EUROPEAN CENTRAL BANK
EUROSYSTEM

Working Paper Series

Jaime Martínez-Martín, Richard Morris,
Luca Onorante, Fabio M. Piersanti

Merging structural and reduced-form models for forecasting: opening the DSGE-VAR box

No 2335 / December 2019

Abstract

The post-crisis environment has posed important challenges to standard forecasting models. In this paper, we exploit several combinations of a large-scale DSGE structural model with standard reduced-form methods such as (B)VAR (i.e. DSGE-VAR and Augmented-(B)VAR-DSGE methods) and assess their use for forecasting the Spanish economy. Our empirical findings suggest that: (i) the DSGE model underestimates growth of real variables due to its mean reverting properties in the context of a sample that is difficult to deal with; (ii) in spite of this, reduced-form VARs benefit from the imposition of an economic prior from the structural model; and (iii) pooling information in the form of variables extracted from the structural model with (B)VAR methods does not give rise to any relevant gain in terms of forecasting accuracy.

Keywords: Bayesian VAR, DSGE models, real time data, forecast comparison.

JEL Classification: C54, E37, F3, F41.

Non-technical summary

The prediction of key macroeconomic time series is an essential input for economic policy decisions in governments and central banks. At the same time the post-crisis environment has posed important challenges for standard forecasting models. Accordingly, the need for new and more flexible tools to monitor and forecast economic developments in real time is increasingly acknowledged.

Macroeconomic forecasting typically follows two different approaches: structural and non-structural. Non-structural methods attempt to exploit the reduced-form correlations between macroeconomic variables, while structural macroeconomic forecasting is grounded on economic theory. This paper aims to evaluate the real-time predictive accuracy of hybrid models at short and long horizons, combining structural and non-structural forecasting methods, to better fit the Spanish economy.

The Spanish economy is of particular interest because the financial crisis of 2008-09, and subsequent sovereign debt crisis pose important challenges for standard forecasting models for at least two reasons: (i) the persistence of the recession and the very gradual recovery could square either with a medium-term phenomenon or with standard business cycle theory; (ii) policy-driven structural changes (i.e., financial regulation, structural reforms, etc.) along with ongoing processes such as the zero lower bound (ZLB), deleveraging, or fiscal consolidation can be naturally understood as different regimes.

Our empirical strategy is twofold. Firstly, we make use of the DSGE modelling approach of Smets and Wouters (2003) to inform the estimation of non-structural models along the lines of the DSGE-VAR method of Del Negro and Schorfheide (2004) and Del Negro *et al.* (2007). Secondly, we use artificial series from the same structural model to expand the variable space where the reduced-form models operate (Augmented-(B)VAR-DSGE) following Fernández- de-Córdoba and Torres (2011).

This New-Keynesian DSGE model with sticky prices and wages was chosen for several reasons. First, the scale of DSGE models has consistently grown over time and those New-Keynesian features that have been incorporated have improved our understanding of key macroeconomic relationships. Second, simpler real business cycle canonical models have proven to track relatively well macroeconomic variables for the Spanish economy at long horizons, but we aim to provide a forecasting assessment within a more complex DSGE model using alternative techniques.

Our empirical findings suggest that the pseudo real-time out-of-sample forecasting performance of the DSGE–VAR model tends to be better than most alternatives under consideration. We further document that: (i) the DSGE model underestimates the growth rate of real variables due to its mean reverting properties in the context of a sample that is difficult to deal with; (ii) in spite of this, reduced-form VARs largely benefit from the imposition of an economic prior. In fact, the optimal prior tightness is surprisingly unchanged in spite of several financial crises unevenly spreading through the several estimation samples; and (iii) on the other hand, adding information extracted from the structural model to (B)VAR methods does not lead to relevant gains in terms of forecasting accuracy. Indeed, the benefit from incorporating the main business cycle drivers (also conditional on the imposition of a non-economic prior) as additional observables in reduced–form models is quite limited. In addition, we show that (iv) forecasts of real variables produced by combinations of structural and non–structural models, especially the DSGE-VAR and the BVAR models, are generally optimal and rational (i.e. unbiased and efficient) in absolute terms. Looking across the sample, we also find that (v) the relaxation of the DSGE prior is particularly effective in some specific periods, coherent with the dynamics of the observables first moments.

1 Introduction

The prediction of key macroeconomic time series is an essential input for economic policy decisions in governments and central banks. As pointed out by Diebold (1998), macroeconomic forecasting follows two different approaches: structural and non-structural. Non-structural methods attempt to exploit the reduced-form correlations between macroeconomic variables, while structural macroeconomic forecasting is grounded on economic theory. The related literature acknowledges that the advantages of VARs and Bayesian VARs make them extremely appealing to macroeconomists: they are easy to estimate, generate out-of-sample forecasts, and are very flexible.[1] However, they embed little (Structural VARs) or no (unrestricted VARs) economic theory.

The alternative to using purely statistical methods is to use a theory-based approach. Structural macroeconomic forecasting is generally based on Dynamic Stochastic General Equilibrium (DSGE) models.[2] However, such models were not considered useful tools for forecasting until very recently. In fact, their forecasting performance was typically underscored by their inability to track and predict co-movements of aggregate time series over the business cycle.

When comparing the real-time forecasting accuracy of structural and reduced-form time series models, no single method can be considered the best at all horizons (Gürkaynak et al., 2013). Simple autoregressive (AR) models tend to be more accurate at short horizons and DSGE models are preferable at long horizons when forecasting output growth, while the opposite is generally true for inflation.

In this paper, we evaluate the real-time predictive accuracy of hybrid models, combining structural and non-structural forecasting methods for the Spanish economy. This idea is not new. Looking at the U.S. economy, Del Negro and Schorfheide (2004) showed how theoretical DSGE models which incorporate rational, forward-looking agents can inform (through priors) reduced-form time series models. In terms of forecasting, Lees et al. (2011) tested the predictive ability of the combination of a small-scale DSGE model and a statistical VAR model which outperformed the Reserve Bank of New Zealand forecasts. Cai et al. (2018), using the New York Fed DSGE model, also showed that "empirical" variants of DSGE models, expanded by including financial variables as observables, perform relatively well compared to both the Blue Chip Survey and the Survey of Professional Forecasters (SPF) in terms of output growth forecasting accuracy.

Against this background, our empirical strategy is twofold. Firstly, we make use of the DSGE

[1] See for instance Sims (1980), Litterman (1986a) and Litterman (1986b).

[2] DSGE models have a strong theoretical background as they are firmly grounded on modern micro-foundations. A broad class of macroeconomic models that spans the standard neoclassical growth model (see King and Rebelo, 1999) as well as New Keynesian monetary models with real and nominal frictions that are based on the work of Christiano et al. (2005) and Smets and Wouters (2003) is encompassed under the term *DSGE model*.

modelling approach of Smets and Wouters (2003) (SW henceforth) to inform the estimation of non-structural models along the lines of the DSGE-VAR method of Del Negro and Schorfheide (2004) and Del Negro et al. (2007). Secondly, we use artificial series from the same structural model to expand the variable space where the reduced-form models operate (Augmented-(B)VAR-DSGE) following Fernández-de-Córdoba and Torres (2011).

The New-Keynesian, SW model with sticky prices and wages was chosen for several reasons. First, the scale of DSGE models has consistently grown over time and those New Keynesian features that have been incorporated have improved our understanding of key macroeconomic relationships. Second, simpler real business cycles (RBC) canonical models have proven to track relatively well macroeconomic variables for the Spanish economy at long horizons, but we aim to provide a forecasting assessment within a more complex DSGE model using alternative techniques.[3]

We test the resulting models in forecasting the Spanish economy. Given that our sample includes the 2008-09 financial crisis and subsequent sovereign debt crisis, this is particularly challenging for at least two reasons: (i) the persistence of the recession and the very gradual recovery could square either with a medium-term phenomenon or with standard business cycle theory; and (ii) policy-driven structural changes (i.e., financial regulation, structural reforms, etc.) along with ongoing processes such as the zero lower bound (ZLB), deleveraging, or fiscal consolidation can be naturally understood as different regimes.

Some DSGE models developed recently for the Spanish economy are those used at the Bank of Spain: BEMOD (Andrés et al., 2006; Andrés et al., 2010; Andrés et al., 2017), and at the Spanish Ministry of Economy and Finance: MEDEA (Burriel et al., 2010) and the REMS model (Boscá et al., 2007; Boscá et al., 2010; Boscá et al., 2018; or Goméz-González and Rees, 2018). As such, our modelling approach can borrow from many years of experience of DSGE modelling of the Spanish economy.

