

# Can a data-rich environment help identify the sources of model misspecification? \*

Francesca Monti  
Bank of England and CfM

January 30, 2015

## Abstract

This paper proposes a method for detecting the sources of misspecification in a DSGE model based on testing, in a data-rich environment, the exogeneity of the variables of the DSGE with respect to some auxiliary variables. Finding evidence of non-exogeneity implies misspecification, but finding that some specific variables help predict certain shocks can shed light on the dimensions along which the model is misspecified. Forecast error variance decomposition analysis then helps assess the relevance of the missing channels. The paper puts the proposed methodology to work both in a controlled experiment - by running a Monte Carlo simulations with a known DGP - and using a state-of-the-art model and US data up to 2011.

**JEL Classification:**C32, C52

**Keywords:** DSGE Models, Model Misspecification, Bayesian Analysis

## 1 Introduction

In recent years Dynamic Stochastic General Equilibrium (DSGE) models have emerged as important tools for forecasting and policy analysis, thanks to their attractive the-

---

\*I am grateful to Marco Del Negro, Wouter Den Haan, Domenico Giannone, Richard Harrison, Lucrezia Reichlin and Raf Wouters for their insightful comments. Corresponding author: Bank of England, Threadneedle Street, EC2R 8AH London, United Kingdom, francesca.monti@bankofengland.co.uk, +442076015397. *Any views expressed are solely those of the authors and so cannot be taken to represent those of the Bank of England or to state Bank of England policy. This paper should therefore not be reported as representing the views of the Bank of England or members of the Monetary Policy Committee or Financial Policy Committee.*

oretical features and their improved time series fit. As these models become more relevant in policy applications, diagnosing their fit becomes crucial. Indeed, despite the great improvements achieved in the empirical performance of DSGE models, misspecification remains a concern even for richly specified models, as documented for example in Del Negro *et al.* (2007).

A small but growing number of researchers, including Del Negro *et al.* (2007) and Del Negro and Schorfheide (2009), have addressed the issue of misspecification in DSGE models. Del Negro *et al.* (2007) depart from Del Negro and Schorfheide (2004) to develop a method for determining the degree of misspecification in a DSGE model by systematically relaxing the implied cross-coefficient restrictions that the DSGE model imposes on its (approximated) VAR representation. This technique successfully highlights the magnitude of the misspecification, but it does not directly address the following question: what are sources of misspecification?

The recent crisis has brought new relevance to this question: key causes of the crisis and its protraction, such as the housing market, financial markets, the labour market and the fiscal sector, are often very stylized or missing in most policy macro-models. We need tools to assess which channels/transmission mechanisms are relevant at various points in time. With this paper I aim to make progress in this direction. Some authors, namely Curdia and Reis (2010) and Del Negro and Schorfheide (2009), have proposed to address the above question by allowing a more flexible and general correlation structure for the shocks and analysing which interactions among the disturbances are preferred by the data. In a similar vein, Inoue, Kuo and Rossi (2014) introduce in the model additional exogenous processes, which they label wedges, and identify potential misspecification via forecast error variance decomposition and marginal likelihood analyses. I instead propose to exploit a data-rich environment, and focus on the interaction between large number of macroeconomic variables and the model to capture the likely sources and magnitude

of the model misspecification.

If a model is well specified, then no other variable should help predict the variables of the model. This intuition underpins, for example, the work of Evans (1992), who questioned the exogeneity of productivity shocks in RBC models, by running bivariate-Granger causality tests between the productivity shock implied by an RBC model and a wide number of relevant macro variables. He found, for the U.S., that money, interest rates and government spending Granger-caused the productivity shocks of an RBC model and his findings contributed to spurring the use of models that incorporated nominal frictions<sup>1</sup>. In this spirit, I propose a method for testing jointly the exogeneity of the variables of the DSGE with respect to some auxiliary variables. The gist of the method is to model the states of the DSGE and auxiliary variables jointly, imposing the restrictions implied by the DSGE as priors. I then verify how much weight is given to these priors in estimation, choosing the tightness of the priors that maximises the marginal data density of the joint model. Finally, I verify what this implies for the parameters of the joint model: if the driving processes of the model, which are assumed to be exogenous in the DSGE, are found to be Granger-caused by some auxiliary variables, then this suggests some form of misspecification<sup>2</sup>. I also measure how relevant the impact of the missing channels is in the dynamics of the model's driving processes, by analysing the forecast error variance decomposition (FEVD) and verifying how much weight is given to other variables. The paper puts the proposed methodology to the test both in a controlled experiment - by running a Monte Carlo simulations with a known data generating process (DGP from now on) - and using a state-of-the-art model and US data up to 2011.

---

<sup>1</sup>Similar results were found for the U.K. by Holland and Scott (1998).

<sup>2</sup>The approach I propose is close in spirit to the methodology proposed by Giannone and Reichlin (2006) to empirically detect if the shocks recovered from the estimates of a structural VAR are truly structural, which is possible only if the shocks are fundamental. Giannone and Reichlin (2006) show that non-fundamentalness can be detected simply by testing whether the VAR is (weakly) exogenous with respect to potentially relevant additional blocks of variables.

I analyse the properties of the Galì, Smets and Wouters model (2012) and find that, despite the richness of its structure, the auxiliary off-model information can account for up to 11% of the forecast error variance decomposition of its driving processes. The investment specific shock appears to be the most misspecified among the shocks, according to the FEVD. I find the investment shock to be Granger-caused by the corporate bond spread, suggesting that this shock does seem to pick up unmodeled aspects of the financial markets and can be seen as a proxy for overall health of the financial system, as stated in Justiniano, Primiceri and Tambalotti (2010)<sup>3</sup>. In order to verify this conjecture, I extend the model to include financial frictions as in Bernanke, Gertler and Gilchrist (1999), along the lines of Del Negro, Hasegawa and Schorfheide (2014) and find that indeed the misspecification is lower, in particular for the investment specific shock.

The paper is structured as follows. Section 2 discusses the link between misspecification and Granger-causality and Section 3 outlines the methodology. In Section 4 I assess the efficacy of the method I propose in a controlled environment, running a Monte Carlo simulation exercise, while Section 5 presents the results for the Galì, Smets and Wouters (2012) model with US data. Section 6 concludes the paper.

## 2 Misspecification and Granger-causality

The concept of misspecification is broad and multiform: it identifies, in a very general way, all issues relating to the fact that some of the features of the econometrician's model are not in line with the true - but unknown - underlying DGP. In this Section I will attempt to formalize the problem of misspecification in the context of (log-)linearized dynamic stochastic general equilibrium models and discuss why I believe that using Granger causality tests and additional information can help us glean

---

<sup>3</sup>Also variables that are informative about the state of the economy, such as inventories, Granger-cause the investment shock.

information on what aspects of our model are misspecified.

Linear or linearized rational expectations models allow a representation for  $y_t$  in the state space form

$$\begin{aligned} H_F(\theta)E_t s_{t+1} + H_C(\theta)s_t + H_B(\theta)s_{t-1} &= \Psi(\theta)\varepsilon_t \\ y_t &= C(\theta)s_t + \xi_t \end{aligned} \tag{1}$$

where  $s_t$  is an  $n \times 1$  vector of possibly unobserved state variables,  $y_t$  is a  $k \times 1$  vector of variables observed by an econometrician, and  $\varepsilon_t$  is an  $m \times 1$  vector of economic shocks impinging on the states, such as shocks to preferences, technologies, agents' information sets, and  $\xi_t$  is an  $l \times 1$  vector of measurement errors.  $H_B(\theta)$ ,  $H_C(\theta)$ ,  $H_F(\theta)$ ,  $\Psi(\theta)$  and  $C(\theta)$  are functions of the underlying structural parameters of the DSGE model,  $\theta$ . The  $\varepsilon_t$ 's are Gaussian vector white noise satisfying  $E(\varepsilon_t) = 0$ ,  $E(\varepsilon_t \varepsilon_t') = I$ ,  $E(\varepsilon_t \varepsilon_{t+j}') = 0$ , for  $j > 0$ . The  $\xi_t$ 's are Gaussian vector white noise measurement errors satisfying  $E(\xi_t) = 0$ ,  $E(\xi_t \xi_t') = R$ ,  $E(\xi_t \xi_{t+j}') = 0$ , for  $j > 0$ : measurement error can be absent, as is often the case with DSGE models, or affect some or all of the variables, that is  $0 \leq l \leq k$ . The assumption of normality is for convenience and allows us to associate linear least squares predictions with conditional expectations.

Model (1) can be solved with methods suggested by, among others, Blanchard and Kahn (1980) and Sims (2002) and cast in the form of a linear state space model, *i.e.*

$$\begin{aligned} s_{t+1} &= A(\theta)s_t + B(\theta)\varepsilon_{t+1} \\ y_t &= C(\theta)s_t + \xi_t, \end{aligned} \tag{2}$$

where the first equation of (2) is often called transition equation and the second one

is the observation equation. A model of this type can be misspecified in many ways. I will start by discussing the misspecification relative to the transition equation, and will then go on to discuss the role of the observation equation when addressing the issue of misspecification. An intrinsic difficulty when dealing with misspecification in DSGE models is that, while we are interested in assessing the specification of the model in its form (1) and the underlying parameters  $\theta$ , we must work with the solved version of the model (2) to have a meaningful interaction between the model and the data.

The various forms of misspecification that can be found in (log-) linearized DSGE model are listed here.

