Group Conference 2004 96, Money Macro and Finance Research Group.

[13] Canova, Fabio (2005): "Methods for Applied Macroeconomic Research ", Princeton University Press, Princeton.

[14] Canova, Fabio and Luca Sala (2009): "Back to Square One: Identification Issues in DSGE Models", *Journal of Monetary Economics*, 56, 431-449.

[15] Canova, Fabio (2012): "Bridging Cyclical DSGE Models and the Raw Data,", *Manuscript*.

[16] Cecchetti, Stephen G., Pok-Sang Lam and Nelson C. Mark (1993): "The Equity Premium and the Risk-Free Rate : Matching the Moments", *Journal of Monetary Economics*, vol. 31(1), pages 21-45.

[17] Clarida, Richard, Jordi Galí and Mark Gertler (2000): "Monetary Policy Rules And Macroeconomic Stability: Evidence And Some Theory," *The Quarterly Journal of Economics*, MIT Press, vol. 115(1), pages 147-180.

[18] Chari, V.V., Patrick Kehoe and Ellen R. McGrattan (2008): "Are Structural VARs with Long-Run Restrictions Useful in Developing Business Cycle Theory?", *Journal of Monetary Economics*, Elsevier, vol. 55(8), pages 1337-1352, November.

[19] Christiano, Lawrence J. and Martin Eichenbaum (1992): "Current Real Business Cycle Theories and Aggregate Labor Market Fluctuations", *American Economic Review*, 82, 430-450.

[20] Christiano, Lawrence J., Martin Eichenbaum and Charles Evans (1998): "Monetary Policy Shocks: What Have We Learned and to What End?", NBER Working Paper No. 6400

[21] Christiano, Lawrence J., Martin Eichenbaum and Charles Evans (2005): "Nominal Rigidities and the Dynamic effects of a Shock to Monetary Policy", *Journal of Political Economy* 113, 1-45.

[22] Christiano, Lawrence J., Martin Eichenbaum and Robert Vigfusson (2006): "Assessing Structural VARs", in: Acemoglu, D., Rogoff, K. and Woodford, M. (Eds.), NBER Macroeconomics Annual 2006, Cambridge: The MIT Press.

[23] Christoffel, Kai, Günter Coenen and Anders Warne (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.

[24] Consolo, Agostino, Carlo A. Favero and Alessia Paccagnini (2009): "On the Statistical Identification of DSGE Models", *Journal of Econometrics,* 150, 99-115.

[25] Cogley, Timothy F and James M. Nason (1994): "Testing the Implications of Long-Run Neutrality for Monetary Business Cycle Models," *Journal of Applied Econometrics*, vol. 9(S), pages S37-70, Suppl. De.

[26] Cooley, Thomas F. and Edward C. Prescott (1976): "Estimation in the Presence of Stochastic Parameter Variation", *Econometrica*, 44, 167-84.

[27] DeJong, David, Beth Ingram and Charles Whiteman (1996): "A Bayesian Approach to Calibration", *Journal of Business Economics and Statistics*, 14, pp 1-9.

[28] DeJong, David, Beth Ingram and Charles Whiteman (2000): "A Bayesian Approach to Dynamic Macroeconomics," *Journal of Econometrics*, 98, pp 203-223.

[29] Del Negro, Marco and Frank Schorfheide (2004): " Priors from General equilibrium Models for VARs", *International Economic Review*, 45, 643-673.

[30] Del Negro, Marco and Frank Schorfheide (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 2006.

[31] Del Negro, Marco, Frank Schorfheide, Frank Smets and Raf Wouters (2007): "On the Fit of New-Keynesisan Models", *Journal of Business, Economics and Statistics, 25,2, 124-162.*

[32] Diebold, Francis, Lee Ohanian and Jeremy Berkowitz (1998): "Dynamic Equilibrium Economies: A Framework for Comparing Models and Data", *Review of Economic Studies*, 65, 433-452.

[33] Doan, Thomas, Robert Litterman and Christopher Sims (1984): "Forecasting and Conditional Projections Using Realistic Prior Distributions", *Econometric Reviews*, 3, pp 1-100.

[34] Dua, Pami, and Ray, Subhash C., (1995): " A BVAR Model for the Connecticut Economy", *Journal of Forecasting* 14(3), 167-180.

[35] Edge, Rochelle and Refet S. Gurkaynak (2011): "How Useful are Estimated DSGE Model Forecasts?" Finance and Economics Discussion Series 2011-11, Board of Governors of the Federal Reserve System (U.S.).

[36] Fernández-de-Córdoba, Gonzalo and José L. Torres (2010): "Forecasting the Spanish Economy with a DSGE Model: An Augmented VAR Approach", *SERIEs-Journal of the Spanish Economic Association.* Forthcoming.

[37] Fernández-Villaverde Jesus, Juan Rubio-Ramirez, Thomas J. Sargent and Mark W. Watson (2007): "A, B, Cs (and Ds) of Understanding VARs", *American Economic Review*, 97, 3,

[38] Fernández-Villaverde Jesus (2009): "The Econometrics of DSGE Models", NBER Working Paper 14677.