This paper provides several contributions. First, we exploit different combinations of a large-scale DSGE model in the New Keynesian tradition with standard non-structural methods such as (B)VARs (i.e. DSGE–VAR) and test them in a forecasting context. Second, we expand the variable space of the reduced-form models with artificial series drawn from the structural model, testing the usefulness of Augmented–(B)VAR–DSGE methods. This is the first attempt - to the best of our knowledge - to use such a class of models for the Spanish economy. Finally, we place our models and their combinations in a forecasting competition, to identify the most appropriate

[3]Our model shares many features with other DSGE models developed at policy-making institutions around the world. Some examples are the Federal Reserve Board (Erceg et al., 2006), the European Central Bank (Warne et al., 2008), the Bank of Canada (Murchison and Rennison, 2006), the Bank of England (Harrison et al., 2005), the Bank of Finland (Kilponen and Ripatti, 2006; Kortelainen, 2002) and the Bank of Sweden (Adolfson et al., 2007).

combinations; in doing so we "open the box" of the models and explain where the advantages of different techniques arise.

Our empirical findings suggest that the pseudo real-time out-of-sample forecasting performance of the DSGE–VAR model tends to be better than the alternatives under consideration. We further document that: (i) the DSGE model based on SW underestimates the growth rate of real variables due to its mean reverting properties in the context of a sample that is difficult to deal with; (ii) in spite of this, reduced-form VARs generally benefit from the imposition of an economic prior. In fact, the optimal prior tightness is surprisingly unchanged in spite of several financial crises unevenly spreading through the several estimation samples; (iii) on the other hand, adding information extracted from the structural model to (B)VAR methods does not lead to relevant gains in terms of forecasting accuracy. Indeed, the benefit from incorporating the main business cycle drivers (also conditional on the imposition of a non-economic prior) as additional observables in reduced–form models is quite limited[4]. In addition, we also show that (iv) forecasts of real variables produced by combinations of structural and non–structural models, especially the DSGE-VAR and the BVAR models, are generally optimal and rational (i.e. unbiased and efficient) in absolute terms. Looking across the sample, we find that (v) the relaxation of the DSGE prior is particularly effective in some specific periods, coherent with the dynamics of the observables first moments.

The remainder of the paper is organised as follows. The DSGE model for Spain is presented in Section 2 along with a description of the data. Reduced-form models are introduced in Section 3. In Section 4 we describe hybrid models: DSGE-VAR and Augmented-(B)VAR-DSGE. In Section 5 we evaluate the pseudo real-time out-of-sample forecasting performance of the estimated models. The final section concludes.

2 The Smets-Wouters model for Spain

We begin this section by describing the main features of the linearised SW DSGE model. For the sake of brevity, we limit ourselves to a broad-level overview of this model taken to the Spanish economy. We then provide details regarding data construction and document the estimation results. The dataset employed is homogeneous across all alternative structural and non-structural models.

[4] Yet, out-of-sample forecasting accuracy is not significant by itself. Section 5 provides further investigation on equal predictive accuracy tests.

2.1 Model Formulation

As we will draw extensively from Del Negro et al. (2007), we stick rigorously to their notation. The original model formulation is coherent with a balanced growth path driven by a stochastic and persistent technological progress. However, given the clear rejection of the implicitly postulated cointegrating relationships[5], we opt for a deterministic trend in technology growth around which real variables fluctuate, thus ruling out the existence of any permanent shock affecting the variables levels. Beyond this, the model embeds many nominal and real frictions shaping representative household and firm optimal decisions.

2.1.1 Households Optimal Choices

Consider an infinite horizon representative consumer with a separable utility function defined over consumption goods (logarithmic and with external habits) and leisure so that saving choices are not directly affected by labour decisions. Labour supplied is differentiated by a working union which exerts some monopolistic power over wages turning into an explicit equation suitable for the introduction of sticky nominal wages in the fashion of Calvo (1983). Additionally, households rent capital services to firms and base their capital accumulation decision on the investment adjustment costs they face.

Consumption dynamics are disciplined by the following Euler equation

$$c_t = \frac{\bar{\gamma}h}{\bar{\gamma}^2 + h^2\beta}c_{t-1} + \frac{\bar{\gamma}h\beta}{\bar{\gamma}^2 + h^2\beta}E_t c_{t+1} - \frac{(\bar{\gamma} - h\beta)(\bar{\gamma} - h)}{\bar{\gamma}^2 + h^2\beta}\chi_t + \left(1 - \frac{\beta h \rho^b}{\bar{\gamma}}\right)\varepsilon_t^b \tag{1}$$

where $\bar{\gamma}$ is the (steady state) deterministic growth rate of real variables, β is the usual stochastic discount factor, and h gives the extent of consumption habit persistence. Equation (1) implies that current consumption, c_t, depends on a weighted average of previous and future expected consumption, and the ex-ante real interest rate implied by the marginal utility of consumption, χ_t, plus the current minus the expected intertemporal preference shock, ε_t^b. The latter introduces an exogenous stochastic disturbance within the optimal saving decision. The intertemporal preference shock is assumed to follow the usual autoregressive process: $\varepsilon_t^b = \rho_b\varepsilon_{t-1}^b + \eta_t^b$.

Marginal utility of consumption is also standard

$$\chi_t = \chi_{t+1} + r_t - E_t\pi_{t+1} \tag{2}$$

[5] In this regard see Section A.2 in appendix.

Investment dynamics are dictated by the investment Euler equation

$$i_t = \frac{1}{1+\beta} i_{t-1} + \frac{\beta}{1+\beta} E_t i_{t+1} + \frac{1}{\bar{\gamma}^2 s'(1+\beta)} (\chi_t^k - \chi_t) + \varepsilon_t^\mu \tag{3}$$

where s' is the steady-state elasticity of the investment adjustment cost function. Also in this case current investment, i_t, is a function of past and expected future investment choices other than the usual Tobin's Q (i.e., $q_t = \chi_t^k - \chi_t$), and the investment specific technology shock whose exogenous process also follows the usual autoregressive form: $\varepsilon_t^\mu = \rho^\mu \varepsilon_{t-1}^\mu + \eta_t^\mu$.

Finally, the arbitrage equation for the value of capital boils down to

$$\chi_t^k = \frac{r^{k,*}}{r^{k,*} + 1 - \delta} \left(E_t \chi_{t+1} + E_t r_{t+1}^k \right) + (1-\delta) E_t \chi_{t+1}^k \tag{4}$$

where $r^{k,*} = \bar{\gamma}/\beta - (1-\delta)$ is the steady state rental rate of capital and δ is the capital depreciation rate. Equation (4) states that the current value of capital, χ_t^k, is determined by its future expected value and the ex-ante real interest rate along with the capital expected remuneration.

2.1.2 Goods Market

The problem of the representative firm is to find optimal values for the utilisation of labour and capital services given the following production function of the Cobb-Douglas form

$$y_t = \alpha k_t^s + (1-\alpha) l_t + \varepsilon_t^z \tag{5}$$

where inputs, capital services (k_t^s) and hours worked (l_t), and TFP dynamics ($\varepsilon_t^z = \rho^z \varepsilon_{t-1}^z + \eta_t^z$) are standard. In this regard, α is the capital share of income.

Given the presence of utilisation costs, the capital services used in current production, k_t^s , are a function of the stock of capital installed in the last quarter and the degree of capital utilisation, u_t, that is

$$k_t^s = k_{t-1} + u_t. \tag{6}$$

Then, by household optimisation choice it turns out that

$$u_t = \frac{r^{k,*}}{a''} r_t^k \tag{7}$$

so that the rate of capital utilisation is proportional to the capital remuneration. Moreover, a'' is the elasticity of capital utilisation cost.

The law of motion of installed capital

$$k_t = \left(1 - \frac{i^*}{k^*}\right) k_{t-1} + \frac{i^*}{k^*} i_t + \frac{i^*}{k^*}(1+\beta)\bar{\gamma}^2 s' \varepsilon_t^{\mu} \tag{8}$$

does not only depend on the flow of investment but also on the investment specific technology shock. Where i^*/k^* is the investment to capital steady state ratio.

Firms' cost minimisation yields:

$$mc_t = (1-\alpha)w_t + \alpha r_t^k - \varepsilon_t^z \tag{9}$$

which crucially depends on the factor price and TFP dynamics.

Prices are sticky but similarly to Del Negro et al. (2007) we abstract from any dynamic price indexation mechanism[6]. Inflation dynamics are summarised by means of the New-Keynesian Phillips curve

$$\pi_t = \beta E \pi_{t+1} + \frac{(1-\xi_p)(1-\beta\xi_p)}{\xi_p} mc_t + \varepsilon_t^{\lambda_f} \tag{10}$$

in this regard, ξ_p represents the duration of the Calvo contracts. Inflation is a positive function of future expected inflation, the current price marginal cost, and depends positively also on the price mark-up shock, $\varepsilon_t^{\lambda_f}$. The exogenous process defining such disturbance is : $\varepsilon_t^{\lambda_f} = \rho^{\lambda_f} \varepsilon_{t-1}^{\lambda_f} + \eta_t^{\lambda_f}$.