1. **Order of the approximation.** Dynamic general equilibrium models are often highly non-linear models. In order to deal with them more easily, it is often standard practice to approximate them with a simpler model, *e.g.* by taking a first or second order Taylor approximation of this model around its equilibrium. This paper focuses exclusively on models that are (log-)linearized around their steady state, *i.e.* exclusively linear models of the type:

$$M : \quad s_{t+1} = A(\theta)s_t + B(\theta)\varepsilon_{t+1}$$

where  $M$  indicates the econometrician's model. It could be the case however that the true DGP,  $T$ , is of a higher order of approximation, *e.g.*

$$T : \quad s_{t+1} = A(\theta)s_t + A(\theta)s_t^2 + B(\theta)\varepsilon_{t+1}.$$

The appropriate level of approximation depends a lot on the problem we want to address: if, for example, we are interested in big shocks that take us very far from steady state, then the linear approximation is clearly ill-suited. Similarly,

if we are interested in accounting for the agents' attitude towards risk in a more rich and realistic manner, e.g. with recursive (Kreps-Porteus-Epstein-Zin-Weil) preferences, then a first order approximation simply will not do. It can be shown that the first order approximation of such model simply collapses to the standard model with constant relative risk aversion preferences, and even in the second order approximation the risk aversion coefficient enters only as a constant. In order for risk aversion and precautionary behavior to play a role in the model, we would need a third order approximation.

2. **Model dynamics.** Let us assume that the true DGP has this form

$$T : \quad s_{t+1} = A(\theta)s_t + B(\theta)\varepsilon_{t+1},$$

while the econometrician's model is

$$M : \quad s_{t+1} = A_M(\theta)s_t + B_M(\theta)\varepsilon_{t+1},$$

where  $A_M$  and  $B_M$  imply different cross-equation restrictions with respect to  $A$  and  $B$ . For example, in the true model agents might have habits in consumption in their utility function. Habits imply that the marginal utility of consumption depends on both *past* and future consumption, while with no habits it should depend exclusively on the current and future consumption. Thus habits entail an additional degree of endogenous persistence.

3. **Missing variables/channels.** Let us now assume that the econometrician's model of the economy is

$$M : \quad s_{t+1} = A_M(\theta)s_t + B_M(\theta)\varepsilon_{t+1},$$

but that the true DGP is actually

$$T : \quad \tilde{s}_{t+1} = \begin{bmatrix} s_{t+1} \\ \sigma_{t+1} \end{bmatrix} = \begin{bmatrix} A_{11}(\theta) & A_{12}(\theta) \\ A_{21}(\theta) & A_{22}(\theta) \end{bmatrix} \tilde{s}_t + \begin{bmatrix} B_1(\theta) \\ B_2(\theta) \end{bmatrix} \varepsilon_{t+1}.$$

$\sigma_t$  is a vector of additional variables important in determining the dynamics of  $s_t$ , but ignored in the econometrician's model. As an example consider a situation in which, in the true model, leisure is in the utility function and agents supply labor according to an optimization problem, while the econometrician uses a simpler model, in which labor supply is completely inelastic. In that case, hours worked and wage are an important determinant of the dynamics of the data, but are completely left out of the econometrician's model. Notice that also the type of misspecification identified in the item 1) can be represented in this form, *i.e.* by having the high-order terms appear as additional variables in the vector  $\sigma_t$ .

4. **Dynamics of the model's driving processes.** In dynamic *stochastic* general equilibrium models some of the states  $s_t$  are represented by the stochastic driving processes of the model (e.g. the technology shock and the investment shock). These are most often modeled as AR(1) processes. It could be the case however a) that in the "true" model the shocks are contemporaneously correlated or b) that the part of the shocks underlying the driving processes are anticipated. Clearly, in both these cases, the model's implied dynamics would be very different. Formally misspecification 4.a) can be represented as the misspecification identified in item 2, while misspecification 4.b) has the same representation as the misspecification identified in item 3 (see e.g. Schmitt-Grohe and Uribe, 2008).

We identified the possible causes of misspecification separately, but obviously various



combinations of the above problems are possible.

Models (1) and (2) imply another layer of complexity in that the states are most often assumed to be unobservable and, thus, need to be inferred using a set of observables  $Y_t$ . Therefore, even in the implausible case in which we knew the “true” model for the dynamics of the states, we still face the non-trivial problem of choosing the right variables to include in our set of observables  $Y_t$ . Canova, Ferroni and Matthes (2014) propose two methods for choosing the the observables of a DSGE model, based, respectively, on the optimisation of parameter identification and on minimisation of the informational discrepancy between the singular and non-singular model.

The conditions that ensure that we can recover  $s_t$  and  $\varepsilon_t$ , given the current values of  $Y_t$  and their (possibly infinite) history are the following:

- $\xi_t = 0 \quad \forall t$ , i.e. that there is no measurement error
- the eigenvalues of  $(A - BD^{-1}\Gamma)$  are in the unit circle, with  $\Gamma = CA$  and  $D = CB$ . This is the so called “Poor Man’s invertibility condition” in Fernandez-Villaverde, Rubio-Ramirez, Stock and Watson (2007).

If any of these two conditions is not met, then it will not be possible to recover the values of the states correctly, even with an infinite amount of past data, *i.e.* the states will be mismeasured to some degree. In this case the econometrician has two options. In case there is no measurement error, he can attempt to find a set observables, which deliver  $\Gamma$  and  $D$  such that the eigenvalues of  $(A - BD^{-1}\Gamma)$  are in the unit circle. An alternative, especially useful in the case of presence of measurement error, is to expand the observables set as in Giannone and Reichlin (2006) and in Boivin and Giannoni (2006). In particular, see Giannone and Reichlin (2006) for an in-depth discussion of how additional information can help better identify the latent states. Notice however that, in principle, if the econometrician’s

model is such that  $\max |eig(A - BD^{-1}\Gamma)| \leq 1$ , then the current and past values of the observables  $Y_t$  are sufficient statistics for the recovery of the states and adding new variables will not add any additional information.

It is clear that, if the econometrician's model has measurement error or does not respect the poor man's invertibility condition, then the finding that some non-model variables Granger-cause the estimated (but surely mismeasured) states of the model is not clearly interpretable. It could be due to the fact that the "true" states are actually non-exogenous, but it could also be the mismeasurement in the states that determines the Granger-causality. In order to avoid this confusion, we work exclusively with models that do not have measurement error<sup>4</sup> and that respect the "Poor man's invertibility condition." In this way, the past and current values of  $Y_t$  are sufficient statistics for determining  $s_t$  and finding Granger causality cannot be attributed to the mismeasurement of the states.

The types of misspecifications 2) and 4.a) can be identified simply relaxing the cross-equation restrictions implied by the model and verifying whether they are rejected by the data as in Del Negro *et al.* (2007) or Del Negro and Schorfheide (2009). Working with Granger-causality<sup>5</sup> and additional information allows us to have a sense of whether our model displays misspecification types 1), 3) and 4.b). Since these three types of misspecification can be all represented in the same way (both higher order terms and anticipated shocks can be represented as additional variables in the state vector), finding Granger-causality does not guide us in determining which of the three types of misspecifications we are dealing with. There is no

---

<sup>4</sup>This is quite common in the DSGE literature, e.g. Smets and Wouters (2007) and Del Negro, Hasegawa and Schorfheide (2014).

<sup>5</sup>A variable  $y_t$  fails to Granger-cause another variable  $x_t$  if  $y_t$  does not help to forecast  $x_t$ . The testable implication of Granger causality we test for is the following: the VAR describing  $x_t$  and  $y_t$  is lower triangular.

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = A(L) \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ u_t \end{bmatrix} \quad (3)$$

where  $\varepsilon_t$  and  $u_t$  white noise, mean zero and serially uncorrelated.  $y_t$  fails to Granger-cause  $x_t$  if and only if the filter  $A(L)$  is lower-triangular.

obvious way to disentangle the misspecification due to omitted variables and that due to not having properly modeled anticipated shocks, because we do not know the true model. What we will be able to say is that there is misspecification and that it involves certain missing variables or missing mechanisms, and we can speculate exactly what is missing. For example, forward-looking variables such as the stock price might Granger-cause real variables because they capture anticipated information on real shocks, while finding evidence of money Granger-causing technology as in Evans (1992) is rather an indication that some crucial economic mechanism has not been modeled properly. Investigating Granger-causality can help us uncover what the sort of mechanisms the shocks are proxying.

Evans (1992) exploited this link between misspecification and Granger-causality to study the empirical validity of RBC models. Running bivariate Granger-causality tests between the productivity shock in a RBC model and a number of key US macroeconomic variables, Evans found strong evidence that productivity shocks were not exogenous with respect to money, interest rates and government spending. Similar results were found for the U.K. by Holland and Scott (1998). Basing the analysis on bivariate Granger-causality tests, however, leaves out completely the possibility of joint dynamics: this could potentially give rise to false-positives in the search for misspecification. For example, in models with more than one shock, it might be the case that letting the shocks interact dynamically among each other reduces amount of Granger-causality found with respect to non-modeled variables. For this reason, I propose to analyze the problem in a joint statistical framework; I propose to model the variables of the DSGE model and the auxiliary variables jointly and to impose the priors coming from the DSGE where possible. Then I suggest to systematically relax the restrictions imposed by the DSGE on the dynamics of the variables of the model among each other and on the zero-restrictions imposed by the DSGE on the coefficients that determine the effect of the auxiliary variables on

the DSGE model.