[39] Fernández-Villaverde Jesus, Guerron-Quintana Pablo A, Rubio-Ramírez Juan (2010): The New Macroeconometrics: A Bayesian Approach, in A. O'Hagan and M. West, eds., Handbook of Applied Bayesian Analysis, Oxford: Oxford University Press.

[40] Fève, Patrick and Francois Langot (1994): "The RBC Models through Statistical Inference: An Application with French Data", *Journal of Applied Econometrics*, 9, S11-S37.

[41] Frisch, Ragnar (1933a): "Editorial", *Econometrica*, 1, 1-5.

[42] Frisch, Ragnar (1933b): "Propagation Problems and Impulse Problems in Dynamic Economics", in *Economic Essay in Honour of Gustav Cassel*, October 20th, 1933, London: Allen&Uniwin.

[43] Gelfand, Alan E., Dey, Dipak K. (1994): "Bayesian model choice: asymptotics and exact calculations", *Journal of the Royal Statistical Society Series B* 56(3):501–514.

[44] Ghent, Andra (2009): "Comparing DSGE-VAR Forecasting Models: How Big are the Differences?", *Journal of Economic Dynamics and Control*, 2009, 33:4, pp. 864-882.

[45] Geweke, John (1977): "The Dynamic Factor Analysis of Economic Time Series", in: D.J. Aigner and A.S. Goldberg (eds.), *Latent Variables in Socio-Economic Models*, North-Holland, Amsterdam.

[46] Geweke, John (1999): "Using Simulation Methods for Bayesian Econometric Models: Inference, Development and Communication", *Econometric Reviews* 18:1 , pp. 1-126.

[47] Good, I. J. (1952): "Rational decisions", *Journal of the Royal Statistical Society Series B* 14(1):107–114.

[48] Gregory, Allan W. and Gregor W. Smith (1991): "Calibration in Macroeconomics," Working Papers 826, Queen's University, Department of Economics.

[49] Ingram, Beth, and Charles Whiteman (1994): "Supplanting the Minnesota Prior - Forecasting Macroeconomics Time Series using Real Business Cycle Model Priors", *Journal of Monetary Economics,* 34, pp 497-510.

[50] Ireland, Peter (2004): "A Method for Taking Models to the Data", *Journal of Economic Dynamics and Control*, 28, pp 1205-1226.

[51] Kadiyala, K. Rao and Sune Karlsson (1997): "Numerical Methods for Estimation and Inference in Bayesian VAR-Models", *Journal of Applied Econometrics*, 12(2), 99-132.

[52] King, Robert G. (2000): "The New IS-LM Model: Language, Logic, and Limits", *Federal Reserve Bank of Richmond Economic Quarterly*, 86, pp 45-103.

[53] Kolasa, Marcin, Michal Rubaszek and Pawel Skrzypczynski (2009): "Putting the New Keynesian DSGE Model to the Real-Time Forecasting Test", *European Central Bank Working Paper Series* n. 1110.

[54] Kydland, Finn E. and Edward C. Prescott (1982): "Time to Build and Aggregate Fluctuations", *Econometrica*, 50, 1345-70.

[55] Kydland, Finn E. and Edward C. Prescott (1991): "The Econometrics of the General Equilibrium Approach to Business Cycles", *Scandinavian Journal of Economics*, 93, 161-78.

[56] Lees, Kirdan, Troy Matheson and Christie Smith (2011): "Open Economy DSGE-VAR Forecasting and Policy Analysis: Head to Head with the RBNZ Published Forecasts", *International Journal of*

Forecasting, Elsevier, vol. 27(2), pages 512-528, April.

[57] Litterman, Robert B. (1981): "A Bayesian Procedure for Forecasting with Vector Autoregressions", Working Paper, Federal Reserve Bank of Minneapolis.

[58] Litterman, Robert B. (1986): "Forecasting Accuracy of Alternative Techniques: A Comparison of U.S. Macroeconomic Forecasts: Comment," *Journal of Business & Economic Statistics, American Statistical Association*, vol. 4(1), pages 17-19, January.

[59] Liu, Guangling D., Rangan Gupta and Eric Schaling (2010): "Forecasting the South African economy: a hybrid-DSGE approach", *Journal of Economic Studies*, Emerald Group Publishing, vol. 37(2), pages 181-195, May.

[60] Lucas, Robert, (1972): "Expectations and the Neutrality of Money", *Journal of Economic Theory* 4(2),103-124.

[61] Maffezzoli, Marco (2000): "Human Capital and International Real Business Cycles", *Review of Economic Dynamics*, 3, 137-165.

[62] Malley, Jim and Woitek, Ulrich (2010) "Technology shocks and aggregate fluctuations in an estimated hybrid RBC model," *Journal of Economic Dynamics and Control*, Elsevier, vol. 34(7), pages 1214-1232, July.

[63] Ravenna, Federico (2007): "Vector Autoregressions and Reduced Form Representations of DSGE models", *Journal of Monetary Economics*, 54, 7, 2048-2064.