Finally, firms' optimal factors choice implies that the rental rate of capital is an inverse function of the capital to labour ratio but is increasing in the real wage

$$r_t^k = -(k_t^s - l_t) + w_t. \tag{11}$$

2.1.3 Labour Market

Given its monopolistically competitive set up, the functioning of the labour market resembles that of the goods market. In this regard, its functioning is described by the optimal wage decision

$$\begin{aligned}\left[1 + \xi_w \beta \nu_l \frac{(1+\lambda_w)}{\lambda_w}\right] w_t = {} & \xi_w \beta \left[1 + \nu_l \frac{(1+\lambda_w)}{\lambda_w}\right] (\widetilde{w}_{t+1} + w_{t+1}) - \left[1 + \nu_l \frac{(1+\lambda_w)}{\lambda_w}\right] \widetilde{w}_t \\ & + (1-\xi_w\beta)\nu_l l_t - (1-\xi_w\beta)\chi_t + \xi_w \beta \left(1 + \nu_l \frac{1+\lambda_w}{\lambda_w}\right) \pi_{t+1} \\ & + \varepsilon_t^{\phi} + (1-\xi_w\beta)(\bar{\gamma}^2 + h^2\beta) \frac{1}{\bar{\gamma}(\bar{\gamma}-h)} \varepsilon_t^b \end{aligned} \tag{12}$$

[6] Del Negro et al. (2007) and Del Negro and Schorfheide (2006) show that abstracting from such a feature does not impair the model's predictive ability.

and the aggregate wage evolution

$$w_t = \frac{1-\xi_w}{\xi_w}\widetilde{w}_t - \pi_t + w_{t-1} \tag{13}$$

As before, we abstract from any dynamic wage indexation mechanism. ξ_w is the average wage contract duration, ν_l the labour supply elasticity, and λ_w is the net wage markup. $\widetilde{w}$ is the flexible price optimal wage. As usual the current wage w depends positively on the expected future real and flexible price wage, hours worked and expected inflation, and depends negatively on the marginal utility of consumption and current flexible price wage. $\varepsilon_t^{\phi} = \rho^{\phi}\varepsilon_{t-1}^{\phi} + \eta_t^{\phi}$ is an intratemporal preference (labour supply henceforth) shock introducing a stochastic exogenous wedge in the optimal labour-consumption decision. This is of course different from ε_t^b which is intertemporal. Positive realisations of both, however, increase the current real wage on impact.

2.1.4 Equilibrium

Households are Ricardian and public expenditure evolves exogenously: $\varepsilon_t^g = \rho^g\varepsilon_{t-1}^g + \eta_t^g$.

Markets clear so that the aggregate resource constraint is

$$y_t = \frac{c^*}{c^*+i^*}c_t + \frac{i^*}{c^*+i^*}i_t + \frac{r^{k,*}k^*}{c^*+i^*}u_t + \varepsilon_t^g \tag{14}$$

where c^* is the steady-state level of consumption.

Finally, a monetary rule is needed to close the model and this becomes a non-trivial issue for a euro area (EA henceforth) country member. In principle, the European Central Bank sets its monetary policy according to the EA fundamentals. This would probably call for a more articulate modelling choice of the external sector which however abstracts from the scope of this paper. In this regard, we might take action in two ways, either we posit that Spain has a domestic monetary authority or we assume that the Taylor rule is completely exogenous from Spanish fundamentals and is tuned according to EA inflation and GDP dynamics. As long as for the implementation of the latter approach we would need to include EA inflation and GDP as additional observables, we strike a balance and therefore opt for the following monetary policy rule:

$$r_t = \rho^r r_{t-1} + (1-\rho^r)\left(\psi_1\pi_t + \psi_2 y_t\right) + \eta_t^r \tag{15}$$

The central bank gradually adjusts the policy rate in response to inflation, whose ψ_1 is the rule weight, and output deviations from its deterministic trend, where ψ_2 is its weight in the Taylor rule. In this regard ρ^r describes the degree of interest rate smoothing. Finally, η_t^r is the monetary policy shock.

2.2 Solving the DSGE model

The model has fifteen endogenous variables: $y_t, c_t, i_t, \chi_t^k, \chi_t, k_t^s, k_t, u_t, r_t^k, mc_t, \pi_t, w_t, \widetilde{w}_t, l_t$, and r_t. Their dynamics are described from equation (1) to (15). Finally, the dynamics of the system are led by seven exogenous disturbances: $\varepsilon_t^a, \varepsilon_t^b, \varepsilon_t^i, \varepsilon_t^g, \varepsilon_t^p, \varepsilon_t^\phi, \eta_t^r$, whose error terms are all normal and identically, independently distributed.

2.3 Data, calibration and estimation

2.3.1 The data

The observables in the model are equivalent to those employed in SW: Gross Domestic Product (GDP_t), consumption ($CONS_t$), investment (INV_t), wage (WAG_t), inflation (P_t), nominal interest rate ($INT.RATE_t$), and hours worked ($HOURS_t$).

Data for GDP, private consumption, investment, hours worked and compensation per hour worked are all taken from the quarterly national accounts, as compiled by the Spanish national statistical institute (INE). Investment refers to total gross fixed capital formation (i.e. all non-financial productive assets and all sectors). Consumption, investment and compensation are deflated using the GDP deflator. GDP, consumption, investment and hours worked are all defined in per capita terms by dividing by population, also sourced from the quarterly national accounts. All series are seasonally and working day adjusted. Inflation refers to the first difference of the log of the GDP deflator. The interest rate is the Spanish interbank overnight offered rate. Consumption, investment, GDP, wages and hours are expressed in 100 times log. The interest rate and inflation rate are expressed on a quarterly basis. The sample time span is from Q3:1980 to Q4:2015.

The SW DSGE model considers seven observables and seven structural shocks, therefore there is no need to add any measurement error. The corresponding set of measurement equations in the model is:

$$Y_t = \begin{bmatrix} dlGDP_t \\ dlCONS_t \\ dlINV_t \\ dlWAG_t \\ dlP_t \\ INT.RATE_t \\ lHOURS_t \end{bmatrix} = \begin{bmatrix} \gamma \\ \gamma \\ \gamma \\ \gamma \\ \bar{\pi} \\ \bar{r} \\ \bar{l} \end{bmatrix} + \begin{bmatrix} y_t - y_{t-1} \\ c_t - c_{t-1} \\ i_t - i_{t-1} \\ w_t - w_{t-1} \\ \pi_t \\ r_t \\ l_t \end{bmatrix} \tag{16}$$

where l and dl represents 100 times log and log difference, respectively; $\gamma = 100\ln(\bar{\gamma})$ is the quarterly balanced growth path net growth rate of GDP, consumption, investment and wage; $\bar{\pi} = 100\ln(\pi^*)$ is the quarterly steady-state inflation rate and $\bar{r} = 100\ln(\beta^{-1}\bar{\gamma}\pi^*)$ is the steady state nominal interest rate.

From now onward, for sake of simplicity, we will refer to $dlGDP_t$,$dlCONS_t$,$dlINV_t$ and dlP_t as ΔY_t, ΔC_t, ΔI_t, and π_t, respectively.

2.3.2 Calibration

As in Del Negro et al. (2007), most of the parameters are estimated, only few are calibrated. They are the capital depreciation rate ($\delta = 0.025$) and the share of fixed costs in production ($\Phi = 0$). In addition, we also impose the Taylor rule weight on output deviation from its trend, ψ_2, to be 0.[7] Given the peculiarity of monetary policy conduct for a single country as a member of the eurozone, this seems a reasonable assumption.[8]

2.3.3 Estimation

The rest of the parameters are estimated using Bayesian methods.[9] Priors elicitation is the same as in Del Negro et al. (2007) with just a few case specific changes due to identification issues. The investment adjustment cost, s', prior standard deviation is slightly smaller (1.25 rather than 1.50) whereas the mean of the labour supply shock prior persistence, ρ^{ϕ}, has been increased (from 0.8 to 0.9) and the standard deviation of σ^{ϕ} has been tightened (from 2 to 1.50).

In the first three columns of Tables 1A and 1B, the priors elicitation is summarised. From columns four to six, the DSGE parameters posterior estimates are reported.

[7] This is done to reduce the degree of misspecification of the monetary policy rule.

[8] However, in all the versions of the SW for EA countries model such parameter magnitude is hardly larger than 0.1, see for instance Del Negro et al. (2007).

[9] Bayesian inference is carried out by means of the Metropolis-Hastings algorithm with one chain of 110,000 draws discarding the first 10,000 draws.