### 3 The Methodology

The gist of the methodology is to model the states of the DSGE model and auxiliary variables jointly as a Bayesian vector autoregression (BVAR). We define our prior for the BVAR based on the belief that the model is well-specified. We then optimally determine tightness for the DSGE priors on their own dynamics and on the zero-restrictions on the auxiliary variables. We then obtain a posterior distribution for the parameters and can then verify which coefficients (or sets of coefficients) are significantly different from zero, that is which auxiliary variables Granger-cause the states of the DSGE model.

The solution of a linear or linearized rational expectations model has the following representation for  $Y_t$  in the state space form

$$\begin{aligned} s_{t+1} &= A(\theta)s_t + B(\theta)\varepsilon_{t+1} \\ Y_t &= C(\theta)s_t + \xi_t \end{aligned} \tag{4}$$

where  $s_t$  is an  $n \times 1$  vector of unobserved state variables,  $Y_t$  is a  $k \times 1$  vector of variables observed by an econometrician, and  $\varepsilon_t$  is an  $m \times 1$  vector of economic shocks impinging on the states, such as shocks to preferences, technologies, and agents' information sets.  $A(\theta)$ ,  $B(\theta)$  and  $C(\theta)$  are functions of the underlying structural parameters of the DSGE model. The  $\varepsilon_t$ 's are Gaussian vector white noise satisfying  $E(\varepsilon_t) = 0$ ,  $E(\varepsilon_t\varepsilon_t') = I$ ,  $E(\varepsilon_t\varepsilon_{t+j}') = 0$ , for  $j > 0$ . The assumption of normality is for convenience and allows us to associate linear least squares predictions with conditional expectations. For notational simplicity we will drop the indication that the matrices A, B, etc. are function of the structural parameters  $\theta$ .

Ideally, the first best solution would be to estimate the joint model and the underlying structural parameters jointly, in the spirit of Del Negro and Schorfheide (2004). However the computational burden of such an approach in a data-rich environment is huge. So I make the following simplification: I estimate the parameters of the model at a first stage. I obtain the solution of the model based on a draw from the posterior distribution or simply choosing the posterior mode<sup>6</sup>, and then I filter the unobservable states of the DSGE with the Carter and Kohn algorithm (1994), which accounts for the uncertainty in the estimation of the states and then treat them as observables.

I am interested in linking the model to a large set of variables, which are commonly available to policy institutions and market participants alike, e.g. industrial production or the flow of funds. Let us consider a panel of additional variables that carry information on current economic conditions. We define by  $X_t = (x_{1,t}, \dots, x_{N,t})'$  the vector of these auxiliary variables. We can model the link between the DSGE and the set of auxiliary variables  $X_t$  as a VAR of order  $p$ , in which the states of the DSGE model enter as exogenous variables. Indeed, if the model is well specified, the states must be exogenous with respect to the auxiliary variables, but it is natural to think that the shocks of the model are driving also the variables that are not explicitly modelled in the DSGE.

$$\begin{aligned} X_t &= \Pi_1 X_{t-1} + \dots + \Pi_p X_{t-p} + \Gamma_0 s_t + \xi_t \\ &= \Pi_1 X_{t-1} + \dots + \Pi_p X_{t-p} + \Gamma_0 A s_{t-1} + \Gamma_0 \Phi \varepsilon_t + \xi_t \end{aligned} \tag{5}$$

We estimate the BVARX( $p$ ) using Bayesian techniques. In brief, we estimate this model using Bayesian methods, based on a modified Litterman prior, *i.e.* we impose

---

<sup>6</sup>For example, in the Monte Carlo exercise of the next Section, each replication involve a draw of  $\theta$  from the posterior distribution.

the following moments for the prior distribution of the coefficients:

$$\begin{aligned}
llE \left[ (\Pi_k)_{i,j} \right] &= \begin{cases} \delta_i, & j = i, k = 1 \\ 0, & \text{otherwise} \end{cases}, & V \left[ (\Pi_k)_{i,j} \right] &= \frac{\lambda^2 \sigma_i^2}{k^2 \sigma_j^2} \\
E \left[ (\Gamma_{0k})_{i,j} \right] &= 0, & V \left[ (\Gamma_k)_{i,j} \right] &= \frac{\sigma_{x_i}^2}{\sigma_{x_j}^2}
\end{aligned} \tag{6}$$

where  $\delta_i = 1$  if the variable is highly persistent, while  $\delta_i = 0$  for variables characterized by a substantial degree of mean-reversion. The lag length  $p$  and the informativeness of the prior can be selected by maximising the marginal data density of the model as in Carriero, Clark and Marcellino (2013) and Giannone, Lenza and Primiceri (2015).

If model (4) is not misspecified, then it should be sufficient to represent the data accurately and the auxiliary variables  $X_t$  should not Granger-cause the exogenous driving forces of the model. Consider stacking equations (4) and (5) into the following model:

$$\begin{bmatrix} s_t \\ \tilde{X}_t \end{bmatrix} = \underbrace{\begin{bmatrix} A^*(\theta) & 0_{n,N} \\ \Gamma & \Pi \end{bmatrix}}_{M^*} \begin{bmatrix} s_{t-1} \\ \tilde{X}_{t-1} \end{bmatrix} + \underbrace{\begin{bmatrix} H^*(\theta) & 0_{n,m} \\ \Phi & V \end{bmatrix}}_{N^*} \begin{bmatrix} \varepsilon_t \\ \xi_t \end{bmatrix}, \tag{7}$$

where  $\tilde{X}_t = [\tilde{X}'_t \tilde{X}'_{t-1} \dots \tilde{X}'_{t-p+1}]$  and the coefficient matrices are defined accordingly.  $\theta$  can be chosen to be the posterior mode or a draw from the posterior distribution.

The key idea is to use this model as a prior for the joint time series model of

$w_t = \begin{bmatrix} s_t' & X_t' \end{bmatrix}'$ , namely

$$\begin{bmatrix} s_t \\ \tilde{X}_t \end{bmatrix} = \underbrace{\begin{bmatrix} A_{11}(\theta) & A_{12} \\ A_{21} & A_{22} \end{bmatrix}}_M \begin{bmatrix} s_{t-1} \\ \tilde{X}_{t-1} \end{bmatrix} + \underbrace{\begin{bmatrix} H(\theta) & 0_{n,m} \\ \Phi & V \end{bmatrix}}_N \begin{bmatrix} \varepsilon_t \\ \xi_t \end{bmatrix} \quad (8)$$

We follow Sims (2008) and postulate that, conditional on  $\theta$ ,  $(M, N)$  have a distribution centered around  $(M^*(\theta), N^*(\theta))$ . We can express these beliefs with a set of dummy observations. We will focus on three types of priors in particular. In the first two types, the  $j$ -th observation consists of a pair  $w_1^j$  and  $w_0^j$  of current and lagged values for the  $w_t$  vector. This pair, combined, gives rise to a term in the log prior density as if  $w_1^j$  were an observation on a  $N(M^*(\theta)w_0^j, N^*(\theta)N'^*(\theta))$ .

With a Type 1 prior we want to impose the belief that  $M$  is close to  $M^*(\theta)$ , *i.e.* that

$$w_1^j = \lambda_j M^*(\theta) w_0^j. \quad (9)$$

There are more ways to specify  $w_1^j$  and  $w_0^j$  for Type 1 prior; we choose the following. Let  $QTQ' = M^*(\theta)$  be the Schur decomposition of  $M^*(\theta)$ : then we define  $w_1^j = QT_{.j}$  and  $w_0^j = Q_{.j}$ , where, for example,  $T_{.j}$  is the  $j$ -th column of  $T$ . This sort of prior enforces belief (9).  $\lambda_j$  is a scaling factor that governs the tightness of the prior, as discussed in more detail below. The dummy observations for Type 2 prior are chosen to impose the belief that  $NN'$  is close to  $N^*N'^*$  and are of the form  $w_1^j = N^*(\theta)e_j$  and  $w_0^j = 0$ , where  $e_j$  is a vector containing all zeros but a one in the  $j$ -th position.

Type 1 and 2 dummy observations are particularly convenient, because they lead to a conjugate prior. Since they yield terms in the log likelihood that are of the same form as the pdf of  $w_t$  in the sample period, Type 1 and 2 dummy observations can, in principal, simply be added to the actual data when estimating the likelihood

and there exists an analytical solution for the problem of maximizing the posterior density or the marginal data density when selecting the tightness. Type 1 priors can also be tightened or loosened by scaling up or down the size of  $w_1^j$  and  $w_0^j$  (using the scaling factor  $\lambda_j$ ), with larger dummy observations implying greater precision of beliefs. An interesting aspect of Type 1 dummy observations is that they make it possible to assign different weight to different dummies and therefore impose different elements of the prior with different tightness. Here I allow for two different scaling factors, one associated to the prior that  $A_{11}$  is close to  $A^*$  ( $\lambda_1$ ) and one associated with the prior that  $A_{12}$  is close to 0 ( $\lambda_2$ ).  $\lambda_1 \rightarrow \infty$  and  $\lambda_2 \rightarrow \infty$  imposes the DSGE prior exactly, such that the dynamics of  $w_t$  implied by priors (12) will coincide with that implied by the DSGE model (7), while  $\lambda_1 \rightarrow 0$  and  $\lambda_2 \rightarrow 0$  gives rise to an uninformative prior. It is also possible to experiment with various combinations of  $\lambda_1$  and  $\lambda_2$ , including the Curdia and Reis (2012) style experiment, in which we loosen the dynamics among the shocks while leaving  $\lambda_2$  very tight.