[64] Rubaszek, Michal and Pawel Skrzypczynski (2008): "On the Forecasting Perfomance of a Small-Scale DSGE Model", *International Journal of Forecasting*, 24, pp. 498-512.

[65] Sargent, Thomas J. and Christopher A. Sims (1977): "Business Cycle Modeling without Pretending to Have Too Much A-Priori Economic Theory", in C. Sims et al. (eds), *New Method in Business Cycle Research*, Federal Reserve Bank of Minneapolis, Minneapolis.

[66] Sargent, Thomas J. (1989):. "Two Models of Measurements and the Investment Accelerator," *Journal of Political Economy*, University of Chicago Press, vol. 97(2), pages 251-87, April.

[67] Schorfheide, Frank (2000): "Loss function-based evaluation of DSGE models," *Journal of Applied Econometrics*, John Wiley & Sons, Ltd., vol. 15(6), pages 645-670.

[68] Schorfheide, Frank, Sill, Keith and Kryshko, Maxym (2010): "DSGE Model-Based Forecasting of Non-Modelled Variables," *International Journal of Forecasting*, Elsevier, vol. 26(2), pages 348-373.

[69] Schorfheide, Frank (2010): " Estimation and Evaluation of DSGE Models: Progress and Challenges", University of Pennsylvania.

[70] Sims, Christopher A. (1980): "Macroeconomics and reality", *Econometrica,* 48(1), pp. 1-48.

[71] Sims, Christopher A. (2002): "Solving Linear Rational Expectations Models", *Computational Economics*, 20 (1-2), 1-20.

[72] Sims, Christopher A. and Tao Zha (1998): "Bayesian Methods for Dynamic Multivariate Models", *International Economic Review*, 39, 949-968.

[73] Smets, Frank and Raf Wouters (2003): "An Estimated Stochastic Dynamic General Equilibrium Model of the Euro Area", *Journal of the European Economic Association*, 1, 1123-75.

[74] Smets, Frank and Raf Wouters (2004): "Forecasting with a Bayesian DSGE Model: An Application to the Euro Area", *Journal of Common Market Studies*, Wiley Blackwell, vol. 42(4), pages 841-867, November.

[75] Smets, Frank and Raf Wouters (2007): "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach", *American Economic Review*, American Economic Association, vol. 97(3), pages 586-606, June.

[76] Soderlind, Paul (1994): "Cyclical Properties of a Real Business Cycle Model," *Journal of Applied Econometrics*, John Wiley & Sons, Ltd., vol. 9(S), pages S113-22, Suppl. De.

[77] Spanos, Aris (1990): "The Simultaneous-Equations Model Revisited: Statistical Adequacy and Identification", *Journal of Econometrics, 44, 87-105.*

[78] Spencer, David E. (1993): " Developing a Bayesian Vector Autoregression Model", *International Journal of Forecasting*, 9:407-421.

[79] Stock, James and Mark Watson (1989), "New indexes of coincident and leading economic indicators", NBER Macroeconomics Annual, 351-393.

[80] Stock, James and Mark Watson (1999) "Forecasting Inflation", *Journal of Monetary Economics,* Vol. 44, no. 2

[81] Stock, James and Mark Watson (2002): "Macroeconomic Forecasting Using Diffusion Indexes", *Journal of Business Economics and Statistics*, XX:II, 147-162.

[82] Stock, James and Mark Watson (2005): "Implications of Dynamic Factor Models for VAR Analysis", NBER Working Paper No. 11467.

[83] Theil, Henry and Arthur S. Goldberg (1961): "On Pure and Mixed Estimation in Economics", *International Economic Review*, 2, pp 65-78.

[84] Theodoridis, Konstantinos (2007): "Dynamic Stochastic General Equilibrium (DSGE) Priors for Bayesian Vector Autoregressive (BVAR) Models: DSGE Model Comparison", Cardiff Business School Working Paper Series, E2007/15.

[85] Todd, Richard M. (1984): "Improving economic forecasting with Bayesian vector autoregression", *Quarterly Review, Federal Reserve Bank of Minneapolis*, issue Fall.

[86] Warne, Anders, Gunter Coenen and Kai Christoffel (2012): "Forecasting with DSGE-VAR Models", Manuscript, European Central Bank.

[87] Watson, Mark (1993): "Measures of Fit for Calibrated Models", *Journal of Political Economy*, 101, 1011-1041.

[88] Woodford, Michael (2003): "*Interest and Prices*", Princeton University Press.

Appendix : The data used to extract factors[13]

Table A1

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

[13] 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.

We describe data used to extract factors in the format adopted by Stock and Watson (2002): series number, long description, short description, transformation code and slow code (The transformation codes are: 1 - no transformation; 2 - first difference; 3 - second difference; 4 - logarithm; 5 - first difference of logarithm and 6 - second difference of logarithm) (The slow codes are: 0 - fast and 1 - slow). The source of the data is the Federal Reserve Economic Data - Federal Reserve Bank of Saint Louis (http://research.stlouisfed.org/fred2/).

Table A1 (continued)

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