TABLE 1A – Structural parameters estimates (prior and posterior distributions).

	Prior Distribution			*DSGE* Posterior			*DSGE–VAR*($\hat{\lambda}$) Posterior		
	Distr.	Mean	St. Dev.	Mean	5%	95%	Mean	5%	95%
α	Beta	0.30	0.10	0.169	0.127	0.211	0.158	0.107	0.216
ξ_p	Beta	0.75	0.10	0.681	0.626	0.734	0.632	0.552	0.725
s'	Gamma	4.00	1.25	4.920	3.791	6.191	4.210	3.603	4.730
h	Beta	0.80	0.10	0.573	0.501	0.643	0.707	0.611	0.793
a''	Gamma	0.25	0.08	0.260	0.221	0.303	0.259	0.219	0.302
ν_l	Gamma	2.00	0.25	1.710	1.347	2.010	1.851	1.495	2.184
ξ_w	Beta	0.75	0.10	0.254	0.185	0.329	0.326	0.233	0.433
λ_w	Gamma	0.30	0.10	0.438	0.280	0.626	0.438	0.264	0.684
r^*	Gamma	0.50	0.10	0.454	0.337	0.577	0.459	0.330	0.600
ψ_π	Gamma	1.70	1.00	1.676	1.546	1.814	1.700	1.544	1.885
ρ^r	Beta	0.80	0.10	0.775	0.735	0.810	0.829	0.791	0.864
$\bar{\pi}$	Normal	0.65	0.20	0.935	0.722	1.145	0.769	0.499	1.062
γ	Gamma	0.50	0.25	0.234	0.198	0.273	0.363	0.287	0.439
g^*	Beta	0.15	0.05	0.127	0.065	0.206	0.140	0.072	0.230
$\bar{L}$	Normal	510.00	10.00	510.269	505.211	515.209	510.192	509.687	510.683

Note: The posterior distributions are obtained using the Metropolis-Hastings algorithm.

TABLE 1B – Shocks processes estimates (prior and posterior distributions).

	Prior Distribution			*DSGE* Posterior			*DSGE–VAR*($\hat{\lambda}$) Posterior		
	Distr.	Mean	St. Dev.	Mean	5%	95%	Mean	5%	95%
ρ^z	Beta	0.80	0.05	0.955	0.9345	0.972	0.812	0.722	0.890
ρ^ϕ	Beta	0.90	0.05	0.956	0.938	0.972	0.958	0.919	0.983
$\rho^{\lambda f}$	Beta	0.80	0.05	0.895	0.841	0.936	0.765	0.634	0.872
ρ^μ	Beta	0.80	0.05	0.825	0.761	0.882	0.691	0.596	0.782
ρ^b	Beta	0.80	0.05	0.835	0.782	0.881	0.756	0.668	0.831
ρ^g	Beta	0.80	0.05	0.973	0.958	0.984	0.877	0.806	0.934
σ^z	Inv.Gam.	1.00	2.00	0.622	0.557	0.695	0.410	0.351	0.477
σ^ϕ	Inv.Gam.	1.00	1.50	3.900	3.212	4.704	2.100	1.717	2.562
$\sigma^{\lambda f}$	Inv Gam.	1.00	2.00	0.242	0.210	0.278	0.228	0.196	0.262
σ^μ	Inv.Gam.	1.00	2.00	0.336	0.291	0.387	0.319	0.269	0.376
σ^b	Inv Gam.	0.20	2.00	0.669	0.563	0.791	0.254	0.191	0.335
σ^g	Inv.Gam.	0.30	2.00	0.638	0.579	0.704	0.282	0.229	0.339
σ^r	Inv.Gam.	0.20	2.00	0.320	0.283	0.360	0.179	0.150	0.213

Note: The posterior distributions are obtained using the Metropolis-Hastings algorithm.

Despite a prior mean of 0.50 and a sample mean of rougly 0.40, the trend growth rate of real variables, γ, is estimated to be just 0.234. This could reflect the difficulties the model encountered trying to deal with the financial crises occurring in our sample. The estimated inflation growth rate, $\bar{\pi}$, is instead slightly higher than both the sample and prior mean, whilst the nominal interest rate is less than 2% on an annual basis, broadly in line with the sample mean.

Similar to SW, the data appear to be very informative about the stochastic processes for the exogenous shocks. In particular the TFP, labour supply and public expenditure exogenous processes turn out to be very persistent.

Finally, the posterior mean of the structural parameters is not too far from the prior, with the only exception of (i) wage stickiness, ξ_w, which is estimated to be surprisingly low (0.254, just one third of the prior mean); and (ii) the wage markup, λ^w, which turns out to be quite large (0.438). Then, also the estimated mean of the investment adjustment cost is higher than its prior mean

(4.920) and the capital share of income is quite lower (0.169) instead.

3 Reduced-Form Models

In this section, we briefly describe the procedure for estimating the non-structural (unrestricted) statistical models: (i) vector autoregressive model of order p, VAR(p); (ii) a Bayesian VAR(p); and (iii) an autoregressive model of order p, AR(p). We follow an agnostic approach and no identification strategy is imposed.[10]

The VAR(p) model is given by a general specification for an N-dimensional vector of time-series y_t which can be defined in compact form as:

$$y_t = C_t + \phi_1 y_{t-1} + ... + \phi_p y_{t-p} + z_t, \quad z_t \sim N(0, \Sigma) \tag{17}$$

where $\phi_1, ..., \phi_p$ are NxN matrices of coefficients on the p-lags of the variables, and C_t is an N-dimensional vector of constants, time trends, or exogenous data series.[11] The reduced-form innovations are collected in the vector z_t , which is assumed to be normally distributed, Σ being the covariance distribution of the VAR errors. In line with SW, $y_t = (y_{1,t}, ..., y_{K,t})'$ is the vector of the K=7 observables and the model is estimated on a rolling windows basis with a lag length (p) equal to 4. Coefficients vary at each estimation window, accordingly.

The Bayesian VAR(p) has the same structure as the VAR(p) but assuming a prior distribution of the parameters, which shrinks the parameters towards univariate autoregressive processes (in our case, a Normal-Wishart prior). The hyperparameters calibration is standard: (i) the autoregressive coefficient is set to 0.8; (ii) the overall tightness hyperparameter, $\lambda_1 = 0.1$; (iii) the cross-variable weighting, $\lambda_2 = 0.5$; and (iv) the lag decay, $\lambda_3 = 2$.[12]

The autoregressive model of order $p = 1$, AR(1), follows from VAR(1) with $K = 1$ for each observable.

4 DSGE Hybrid Models

In this section, we define two different types of DSGE hybrid models. They refer to those models combining elements, or information, from structural and reduced–form models. In this regard,

[10] All models employ the same series as in the left-hand side of Eq. (16).

[11] The combination of variables in levels and growth rates may lead to problems with trends and cointegrating relationships. Yet, stationarity has been guaranteed (when needed).

[12] As robustness checks, Bayesian VAR estimates are largely unchanged by exploiting different sets of non-economic priors.

our approach is twofold. First, based on the seminal work of Del Negro and Schorfheide (2004), the DSGE model is used as an economic prior for estimating a non–structural model (i.e., VARs). Such class of models is labelled by the related literature as DSGE–VAR. The second type of DSGE hybrid model is based on the methodology proposed by Fernández-de-Córdoba and Torres (2011), where the VAR information set is broadened to include those series retrieved from the DSGE model as additional observables. We refer to this class of models as Augmented-(B)VAR-DSGE.

4.1 DSGE–VAR

The DSGE–VAR model estimation, based on Del Negro and Schorfheide (2004), implicitly restricts the DSGE model parameters space so that it is possible to draw posterior inferences. They aimed at implicitly searching the set of DSGE parameters for which the distance between the VAR estimate and the vector autoregressive representation of the DSGE model is sufficiently small. In this regard, the hyperparameter λ, whose bounds are such that $\lambda \in [\lambda^{MIN}, +\infty)$, reflects the extent to which the DSGE model performs in terms of forecasting accuracy. The smaller is λ, the worse the DSGE model fits the data.

Loosely speaking, Del Negro and Schorfheide (2004) state that the economic prior can be thought of as fitting a VAR on artificial time series data simulated from a DSGE, thus yielding the reduced-form representation of the structural model of interest. Then, estimating a VAR on a combination of the dummy and true observations would yield the DSGE–VAR parameters. The main drawback, they suggest, is that repeated applications of such a procedure would lead to an unpleasant stochastic variation in the prior distribution. In order to eschew the occurrence of any stochastic noise, the empirical moments stemming from the artificially generated data are replaced by their theoretical counterpart extracted from an approximation of the VAR representation of the DSGE. The extent to which this approach is reliable rests on the fact that the above finite order VAR approximation of the DSGE is close enough to its infinite lags counterpart.