The precision of prior beliefs about the connection of M and N with  $M^*(\theta)$  and  $N^*(\theta)$  can be controlled in different ways depending on the type of prior. For each of the two types, the repetition of the dummy observations will tighten the prior.

Once we have selected the tightness of  $\lambda_1$  and  $\lambda_2$ , we can obtain the posterior distribution for M via Gibbs sampling and then verify for which coefficients in  $A_{21}$  the 90% credible set does not include zero<sup>7</sup>. We can also measure how relevant the impact of the missing channels is in the dynamics of the model's driving processes, for example by analysing a forecast error variance decomposition (FEVD) and verifying how much weight is given to other variables.

---

<sup>7</sup>The 90% credible set is the narrowest interval around the mode of the posterior distribution such that the probability that a coefficient lies within that interval is 90%.



## 4 The Monte Carlo Exercise

### 4.1 The Simulation Design

To provide further insights on the methodology, I investigate its workings, and show its effectiveness, in a simulation environment. The number of replications for the Monte Carlo experiment is 2000. The data generating process (DGP) is a variant of the Galí, Smets and Wouters (2012) model that incorporates financial frictions as in Bernanke, Gertler and Gilchrist (1999). The Galí, Smets and Wouters (2012; GSW henceforth) model reformulates the well known Smets-Wouters (2007; henceforth SW) framework by embedding the theory of unemployment proposed in Galí (2011a,b). The main difference between GSW and SW is the explicit introduction of unemployment, and the use of a utility specification that parameterizes wealth effects, along the lines of Jaimovich and Rebelo (2009). I add the financial frictions building on the work of Del Negro, Hasegawa and Schorfheide (2014). In this setup, banks collect deposits from households and lend to entrepreneurs, who are hit by idiosyncratic shocks to their net wealth. The entrepreneurs use a mix of these funds and their wealth to acquire physical capital, but because of their idiosyncratic shocks, their revenues may be too low to repay the loans. The banks therefore protect themselves charging a spread over the deposit rate, which will be a function of the entrepreneurs' leverage and riskiness.

Each replication in the Monte Carlo exercise starts by taking a draw  $\theta$  from the posterior distribution of the model and simulating data from the solution of the model based on  $\theta$ . The model generates data on the following 9 variables: GDP growth, consumption growth, investment growth, real wage growth, inflation, employment, the unemployment rate, the policy rate and a measure of spreads. I present the main log-linearized equations of the model in Appendix A and refer to Galí, Smets and Wouters (2012) for an in depth discussion of the model and

Del Negro, Hasegawa and Schorfheide (2014) for the part on the financial frictions. We then assume that the econometrician estimates a prototypical new-Keynesian model, such as the one in An and Schorfheide (2007), which is characterized by the following log-linearized equations.

$$\begin{aligned}
y_t &= E_t y_{t+1} + g_t - E_t g_{t+1} - \frac{1}{\tau}(r_t - E_t \pi_{t+1} - E_t z_{t+1}) \\
\pi_t &= \kappa(y_t - g_t) + \beta E_t \pi_{t+1} \\
r_t &= \rho_r r_{t-1} + (1 - \rho_r)\phi_\pi \pi_t + (1 - \rho_r)\phi_y * (y_t - g_t) + mpt
\end{aligned} \tag{10}$$

Deviations of output from its steady state are denoted by  $y$ ,  $\pi$  is inflation and  $r$  is the interest rate, while  $z_t$ ,  $g_t$  and  $mpt$  are, respectively, the technology shock, the government spending shock and the monetary policy shock. The shocks are assumed to be AR(1) processes. This model uses data on GDP growth, inflation and the interest rate to estimate the history of the shocks driving the model, so part of the data generated by our DGP is unused. I treat the data on consumption growth, investment growth, real wage growth, employment, the unemployment rate and the spread as the auxiliary variables  $X_t$  in our methodology. To these I add a completely uninformative series (an AR(1) process), to make sure the methodology can correctly distinguish between relevant and irrelevant additional information. I estimate a BVARX as (5) on these variables, imposing a lag length of  $p = 1$  and selecting the tightness of the prior by maximising the marginal data density, following the methodology proposed in Giannone, Lenza and Primiceri (2015)<sup>8</sup>.

I draw from the posterior distribution for the GSW with financial frictions and generate data from the model implied by those parameters. I then use the econometrician's model (10) to filter the simulated data for GDP growth, inflation and

---

<sup>8</sup>As discussed in the previous section, it is also possible to make the lag length an argument of the maximisation problem as shown in Carriero *et al.* (2013), who however show that one quarter (or 3 or 4 months for a monthly model) is preferred for a big part of the sample. Also Giannone *et al.* (2015) fix the lag length rather than choosing it optimally.

the interest rate and produce some estimates of the shock processes  $s_t = [z_t' b_t' m p_t']'$ .

I then stack these with the auxiliary variables  $X_t$  and estimate a VAR for:

$$\begin{bmatrix} s_t \\ X_t \end{bmatrix} = \underbrace{\begin{bmatrix} A_{11}(\theta) & A_{12} \\ A_{21} & A_{22} \end{bmatrix}}_M \begin{bmatrix} s_{t-1} \\ X_{t-1} \end{bmatrix} + \underbrace{\begin{bmatrix} H(\theta) & 0_{n,m} \\ \Phi & V \end{bmatrix}}_N \begin{bmatrix} \varepsilon_t \\ \xi_t \end{bmatrix}. \quad (11)$$

As described in the previous section, the priors used for estimating the BVAR are centered on the belief that the model is well-specified, *i.e.* that  $A_{11} = A^*(\theta)$  and  $A_{12} = 0$ . We use two different scaling factors,  $\lambda_1$  and  $\lambda_2$ , for these priors and select them by choosing the values that maximize the marginal likelihood of model (11).

## 4.2 The Simulation Results

The first way of assessing whether there is misspecification is to look at the forecast error variance decomposition (FEVD) of the econometrician's model's shocks in (11). For each Monte Carlo replication I compute, for the mean of the posterior distribution of  $M$  in model (11), how much of the long-run (30 quarters) variance of each shock is explained by its own dynamics, by other shocks, and by the auxiliary variables. If the model were well specified, then the auxiliary variables and the other shocks should not contribute at all to the FEVD. We know however that the econometrician's model is misspecified along several dimensions, so, unsurprisingly, I find that other variables will contribute a lot to the FEVD. Table 1 presents the mean FEVD across all Monte Carlo replications. To give an idea of the distribution of the results across different replications, Figures 4.2-4.2 show the distribution of outcomes for the contribution of each auxiliary variable to the FEVD of each of the econometrician's model's shocks. The results clearly point out that the biggest misspecification in the econometrician's model is the absence of capital and therefore investment, which in the DGP has an important role both intertemporally and

Table 1: Forecast error variance decomposition at a 30-quarter horizon, in percentage points

	z	g	mp	$\Delta C$	$\Delta I$	$\Delta w$	N	UR	SP	AR
z	69.81	6.52	4.84	1.65	12.23	0.48	1.15	1.67	0.13	1.02
g	2.45	61.22	1.89	4.5	20.10	1.20	3.64	2.32	0.82	1.68
mp	4.96	5.56	68.58	1.87	14.10	0.58	1.14	1.52	0.57	0.98

intratemporally, for households and firms. The government spending shock, which in the econometrician’s model is a sort of catch-all for various demand shocks, is also affected by labour market variables and consumption growth. Finally notice that virtually no contribution to the FEVD of any of the shocks comes from the spurious AR process we had added.

For each of the 2000 Monte Carlo replications, it is possible to study the posterior distribution of the coefficients in the matrix  $A_{12}$  of model (11) and ask whether its 90% credible set contains zero. Finding that it does not, is suggestive of non-exogeneity of the shocks with respect to the auxiliary variables in question. For example, Figure 4.2 reports, for a sample replication, the posterior distribution of the coefficients in the row of the  $A_{12}$  matrix that impinge on  $g_t$ . In line with the FEVD, it is clear that both investment and labour market variables are not exogenous to this shock.