For the sake of clarity, let us consider ϕ and Σ_u as the VAR parameters and θ the DSGE model parameters.[13] Overall our prior has the hierarchical structure

$$p(\phi, \Sigma_u, \theta) = p(\phi, \Sigma_u|\theta)p(\theta)$$

Therefore the DSGE–VAR posterior can be factorised in the posterior density of the VAR parameters given the DSGE model parameters and the marginal posterior density of the DSGE model parameters.[14]

[13]The notation adopted in this section strictly follows Del Negro and Schorfheide (2004).

[14]Thus, in a sense, this procedure is equivalent to estimating a Bayesian VAR whose priors are assessed by means

$$p(\phi, \Sigma_u, \theta|Y) = p(\phi, \Sigma_u|Y, \theta)p(\theta|Y)$$

without any loss of generality, this boils down to

$$p(\phi, \Sigma_u, \theta|Y) = p(\phi, \Sigma_u|Y)p(\theta|\phi, \Sigma_u)$$

giving the extent to which the DSGE–VAR estimation allows to draw posterior inferences about the DSGE model parameters.

Then, let $Y'X$, $X'X$, $Y'Y$ be the actual sample moments, and let $\lambda T\Gamma^*_{yx}(\theta)$, $\lambda T\Gamma^*_{xx}(\theta)$ and $\lambda T\Gamma^*_{yy}(\theta)$ be the expected values of the (scaled) population moments of the DSGE artifical data.

It turns out that the maximum-likelihood estimate of ϕ and Σ_u, i.e., $\widetilde{\phi}(\theta)$ and $\widetilde{\Sigma}_u(\theta)$, are obtained considering both the artificial and actual sample, that is

$$\widetilde{\phi}(\theta) = \left(\lambda T\Gamma^*_{xx}(\theta) + X'X\right)^{-1}\left(\lambda T\Gamma^*_{xy}(\theta) + X'Y\right) \tag{18}$$

$$\begin{aligned}\widetilde{\Sigma}_u(\theta) = &\frac{1}{(\lambda+1)T}\Big[\left(\lambda T\Gamma^*_{yy}(\theta) + Y'Y\right) \\ &- \left(\lambda T\Gamma^*_{yx}(\theta) + Y'X\right)\left(\lambda T\Gamma^*_{xx}(\theta) + X'X\right)^{-1}\left(\lambda T\Gamma^*_{xy}(\theta) + X'Y\right)\Big]\end{aligned} \tag{19}$$

Since conditional on θ the DSGE model prior and the likelihood function are conjugate, it can be shown that the posterior distribution of ϕ and Σ_u is of the Inverted Wishart–Normal form

$$\Sigma_u|Y, \theta \sim \mathcal{IW}\left((\lambda+1)T\widetilde{\Sigma}_u(\theta), (1+\lambda)T - k, n\right) \tag{20}$$

$$\phi|Y, \Sigma_u, \theta \sim \mathcal{N}\left(\widetilde{\phi}(\lambda), \Sigma_u \otimes \left(\lambda T\Gamma^*_{xx}(\theta) + X'X\right)^{-1}\right) \tag{21}$$

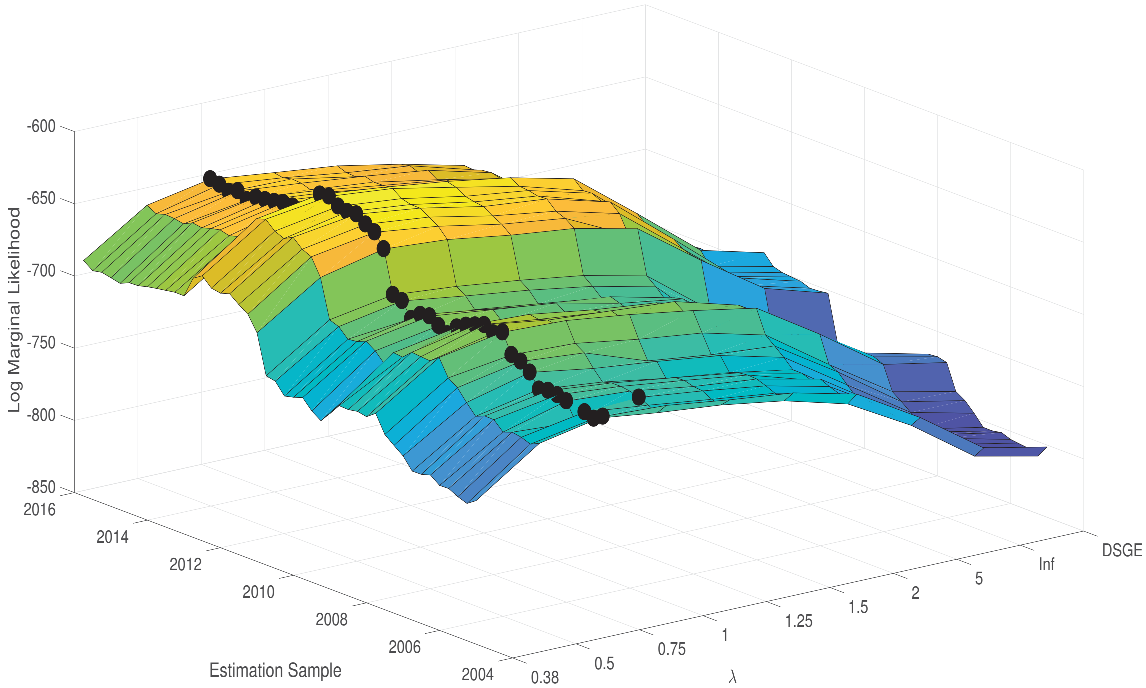
4.1.1 Optimal economic prior tightness

In principle, an optimal λ can be chosen to maximize the log of the marginal data density. However, since our sample involves several episodes of financial disruption, it is of high interest to ensure that the degree of shrinkage of the imposed economic prior is relatively stable along the rolling estimation schemed adopted. In this regard, Figure 1 displays the dynamic evolution of the DSGE–VAR log-marginal likelihood over a grid of values for λ and along the rolling estimation sample evolution. Black dots identify the highest log-marginal likelihood for each rolling sample across

of a DSGE model. In this regard it is possible to refer to λ as the prior weight.

several values of λ. It turns out that the optimal DSGE prior weight is $\hat{\lambda} = 0.75$.[15]

FIGURE 1. Dynamic DSGE prior weight evolution.



Note: Dynamics DSGE-VAR log-marginal likelihood evaluation over different values of λ following the sample rolling window scheme. The sample spans from 1981q3 to 2015q4, with a window size equal to 95 periods. Black dots identify the highest log marginal likelihood for each of the estimation samples.

Quite surprisingly, but similarly to Del Negro and Schorfheide (2006), the typical inverse U-shape pattern of the DSGE–VARs log-marginal likelihood is very stable over time. The implications to be drawn are twofold. First, if DSGE imposed restrictions cannot be completely accepted, they cannot be fully rejected either. Second, since it turns out that the optimal prior weight is always $\hat{\lambda} = 0.75$ (except for one single evaluation sample), this means that more difficult evaluation samples including the financial crises do not call for the imposition of a looser economic prior. Further, we highlight that the log-marginal likelihood of the $DSGE{-}VAR(\lambda = \infty)$, i.e. the finite lag order VAR representation of the $DSGE$, is fairly close to the true $DSGE$ one hinting that the truncation bias is of a fairly small degree.

Finally, for benchmarking purposes, a brief comparison to Del Negro et al. (2007) is due since it is the most closely related approach. As already mentioned, originally their methodology relies on a specific cointegrating framework (i.e., DSGE–VECM model) for the euro area by means of the estimation of the SW DSGE model to investigate its in-sample fitting and forecasting performance.

[15]The lowest possible value of λ, i.e. $\lambda^{MIN} = 0.38$, is such that the economic prior is proper and non-degenerate.

Yet it is worthwhile to highlight the fact that the cointegrating restrictions they imposed are essentially rejected also by the US data, inducing a worse forecasting performance of the model (although they seem to help for impulse responses identification). As far as the main interest of this paper is forecasting the Spanish business cycle, and since the imposition of the above cointegrating restrictions fits particularly badly with the Spanish data[16], the canonical DSGE–VAR formulation is considered.

4.1.2 DSGE–VAR estimation results

Posterior $DSGE\text{–}VAR(\hat{\lambda})$ estimation results are reported in Tables 1A and 1B (columns 7-9). As already mentioned, the distance between DSGE and DSGE–VAR parameters is instructive about the DSGE model misspecification degree. In particular, it emerges that the degree of habit in consumption, h, was understated by the DSGE model estimation by almost 25%. Also the wage contract duration, ξ_k, is now longer in the DSGE–VAR than before. The deterministic growth rate of the economy, γ, has increased as well, being now more in line with the sample mean. Coming to the exogenous processes, their persistence is now sensibly reduced except for ρ^{ϕ} which is virtually unchanged. Finally, the shocks' standard deviations, apart from the IST and price mark-up shocks, have all been revised downwards, particularly the public expenditure and labour supply shocks, which are now considerably smaller than previously estimated.

Finally, a remark is in order. These results do not have any pretense of being general, indeed they must be evaluated in the precise context of this analysis and specifically to our dataset.

4.2 Augmented–(B)VAR–DSGE

The Augmented–(B)VAR–DSGE estimation procedure is inspired by the work of Fernández-de-Córdoba and Torres (2011). The key idea behind this approach is that an unrestricted (B)VAR contains only limited information on business cycle determinants in contrast with the rich structural dynamics disciplined by DSGE models. Thus, a natural step to enrich the (B)VAR information content is to include unobserved variables produced by the DSGE model. However, our modelling approach differs from that of Fernández-de-Córdoba and Torres (2011) at least along three key dimensions. Firstly, the underlying DSGE model we use to augment the (B)VAR variables dimension is far richer than the simple Real Business Cycle (RBC) model they use. Secondly, our framework allows us to choose from among many different unobserved variables exploiting the features of the underlying structural model. Third, the way artificial data is extracted from the DSGE model is different, since it exploits more deeply the advantages of confronting the structural

[16] In this regard see Section A.2 in Appendix.

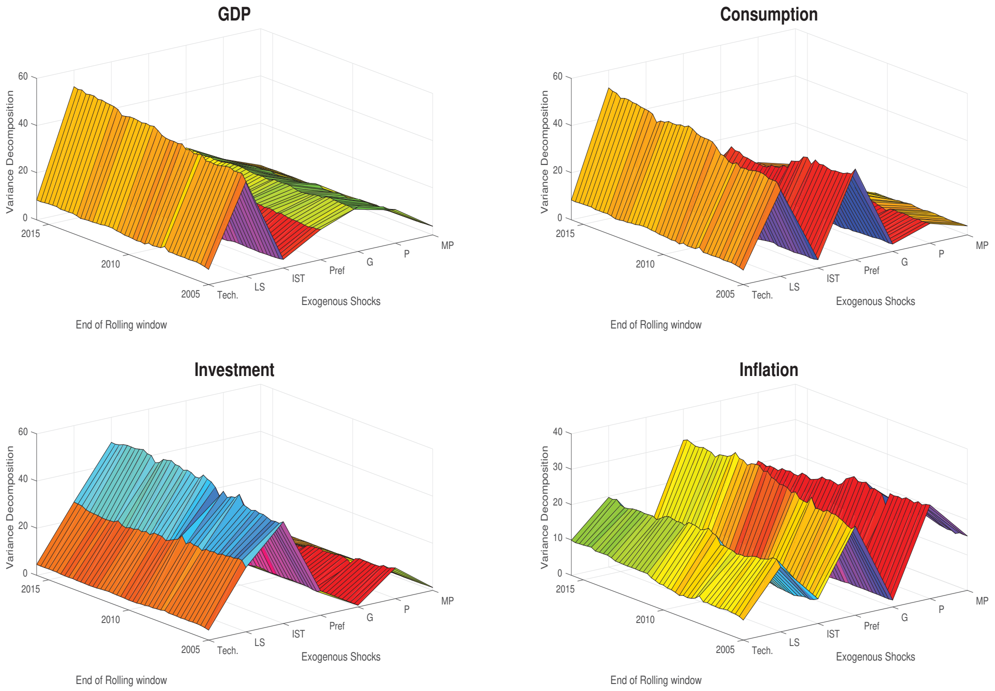
model with the data. A set of different structural shocks and all their possible combinations are explored to enlarge the (B)VAR estimation space. All structural shocks (as in SW) are considered as unobserved variables: the total factor productivity (Tech.) shock, ε_t^z, the intertemporal preference (Pref.) shock, ε_t^b, the public expenditure (G) shock, ε_t^g, the investment specific technology (IST) shock, ε_t^μ, the monetary policy shock (MP) shock, ε_t^r, the inflation (P) shock, $\varepsilon_t^{\lambda_f}$, and the labour supply (LS) shock, ε_t^ϕ.

The (smoothed) shocks are extracted from the DSGE model following the rolling window estimation, so that the extracted series vary in each vintage of the Augmented–(B)VAR estimation. All priors are homogeneous with those of the BVAR.

4.2.1 DSGE Variance Decomposition and Augmented-(B)VAR Model Selection

In Figure 2, we plot the variance decomposition of shocks for GDP, consumption, investment and inflation over different estimation samples. As shown, the labor supply shock is always among the most relevant business cycle drivers for all of the variables of interest, although investment and inflation are led by the IST and cost push-up shocks, respectively. Overall a relevant role is played by both preference and cost push-up shocks. In this regard we expect the labour supply, preference and inflation shocks to convey the most valuable information to the Augmented-(B)VAR model.

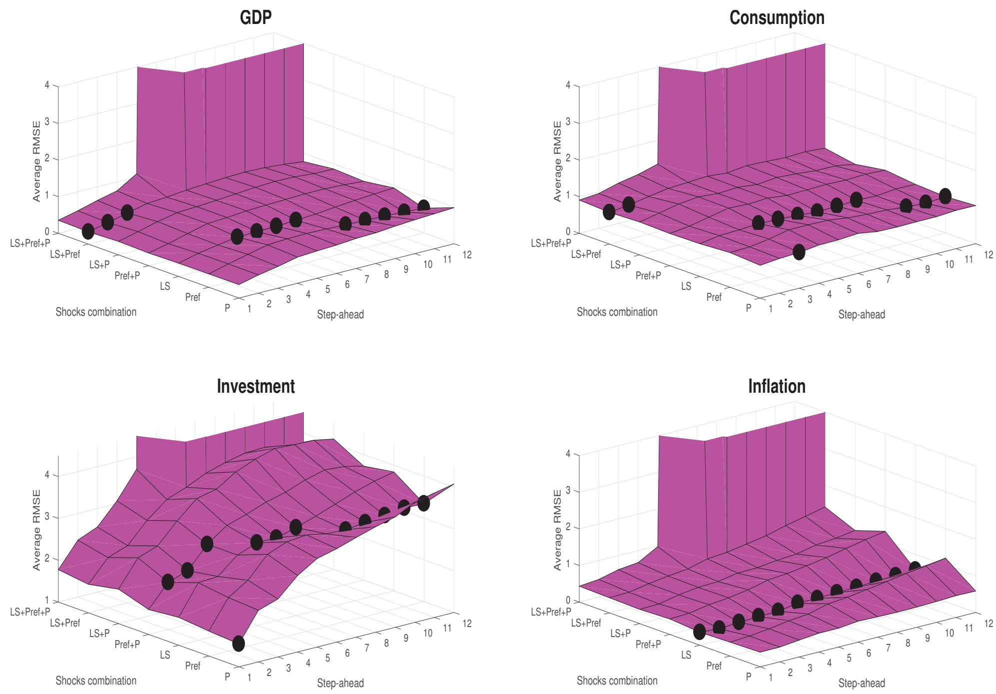
FIGURE 2. Exogenous Shocks Contribution to Each Variable Variance Decomposition (%).