## 5 The Application

I now apply the proposed methodology to the Galí, Smets and Wouters (2012) model, which reformulates the well known Smets-Wouters (2007) framework by embedding the theory of unemployment proposed in Galí (2011a, 2011b). I present the main log-linearized equations of the model in Appendix A and refer to Galí, Smets and Wouters (2012) for an in depth discussion of the model. The model includes the following 8 shocks:

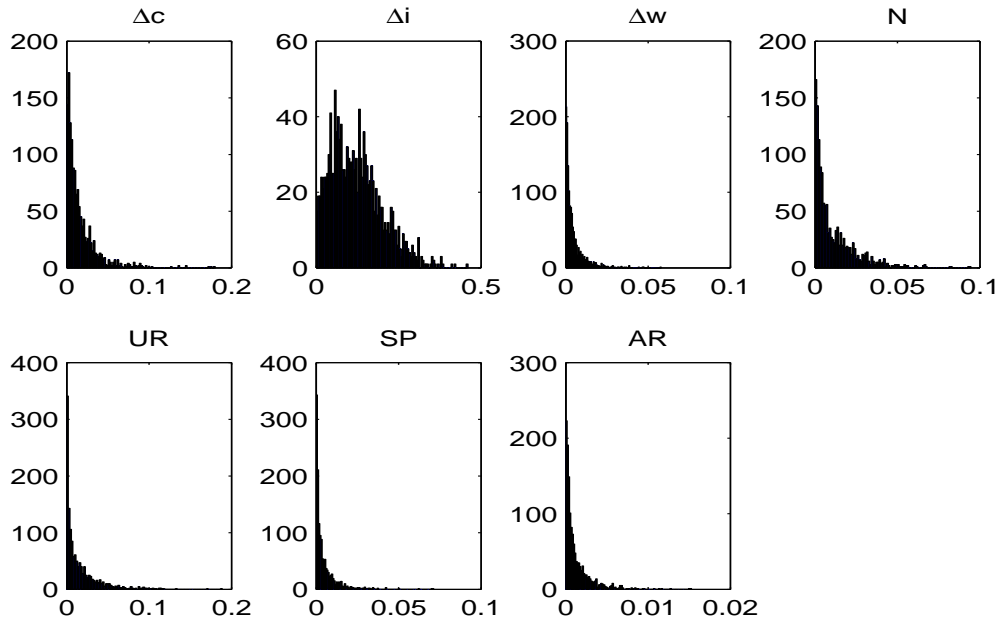


Figure 1: Distribution of FEVD contributions of the auxiliary variables to the technology shock  $z_t$  for the 2000 Monte Carlo replications

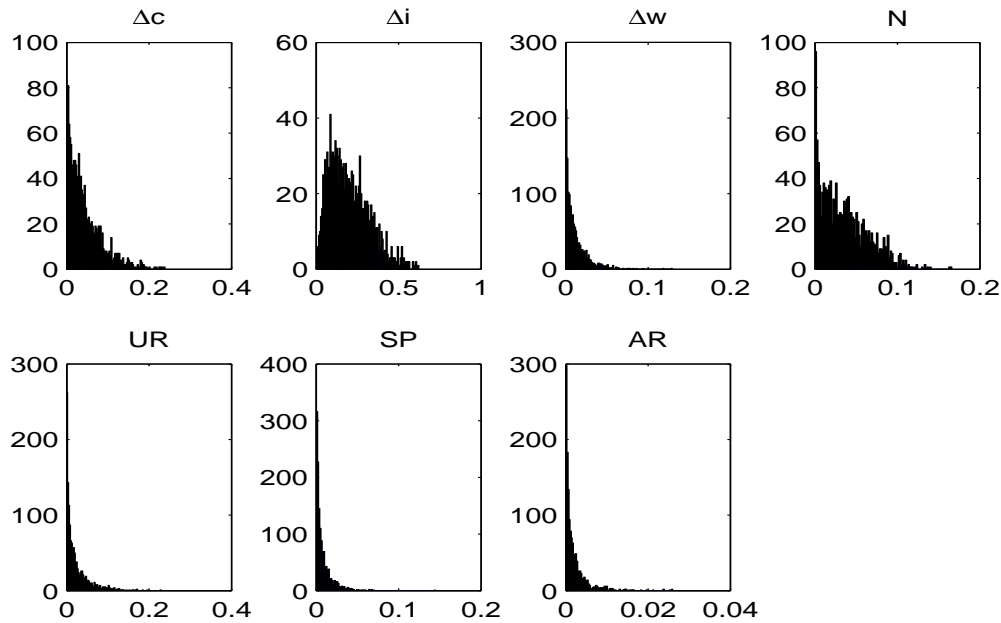


Figure 2: Distribution of FEVD contributions of the auxiliary variables to the government spending shock  $g_t$  for the 2000 Monte Carlo replications

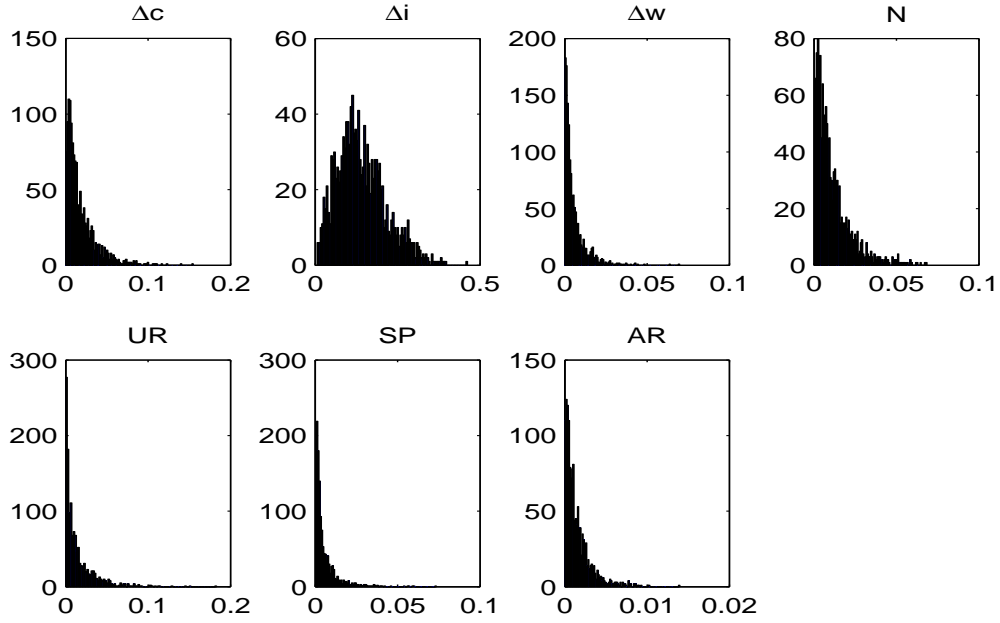


Figure 3: Distribution of FEVD contributions of the auxiliary variables to the monetary policy shock  $mp_t$  for the 2000 Monte Carlo replications

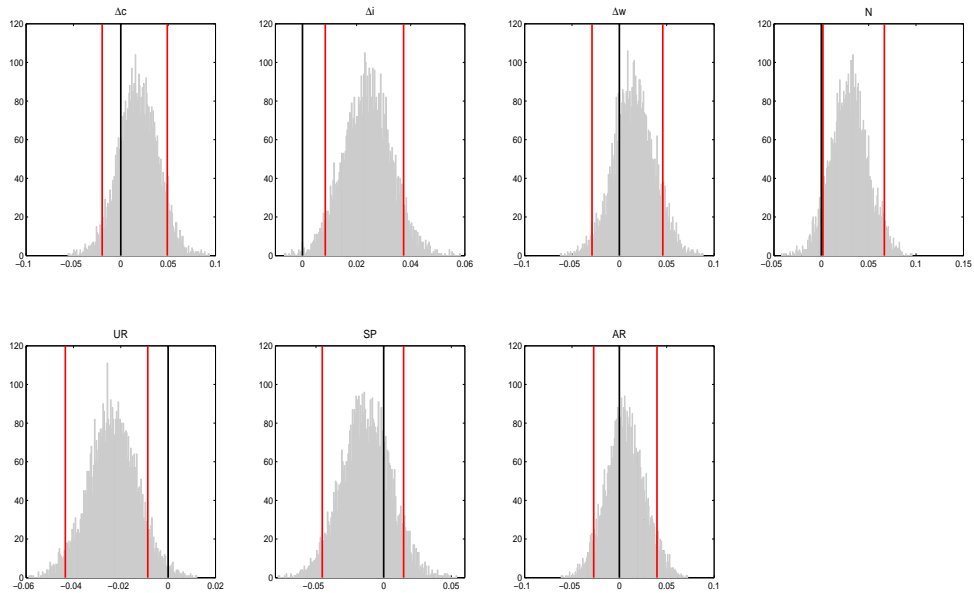


Figure 4: Distribution of the coefficients of the row in the  $A_{12}$  block of matrix M that load the auxiliary variables in the title of each subplot onto the government spending shock  $g_t$  and their 90% credible set. Under the DSGE prior these coefficients are zero.

- neutral technology shock:  $a_t = \Delta \ln(A_t) = (1 - \rho_a)\gamma + \rho_a a_{t-1} + \varepsilon_t^a$
- price mark-up shock:  $\lambda_t^p = \rho_p \lambda_{t-1}^p + \varepsilon_t^p - \theta_p \varepsilon_{t-1}^p$
- wage mark-up shock:  $\lambda_t^w = \rho_w \lambda_{t-1}^w + \varepsilon_t^w - \theta_w \varepsilon_{t-1}^w$
- risk premium shock:  $b_t = \rho_b b_{t-1} + \varepsilon_t^b$
- investment shock:  $q_t = \rho_q q_{t-1} + \varepsilon_t^q$
- government spending shock:  $g_t = (1 - \rho_g) \ln g + \rho_g g_{t-1} + \varepsilon_t^g$
- monetary policy shock:  $r_t = \rho_r r_{t-1} + \varepsilon_t^r$
- labor supply shock:  $\chi_t = \rho_\chi \chi_{t-1} + \varepsilon_t^\chi$

I then link the model to the variables in Table 2, *i.e.* some of the main macro variables followed by the markets and economic research institutions like the NBER.