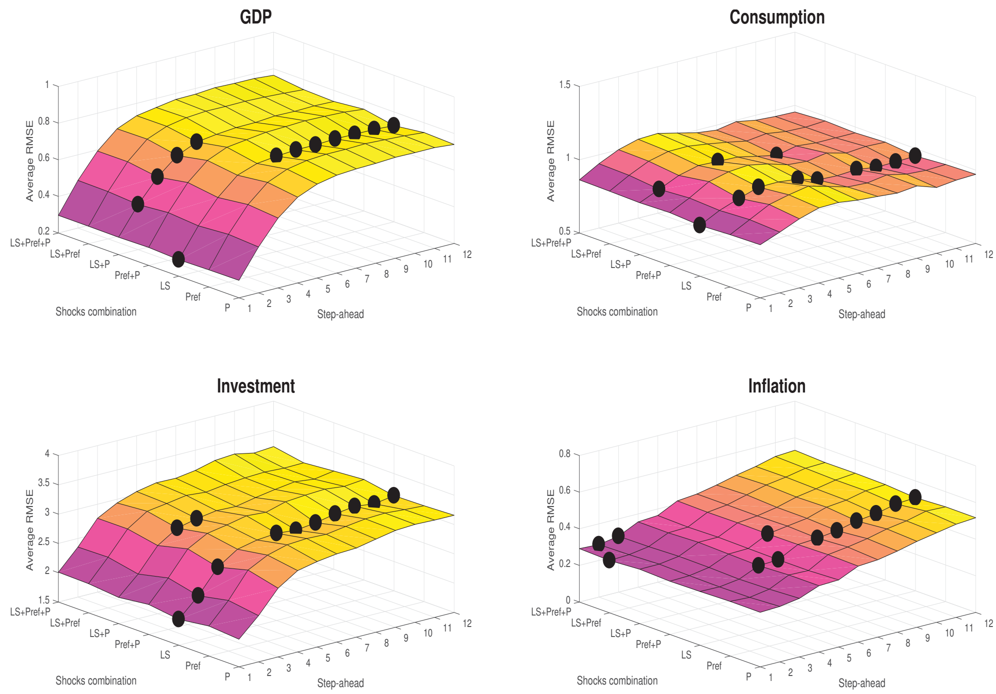
Note: Exogenous shocks (Tech refers to TFP shock, LS refers to labour supply shock, IST refers to investment specific technology shock, Pref refers to preference shock, G refers to public expenditure shock, P refers to cost push-up shock, and MP refers to monetary policy shock) variance decomposition contributions for output growth (ΔY_t), consumption (ΔC_t), investment (ΔI_t) and inflation (π_t) along all the evaluation samples retrieved from the Bayesian estimation of SW DSGE model for Spain.

FIGURE 3. Average RMSEs for different Augmented-(B)VAR models.

(A) Augmented-VAR



(B) Augmented-BVAR



Note: Root Mean Square Forecast Errors (RMSFEs) of GDP (ΔY_t), consumption (ΔC_t), investment (ΔI_t) growth rates and inflation (π_t) along different forecast horizons for several Augmented-(B)VARs. Selected exogenous shocks are LS, Pref and P (LS refers to labour supply shock, Pref refers to preference shock, P refers to cost push-up shock) and all combinations. Black dots identify the lowest average RMSE at each step ahead across different models. The sample spans from 1981q3 to 2015q4. Window size = 95 periods.

Since we integrate the (B)VAR variables space with different sets of additional (artificial) observables, we need to choose adequate metrics in order to proceed with the model selection. This creates some difficulties when we consider the Augmented -BVAR estimation since comparing the log-marginal densities among different models is not feasible as the set of observables is not kept constant. In addition, such metrics should also have the virtue of being constant across both VARs and BVARs estimation. In this regard, we opt to select those models providing the lowest RMSFEs along different forecasting horizons.

In Figure 3, we report the forecasting performance of seven different Augmented-(B)VARs according to all potential combinations of the previously selected smoothed exogenous processes. Black dots identify the most accurate model, in terms of forecasting performance, for each step ahead. Every single Augmented-(B)VAR model differs depending on the unobserved variable considered. As expected, Augmented-VARs forecast errors are clearly explosive when a combination of smoothed shocks is added to the set of observables for the VAR estimation (see Panel A in Figure 3). All in all, our variance decomposition simulation suggests that the labour supply shock conveys the most relevant information for all variables, and particularly for inflation. Surprisingly, the preference shock does not help to improve the forecasting accuracy of consumption dynamics at short horizons and its contribution to the other variables is quite limited. The inflation shock does help to improve investment forecasts at steps ahead h=1 but is constantly taken over by the labour supply shock for inflation itself. Lastly, TFP shocks do not seem to play any particularly prominent role.

Turning to the Augmented-BVAR model, the curse of dimensionality is easily overcome. However the overall picture is virtually unchanged (see Panel B in Figure 3). In this regard we select the Augmented-(B)VAR formulation where only the labour supply smoothed exogenous process is included as additional observable.

5 Forecast evaluation

This section analyses the pseudo real–time out-of-sample performance of the competing models over the four key macroeconomic variables for the Spanish economy.[17] To simulate the predictive accuracy we estimate all competing models over a rolling sample of 95 quarters from 1981Q3 to 2015Q4. The evaluation sample starts at 2005Q2 and ends at 2018Q4, divided into 44 periods with 12–quarters ahead blocks (evaluated at horizons h=1,...,12 quarters ahead).

The absolute and relative accuracy of point forecasts is evaluated as follows. To begin with, we

[17] A purely real–time out–of–sample exercise has not been conducted mainly due to data availability limitations.

proceed with a preliminary comparison of the predictive ability of point forecasts, which is summarized by computing RMSFEs. Secondly, we undertake an evaluation of the absolute predictive ability of each model. We determine the correct specification of forecasting points using several tests. The tests we consider include tests of optimality, in particular, forecast unbiasedness and efficiency tests, commonly referred to as tests of forecast rationality. This is done by means of the Mincer and Zarnowitz (1969) test statistic (MZ henceforth), following the framework provided by West and McCracken (1998). Then, it is interesting additionally to evaluate the absolute predictive ability over time. We thus detect the presence of potential locally non-rational forecasts exploiting the rationality fluctuation test developed by Rossi and Sekhposyan (2016).

In a second step, we compare the relative forecasting performance of all competing models under a two-fold approach. First, we consider tests of relative forecasting performance based on the difference of the RMFSE of competing models by making use of the Diebold and Mariano (1995) test statistic (DM henceforth) and the critical values proposed by Giacomini and White (2006). The performance is judged based on a quadratic loss function. Secondly, in order to tackle the potential presence of parameter instability in the predictive accuracy, we explore each model's relative forecasting stability over time based on the fluctuation test of Giacomini and Rossi (2010).

5.1 A preliminary comparison

An initial comparison of the real-time predictive accuracy of the models is summarized in Figure 4.

We compute the RMSFE of each model m over the full out–of–sample period[18] defined as:

$$RMSFE^{m} \equiv \sqrt{P^{-1}\sum_{t=R}^{T}\hat{L}_{t+h}^{(m)}}$$

for each forecast horizon h=[1,...,12] and each variable: GDP, consumption, investment and inflation. Figure 4 displays the RMSFE for seven different structural, non-structural and hybrid models: (i) VAR; (ii) BVAR; (iii) DSGE; (iv) DSGE–VAR; (v) Augmented–VAR adding the labor supply shock as additional observable (AVAR henceforth); (vi) Augmented–BVAR with the labor supply shock as additional observable (ABVAR henceforth); and finally (vii) a naive AR(1). VAR and alike models have a lag order of four.

Some interesting patterns emerge at first sight. First, our out–of–sample results suggest that the naive AR(1) benchmark performs relatively well when forecasting GDP growth and inflation

[18] The evaluation samples range from 2005Q2 to 2018Q4, for a total of 44 forecast series, each model m is estimated over a rolling window sample of 95 data points