The first step of the exercise requires estimating the parameters  $\theta$  of the DSGE model: I use the same data set (on the sample 1983-2011) and priors as Galí, Smets and Wouters (2012) and obtain a posterior distribution for  $\theta$ . Given these parameters the estimates of the underlying states of the DSGE model,  $\hat{s}_t$ , can be extracted. As already mentioned above, I now treat the estimated states as observables and use them to estimate BVARX on the auxiliary variables collected in table 2. Let  $X_t$  denote the  $20 \times 1$  vector of the auxiliary variables. I estimate the following VAR(1) for  $X_t$ :

$$X_t = \Pi_1 X_{t-1} + \Gamma \hat{s}_t + \xi_t,$$

where  $\hat{s}_t$  are the estimated states of the DSGE model. The estimation of the BVAR is performed as described above. In this way we obtain  $\Pi_1^*$  and  $\Gamma^*$ , which I will use as a prior for the joint model.

	Variable	Mnemonic
1	Industrial Production	IPTOT
2	Capacity Utilization	CUTOT
3	Purchasing Managers Index	PMI
4	Employment on nonaggregate payrolls: Total	PYRLTOT
5	Average hourly earnings: Total nonagricultural	HRLYEGSTOT
6	Sales: Mfg. & Trade: Total (mil of chained 96\$)	SALESTOT
7	Personal Consumption Expenditures: Total (bil of chained 96\$)	PCETOT
8	Privately-owned housing, started: Total	HSTARTS
9	Inventories: Mfg. & Trade, Total (mil of chained 96\$)	INVTOT
10	S & P composite	SPCOMP
11	Interest rate: 1-year Treasury (constant maturity)1	YTBILL
12	Primary market yield on 30-year fixed mortgage	MORTG
13	M2	M''
14	Loans and Securities @ all comm. banks: Total (mil of \$)	LSTOT
15	Loans and Securities @ all comm. banks: Securities, total (mil of \$)	LSSEC
16	Loans and Securities @ all comm. banks: Consumer loans (mil of \$)	LSCONS
17	PPI: crude materials	PPICRUDEMAT
18	Philadelphia Business Outlook Survey: General activity	PHBOSGA
19	Spread: Bba - Aaa rates	SPREAD

Table 2: The auxiliary variables

Let us now stack the DSGE model and the model for the auxiliary variables.

$$\begin{bmatrix} s_t \\ X_t \end{bmatrix} = M \begin{bmatrix} s_{t-1} \\ X_{t-1} \end{bmatrix} + N \begin{bmatrix} \varepsilon_t \\ \xi_t \end{bmatrix}$$

We postulate that  $(M, N)$  have a distribution centered around  $(M^*(\theta), N^*(\theta))$ ,

where

$$M^* = \begin{bmatrix} A^*(\theta^*) & 0 \\ \Gamma^* A^*(\theta^*) & \Pi_1^* \end{bmatrix}, \quad N^* = \begin{bmatrix} H^*(\theta^*) & 0 \\ \Gamma^* B^*(\theta^*) & V^* \end{bmatrix} \quad (12)$$

As mentioned in the previous section, we implement the prior using the method of generalized dummy observations developed by Sims (2008) and, in particular, we use Type 1 and 2 priors only. The tightness of the prior is governed with a scaling



Table 3: Forecast error variance decomposition at a 30-quarter horizon, in percentage points

	b	q	g	a	r	$\chi$	$\eta_w$	$\eta_p$	Aux
b	81.46	1.42	3.63	2.09	0.02	3.03	0.58	0.00	7.76
q	9.41	49.16	8.51	13.35	0.00	6.39	1.51	0.00	11.67
g	2.08	0.53	62.18	25.32	0.01	5.19	0.42	0.00	4.26
a	11.24	3.54	3.34	72.95	0.00	3.45	0.62	0.00	4.26
r	6.36	1.08	0.45	4.38	0.00	0.15	83.90	0.00	3.67
$\chi$	21.71	4.36	0.87	34.69	0.00	27.80	1.75	0.00	8.81
$\eta_w$	7.23	0.70	1.18	6.68	0.25	1.38	1.97	74.04	6.56
$\eta_p$	3.31	0.24	5.96	32.02	34.91	13.10	0.14	0.00	10.32

parameter. In particular, we will allow the parameter governing the tightness of the prior on the first  $n_w$  columns of  $M$ , denoted by  $\lambda_1$  to differ from the tightness parameter for the other columns of  $M$ , denoted by  $\lambda_2$ . In practice, with respect to  $w_t$ , the parameter  $\lambda_1$  governs the tightness of the priors coming from the DSGE model, while  $\lambda_2$  is the tightness parameters for the zero-restrictions on the coefficients that determine the impact of the auxiliary variables on the states of the model.

I determine the values of  $\lambda_1$  and  $\lambda_2$  that maximize the marginal data density and find that these values push the model away from the prior that the GSW shocks are exogenous to the auxiliary variables. Therefore I measure how much of the variance of the model's shocks' is explained by other shocks and the auxiliary variables, that is a forecast error variance decomposition (FEVD), for example evaluated at the posterior mode. Table 3 reports the FEVD at a 30 quarters horizon. This table highlights that despite the richness of the model's dynamics, the auxiliary variables contribute to the total variability of the model's driving process. In the case of the investment specific shock  $q_t$  the auxiliary variables determine up to 12% of the shock's variance.

Focussing on the investment specific shock, I can pitch into the lively debate regarding the importance of investment shocks in driving the business cycle and

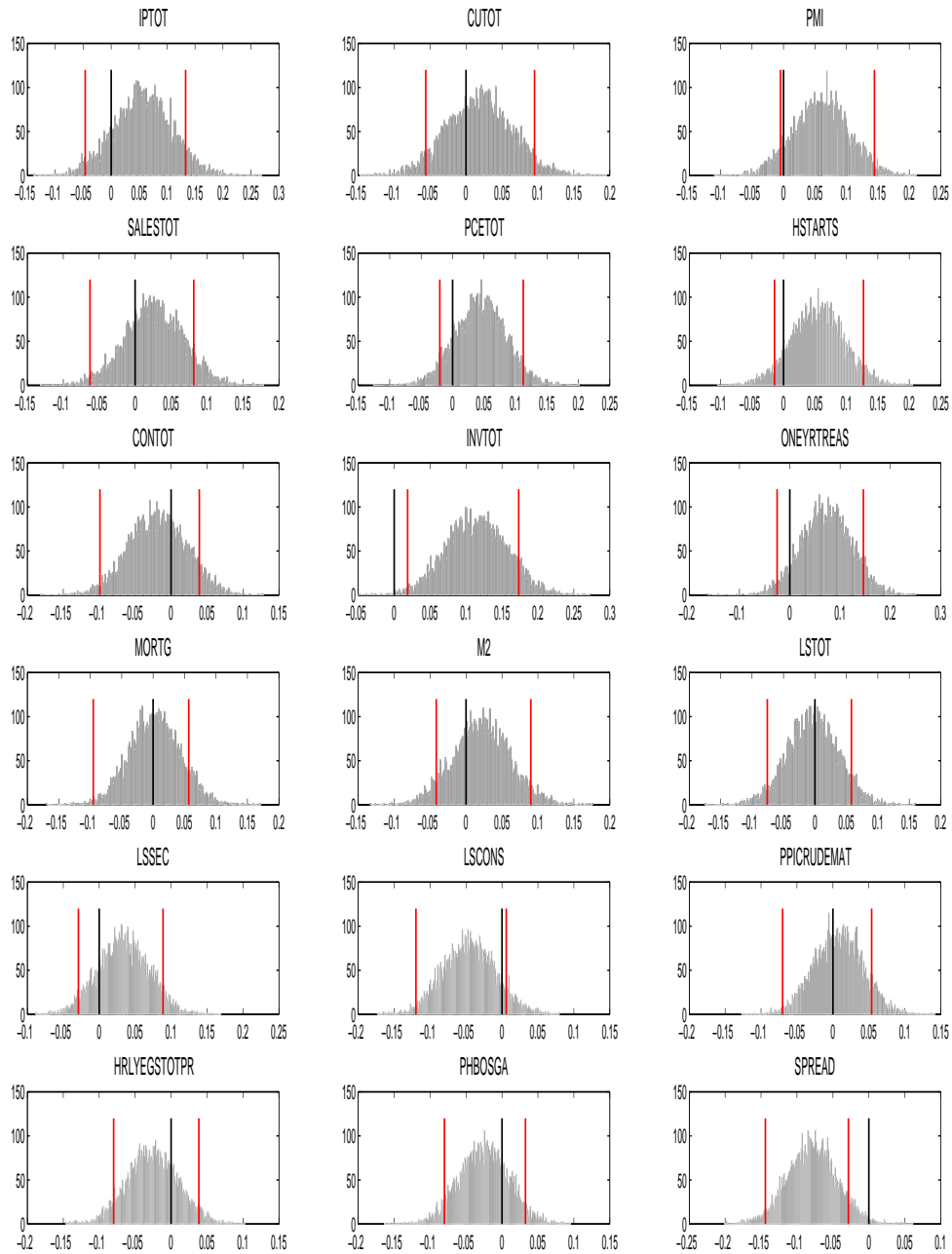


Figure 5: Distribution of the coefficients of the row in the  $A_{12}$  block of matrix  $M$  that load the auxiliary variables in the title of each subplot onto the investment specific shock  $q_t$  and their 90% credible set. Under the DSGE prior these coefficients are zero.

Table 4: Forecast error variance decomposition at a 30-quarter horizon, in percentage points

	b	q	g	a	r	$\chi$	$\eta_w$	$\eta_p$	nw	Aux
b	76.39	9.99	3.25	5.20	0.47	0.05	0.00	0.36	0.65	3.64
q	3.79	73.47	4.96	11.12	0.41	0.05	0.01	1.12	0.17	4.88
g	6.17	12.07	49.89	26.67	0.17	1.2	0.00	0.60	0.42	2.79
a	2.85	1.09	0.15	86.83	0.84	1.75	0.04	0.59	0.15	5.71
r	0.92	5.01	1.37	13.67	0.17	0.12	75.59	0.28	0.04	2.83
$\chi$	6.37	25.95	17.26	25.25	0.65	20.26	0.03	0.07	0.34	3.82
$\eta_w$	0.31	5.93	2.97	5.82	0.36	0.17	0.02	79.14	0.09	5.20
$\eta_p$	2.75	2.36	2.95	24.98	61.92	0.94	0.00	0.87	0.15	3.07
nw	14.82	14.32	12.01	17.87	0.50	1.50	0.01	0.17	36.50	2.28

their key features (e.g. Fisher 2006 and Schmitt-Grohe and Uribe 2008). Figure (??) reports the distribution of the coefficients of the row in the  $A_{12}$  block of matrix  $M$  that load onto the investment specific shock  $q_t$ . The investment shock affects the way in which investment goods are transformed into productive capital, so it is not surprising to find that it is Granger-caused by variables such as the spread between Aaa and Baa graded corporate bonds and by variables that are informative about the state of the economy, such inventories. The investment shock does seem to pick up unmodeled aspects of the financial markets and can be seen as a proxy for overall health of the financial system, as stated in Justiniano, Primiceri and Tambalotti, 2010.

I verify whether adding a financial accelerator in a model of this type can help address some of the misspecification, using the variant of the GSW with the financial frictions that I used as DGP in the Monte Carlo exercise to filter the data. The misspecification in the driving processes of this model is substantially lower, as shown in table 4

## 6 Conclusions

In this paper I propose a method for analyzing the driving processes of a DSGE model, in order to understand what these proxy and to gauge what sort of missing mechanism is determining the misspecification, if any. The idea is to model the states of the DSGE and auxiliary variables jointly, imposing the restrictions implied by the DSGE as priors, and then verify how much weight is given to the priors in estimation. In particular, I identify the tightness of the prior that maximises the marginal data density of the joint model and I verify what this implies for the parameters of the joint model. If the driving processes of the model, which are assumed to be exogenous in the DSGE model, are found to be Granger-caused by some auxiliary variables, then we are in the presence of some form of misspecification. Forecast error variance decompositions (FEVDs) can help us assess how important this misspecification is, by indicating how much of the variance of the driving processes of the model is determined by off-model variables. The methodology allows to also relax restrictions governing the dynamic relations between the disturbances driving the model in order to check whether these, rather than additional variables, would be the best way to capture the misspecification.

I assess the methodology's efficacy in a Monte Carlo exercise, where the true data generating process is known. I also apply the proposed methodology to a state-of-the-art DSGE model, the Galí, Smets and Wouters (2012). The joint analysis of the auxiliary information and the model is very informative regarding the properties of the underlying structural shocks of the model. Not only can it identify if some shocks are misspecified, but it can also shed light on what these shocks proxy or capture. For example, I find support for Justiano, Primiceri and Tambalotti's (2010) claim that the investment shock is a proxy for the overall health of the financial system.

## References

- [1] An, S. and F. Schorfheide, 2007. “Bayesian Analysis of DSGE models,” *Econometric Reviews*, 26(2-4), 113-172.
- [2] Banbura, M., D. Giannone and L.Reichlin 2010. “Large Bayesian VARs,” *Journal of Applied Econometrics*,25(1), 71-92.
- [3] Bernanke, B., M. Gertler and S. Gilchrist (1999). “The Financial Accelerator in a Quantitative Business Cycle Framework,” in *Handbook of Macroeconomics*, ed. by J.B. Taylor and M. Woodford, vol. 1, North Holland, Amsterdam.
- [4] Blanchard, O.J. and C.M. Kahn 1980. “The Solution of Linear Difference Models under Rational Expectations,” *Econometrica* 48(5), 1305-1311
- [5] Boivin, J. and M. Giannoni 2006. “DSGE Models in a Data-Rich Environment,” NBER Working Papers 12772, National Bureau of Economic Research, Inc
- [6] Canova, F., F. Ferroni and C. Matthes, 2014. “Choosing the variables to estimate singular DSGE models,” *Journal of Applied Econometrics*, 29, 1099–1117.
- [7] Carriero, A. T. Clark and M. Marcellino, 2013. “Bayesian VARs: Specification Choices and Forecast Accuracy,” *Journal of Applied Econometrics*, doi: 10.1002/jae.2315
- [8] Carter, C.K. and R. Kohn, 1994. “On Gibbs Sampling for State Space Models,” *Biometrika*, 81, 541-553
- [9] Christiano, L.J., M. Eichenbaum, and C.L. Evans. 2005. “Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy.” *Journal of Political Economy* 113 (1): 1–45.
- [10] Curdia V. and R. Reis, 2010. “Correlated Disturbances and U.S. Business Cycle.” NBER Working paper #15774
- [11] Del Negro, M., R. Hasegawa and F. Schorfheide (2014). “Dynamic Prediction Pools: An Investigation of Financial Frictions and Forecasting Performance,” NBER Working Paper No. 20575.
- [12] Del Negro, M. and F. Schorfheide, 2009. “Monetary Policy Analysis with Potentially Misspecified Models.” *American Economic Review*, 99:4, 1415-1450.
- [13] Del Negro M. and F. Schorfheide 2004. “Priors from General Equilibrium Models for VARs,” *International Economic Review*, 45 (2004): 643–673.
- [14] Del Negro, M., F. Schorfheide, F. Smets and R. Wouters (2007), “On the Fit of New Keynesian Models,” *Journal of Business and Economic Statistics* 25:123-43.

- [15] Evans, C., 1992. "Productivity shocks and business cycles," *Journal of Monetary Economics*, 29(2), 191-208.
- [16] Fernandez-Villaverde, J., J. Rubio-Ramirez, T.J. Sargent and M.W. Watson, 2007. ABCs (and Ds) of Understanding VARs. *American Economic Review*, 97(3), 1021-1026.
- [17] Galí, J., F. Smets and R. Wouters, 2012. "Unemployment in an Estimated New Keynesian Model," *NBER Macroeconomics Annual*, University of Chicago Press, vol. 26(1), pages 329 - 360.
- [18] Galí, J. 2011a. "The Return of the Wage Phillips Curve," *Journal of the European Economic Association*, 9 (3), 2011, 436-461
- [19] Galí, J. 2011b. "Monetary Policy and Unemployment," in B. Friedman and M. Woodford (eds.) *Handbook of Monetary Economics*, vol. 3A, Elsevier B.V., 2011, 487-546
- [20] Giannone, D., M. Lenza and G. Primiceri, 2015. "Prior Selection for Vector Autoregressions," *Review of Economics and Statistics*, forthcoming.
- [21] Giannone, D. and L. Reichlin, 2006. "Does information help recovering structural shocks from past observations?," *Journal of the European Economic Association*, 4(2-3), 455-465
- [22] Holland, A. and A. Scott, 1998. "The Determinants of UK Business Cycles," *Economic Journal*, Royal Economic Society, vol. 108(449), pages 1067-92.
- [23] Inoue, A., C. Kuo and B. Rossi, 2014. "Identifying the Sources of Model Misspecification," CEPR Discussion Paper No. 10140
- [24] Justiniano, A., G. Primiceri and A. Tambalotti, 2010. "Investment Shocks and Business Cycles," *Journal of Monetary Economics*, 57(2),132-145.
- [25] Justiniano, A., G. Primiceri and A. Tambalotti, 2009. "Investment Shocks and the Relative Price of Investment" CEPR Discussion Paper 7598.
- [26] Klein, P. 2000. "Using the Generalized Schur Form to Solve a System of Linear Expectational Difference Equations," *Journal of Economic Dynamics and Control*, vol. 24(10), 1405-1423
- [27] Mortensen, D. and C. Pissarides, 1994. "Job Creation and Job Destruction in the Theory of Unemployment", *Review of Economic Studies* 61(3), pp. 397-415.
- [28] Smets, F. and R. Wouters, 2007. Shocks and Frictions in US Business Cycles: a Bayesian DSGE Approach. *American Economic Review*, 97(3),586-606.
- [29] Smets, F. and R. Wouters, 2004. Forecasting with a Bayesian DSGE Model: an Application to the Euro Area. *Journal of Common Market Studies*, 42(4), 841-867.

- [30] Smets, F. and R.Wouters. 2003. “An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area.” *Journal of the European Economic Association* 1 (5): 1123–175.
- [31] Sims, C.A., 2008. “Making Macro-Models Behave Reasonably,” mimeo.
- [32] Sims, C.A., 2002a. Solving Linear Rational Expectations Models. *Computational Economics*, Springer, 20(1-2), 1-20.

# Appendix A

Here we summarize the key log-linear equations of the GSW model. We refer to Galí, Smets and Wouters (2012) for a more detailed description of the model.

- Consumption Euler equation:

$$\hat{c}_t = c_1 E_t [\hat{c}_{t+1}] + (1 - c_1) \hat{c}_{t-1} - c_2 \left( \hat{R}_t - E_t [\hat{\pi}_{t+1}] - \hat{\varepsilon}_t^b \right)$$

with  $c_1 = (h/\tau)/(1 + (h/\tau))$ ,  $c_2 = (1 - h/\tau)/(1 + (h/\tau))$  where  $h$  is the external habit parameter.  $\hat{\varepsilon}_t^b$  is the exogenous AR(1) risk premium process.

- Investment Euler equation:

$$\hat{i}_t = i_1 \hat{i}_{t-1} + (1 - i_1) \hat{i}_{t+1} + i_2 \hat{Q}_t^k + \hat{\varepsilon}_t^q$$

with  $i_1 = 1/(1 + \beta)$ ,  $i_2 = i_1/(\tau^2 \Psi)$  where  $\beta$  is the discount factor and  $\Psi$  is the elasticity of the capital adjustment cost function.  $\hat{\varepsilon}_t^q$  is the exogenous AR(1) process for the investment specific technology.

- Aggregate demand equals aggregate supply:

$$\hat{y}_t = \frac{c_*}{y_*} \hat{c}_t + \frac{i_*}{y_*} \hat{i}_t + \hat{\varepsilon}_t^g + \frac{r_*^k k_*}{y_*} \hat{u}_t \quad (13)$$

$$= \mathcal{M}_p \left( \alpha \hat{k}_t + (1 - \alpha) \hat{L}_t + \hat{\varepsilon}_t^a \right) \quad (14)$$

with  $\mathcal{M}_p$  reflecting the fixed costs in production which corresponds to the price markup in steady state.  $\hat{\varepsilon}_t^g, \hat{\varepsilon}_t^a$  are the AR(1) processes representing exogenous demand components and the TFP process.

- Price-setting under the Calvo model with indexation:

$$\hat{\pi}_t - \gamma_p \hat{\pi}_{t-1} = \pi_1 (E_t [\hat{\pi}_{t+1}] - \gamma_p \hat{\pi}_t) - \pi_2 \mu_t^p + \hat{\varepsilon}_t^p$$

with  $\pi_1 = \beta$ ,  $\pi_2 = (1 - \theta_p \beta)(1 - \theta_p)/[\theta_p(1 + (\mathcal{M}_p - 1)\varepsilon_p)]$  and  $\theta_p$  and  $\gamma_p$  are, respectively, the probability and indexation of the Calvo model, and  $\varepsilon_p$  the curvature of the aggregator function. The price markup  $\mu_t^p$  is equal to the inverse of the real marginal  $\hat{m}c_t = (1 - \alpha)\hat{w}_t + \alpha r \hat{k}_t - \hat{A}_t$ .

- Wage-setting under the Calvo model with indexation:

$$\pi_t^w = \gamma_w \pi_{t-1}^w + \beta E_t [\pi_{t+1}^w - \gamma_w \pi_t^w] - \lambda_w \phi u_t + \lambda_w \mu_t^w$$

where the unemployment rate  $u_t = l_t - n_t$  is defined so as to include all the individuals who would like to be working (given current labour market conditions, and while internalizing the benefits that this will bring to their households) but are not currently employed.



- Capital accumulation equation:

$$\hat{k}_t = \kappa_1 \hat{k}_{t-1} + (1 - \kappa_1) \hat{i}_t + \kappa_2 \hat{\varepsilon}_t^q$$

with  $\kappa_1 = 1 - (i_*/\bar{k}_*)$ ,  $\kappa_2 = (i_*/\bar{k}_*)(1 + \beta)\Psi$ . Capital services used in production is defined as:  $\hat{k}_t = \hat{u}_t + \hat{k}_{t-1}$

- Optimal capital utilisation condition:

$$\hat{u}_t = \frac{1 - \phi}{\phi} \hat{r}_t^k$$

with  $\phi$  being the elasticity of the capital utilisation cost function.

- Optimal capital/labour input condition:

$$\hat{k}_t = \hat{w}_t - \hat{r}_t^k + \hat{L}_t$$

- Monetary policy rule:

$$\hat{R}_t = \rho_r \hat{R}_{t-1} + (1 - \rho_r)(r_\pi \hat{\pi}_t + r_y ygap_t) + r_{\Delta y} \Delta y_t + \varepsilon_t^r$$

where  $ygap_t = y_t - yflex_t$  is the difference between actual output and the output in the flexible price and wage economy in absence of distorting price and wage markup shocks.

- In practice, as Del Negro, Hasegawa and Schorfheide (2014) show for the SW, adding the financial frictions to this model simply amount to replacing the equation for the value of the capital stock with the following conditions:

$$E_t \left[ \hat{R}_t^k - \hat{R}_t \right] = b_t + \zeta_{sp,b} (\hat{Q}_t^k + \bar{k}_t - n_t) + \sigma_{\omega,t}$$

$$\hat{R}_t^k - \pi_t = \frac{r_*^k}{r_*^k + 1 - \delta} r_t^k + \frac{1 - \delta}{r_*^k + 1 - \delta} \hat{Q}_t^k - \hat{Q}_{t-1}^k$$

$$n_t = \zeta_{nrk} (\hat{R}_t^k - \pi_t) - \zeta_{nr} (\hat{R}_t - \pi_t) + \zeta_{nqk} (\hat{Q}_{t-1}^k + \bar{k}_{t-1}) + \zeta_{nn} n_{t-1} - \frac{zeta_{\sigma} n_{\sigma}}{sp\sigma} \sigma_{\omega,t-1},$$

which define respectively the spread, the return on capital and the evolution of the entrepreneurial net worth. Unlike Del Negro, Hasegawa and Schorfheide (2014) we estimate the parameters in this last equation directly. The measure of spreads in the observables is related to the model variables  $E_t \left[ \hat{R}_t^k - \hat{R}_t \right]$  as follows:

$$Spread = SP^* + 100 + E_t \left[ \hat{R}_t^k - \hat{R}_t \right]$$

We calibrate the  $\delta$ ,  $\frac{c}{g}$  and  $h$  to standard values of 0.025, 0.18 and 0.7 respectively, while we calibrate the following parameters to their mean posterior

values in GSW (2012):  $\beta = (0.31/100 + 1)^{-1}$ ,  $\Psi = 3.96$ ,  $\alpha = 0.17$ ,  $M_p = 1.74$ ,  $M_w = 1.22$ ,  $\zeta_p = 10$ ,  $\rho_{chi} = 0.99$ , and  $cgy = 0.69$ .

The priors and posterior modes and means of the estimated parameters are reported in the following table.

	Prior Distribution			Posterior Distribution	
	Distr.	mean	st.dev	mode	mean
$\nu$	B	0.5	0.2	0.05	0.005
$\rho_\pi$	N	1.5	0.125	1.70	1.69
$\rho_{ygap}$	N	0.12	0.01	0.09	0.10
$\rho_{\Delta ygap}$	N	0.12	0.01	0.11	0.13
$\theta_w$	B	0.5	0.1	0.81	0.79
$\theta_p$	B	0.5	0.1	0.95	0.95
$\gamma_p$	B	0.5	0.1	0.24	0.26
$\gamma_w$	B	0.5	0.1	0.67	0.61
$\psi$	B	0.5	0.15	0.25	0.30
$\rho_r$	B	0.75	0.10	0.79	0.83
$\phi$	N	2	0.5	4.31	4.26
$SP^*$	N	2	0.5	1.49	1.45
$\Pi^*$	G	0.62	0.1	0.66	0.66
$l^*$	N	0	0.1	0.02	0.04
$\zeta_{spb}$	B	0.2	0.1	0.15	0.12
$\zeta_{rk}$	B	0.2	0.1	0.22	0.45
$\zeta_{nr}$	B	0.2	0.1	0.19	0.11
$\zeta_{nq}$	B	0.2	0.1	0.05	0.02
$\zeta_{nn}$	B	0.8	0.1	0.88	0.9
$\zeta_{n\sigma}$	N	2	0.5	5.12	4.54
$\tau$	N	0.40	0.1	0.39	0.35
$\rho_b$	B	0.5	0.2	0.98	0.97
$\rho_q$	B	0.5	0.2	0.82	0.71
$\rho_g$	B	0.5	0.2	0.99	0.99
$\rho_a$	B	0.5	0.2	0.99	0.99
$\rho_{ms}$	B	0.5	0.2	0.02	0.02
$\rho_p$	B	0.5	0.2	0.65	0.65
$\rho_w$	B	0.5	0.2	0.19	0.20
$\rho_{nw}$	B	0.5	0.2	0.98	0.99
$\sigma_b$	U	2.5	1.44	0.21	0.12
$\sigma_q$	U	2.5	1.44	0.71	0.92
$\sigma_g$	U	2.5	1.44	0.47	0.49
$\sigma_a$	U	2.5	1.44	0.43	0.42
$\sigma_p$	U	2.5	1.44	3.18	3.61
$\sigma_r$	U	2.5	1.44	0.24	0.23
$\sigma_w$	U	2.5	1.44	2.65	2.56
$\sigma_\chi$	U	2.5	1.44	1.52	1.30
$\sigma_{nw}$	U	2.5	1.44	0.08	0.10

Table 5: Prior and posterior distribution of the parameters of the model estimated over the period 1959Q2 to 2006Q4.