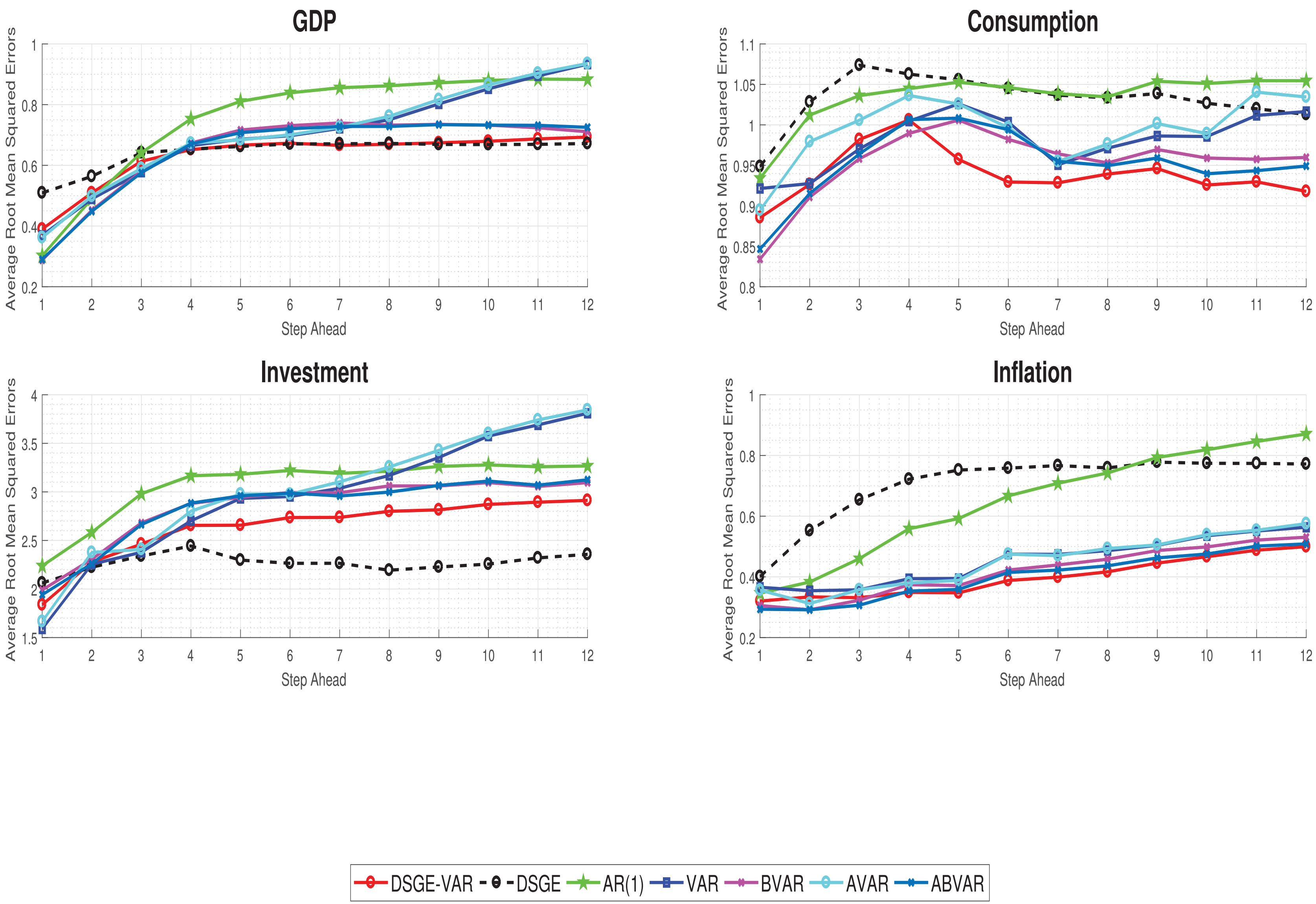
at shorter horizons, but it worsens as the forecast horizon becomes larger. Second, it is worthy to highlight that the DSGE–VAR is very competitive along all dimensions. In particular, it performs extremely well for GDP (h=4:12) consumption (h=5:8) and for inflation, emerging as the best available alternative from step ahead h=4 onwards. Third, coming to the DSGE model, it is the worst alternative for GDP growth forecasting until h=3, becoming then, together with the DSGE–VAR, the best one from step ahead h=4 until the end. The forecast accuracy for consumption is particularly bad along all the steps ahead, whilst surprisingly for investment it is the best performer from step ahead h=3 onward, and by a fairly large extent from step ahead h=5. Finally, its forecasting performance for inflation is, not surprisingly, very bad. Fourth, (B)VARs and A(B)VARs predictive accuracy is almost identical (the only exception is when forecasting consumption for (A)VARs), suggesting that the informative content of the exogenous process retrieved from the DSGE model does not add any valuable information to improve the forecasting properties of reduced-form models.

In general, the elicitation of an economic prior for the VAR model displays particularly evident beneficial effects. Indeed, the DSGE–VAR forecasting accuracy is always comparable to, if not better than, a more canonical BVAR subject to the imposition of a purely statistical prior. In this regard, it is worthy to point out that Fernández-de-Córdoba and Torres (2011) report very large average RMSFEs for the DSGE–VAR.[19]. In their work the underlying structural model was a simple RBC, whilst in our case it has a fairly higher degree of complexity. In this sense, having a more structured DSGE model as economic prior seems to be of great help, especially if we consider that, differently from them, our sample is less suitable to a forecast evaluation exercises since it embeds several episodes of financial market disruptions. By contrast, the opposite seems to be true for the Augmented–(B)VAR: considering a DSGE model with many shocks (and therefore observables) from which exogenous processes are extracted seems to fragment rather than condensate the quality of the information conveyed to the reduced–form models.

We can therefore argue that, until now, there is no consistently "best" forecasting model, as the forecasting methods' relative performance depends both on the variable of interest and the forecast horizon, but the DSGE-VAR model appears to outperform all competing models on average across forecasting horizons and observables. Accordingly, in the following section we investigate formally and in more depth the reason why and the extent to which the DSGE–VAR is (not) preferable to its alternatives.

[19]Our DSGE–VAR model is labelled as VAR-DSGE in Fernández-de-Córdoba and Torres (2011) in spite of the conventional wisdom. In order to avoid confusion, we stick to the usual literature labelling of our hybrid model, that is DSGE–VAR.

Figure 4. Real-time forecasting accuracy, Root Mean Squared Forecast Errors (RMSFEs).



Note: The forecast period is from 2005q2-2008q1 to 2016q2-2018 q4 (Section 5.1 provides the details of the forecast comparison exercise.). The top and bottom panels compare the average RMSFEs for the SW DSGE models, DSGE–VAR ($\hat{\lambda} = 0.5$), DSGE, A(B)VAR, and reduced–form models, (B)VAR and AR(1), for one through four quarters ahead for GDP (ΔY_t), consumption (ΔC_t), investment (ΔI_t) growth rates and inflation (π_t).

5.2 Empirical results based on forecast rationality

5.2.1 Mincer and Zarnowitz's (1969) Tests

The aim of the analysis conducted until now was to initially assess the models forecasting accuracy based on RMSFEs. However, our results abstract from any judgment concerning the evidence of forecasts unbiasedness and efficiency. In this section, we test whether our models produce systematically irrational forecasts.[20] In particular, we test the unbiasedness and efficiency properties of each model's forecasts. To do this, we exploit the MZ (Mincer and Zarnowitz, 1969) rationality test consistently with the framework developed by West and McCracken (1998). This test focuses on the null hypothesis:

$$H_0 : \theta = \theta_0 \text{ vs } H_A : \theta \neq \theta_0, \quad \text{where } \theta_0 = 0$$

θ is the true coefficient from the general regression

$$v_{t+h} = \hat{g_t}' \cdot \theta + \eta_{t+h}, \quad t = R, ..., T \tag{22}$$

where v_{t+h} is the estimated h–step ahead forecast error and $\hat{g}_t$ is a matrix of related forecasts and a constant term, and R is the in–sample size of the dataset. Let $\hat{\theta}_p$ denote the estimate of θ in regression (22) and P be the out–of–sample size of the dataset. Then, let us consider the following Wald test

$$\mathcal{W}_P = P\left(\hat{\theta}_p - \theta_0\right)' \hat{V}_{\theta,P}^{-1}\left(\hat{\theta}_p - \theta_0\right) \tag{23}$$

where $\hat{V}_{\theta,P}$ is a consistent estimate of the long run variance of $\sqrt{P}\hat{\theta}_P$. Then, if the null hypothesis is not rejected the forecast can be assumed to be rational. Table 2 reports the MZ statistics for the variables of interest for the estimated models along increasing forecast horizons.[21] For GDP, apart from steps ahead h=8,12, the DSGE–VAR forecasts can be assumed to be rational. The picture is almost unchanged for DSGE, VAR and BVAR forecasts. For AR model, forecasts are rational only for the first two forecast horizons. For consumption instead, only DSGE-VAR forecasts never reject the null, whilst DSGE's always reject. VAR forecasts do not reject the null for $h >=4$, BVAR predictions instead are rational for $h=$ 1,4,6. Finally, for AR model the null is rejected for $h >2$. By contrast, for investment only DSGE predictions are always rational, except for h=12, while for the others rationality is always rejected apart from DSGE–VAR and VAR at

[20] The models generated forecasts are reported in Figures 19, 20, 21 and 22 in Appendix (Section A.5)

[21] For sake of tractability, we abstract from AVAR and ABVAR inclusion since their performance is in line with BVAR and VAR.

one-step-ahead horizon, and AR for the first two steps ahead. Finally, the inflation forecasts are always biased and inefficient, the only exception is for BVAR at the one-step-ahead horizon.

TABLE 2 – (HAC Robust) Mincer and Zarnowitz's (1969) Test Statistics

	Steps Ahead (h)					
	1	2	4	6	8	12
			GDP(ΔY_t)			
DSGE-VAR	0.12	0.44	1.25	1.45	5.52*	9.14**
DSGE	3.53	0.52	1.76	4.15	6.13**	6.36**
VAR	2.57	3.20	4.52	2.42	6.51**	4.56
BVAR	2.24	2.62	2.83	2.81	5.48*	26.31***
AR	0.87	1.70	6.00**	12.42***	9.38***	7.83**
			CONSUMPTION(ΔC_t)			
DSGE-VAR	3.98	4.10	4.6	2.34	2.14	4.46
DSGE	10.14***	12.98***	11.71***	13.64***	24.65***	39.50***
VAR	16.17***	11.02***	9.74***	4.55	3.76	4.23
BVAR	3.83	5.74*	3.64	4.36	6.27**	10.27***
AR	3.67	4.54	7.26**	17.79***	21.74***	28.05***
			INVESTMENT(ΔI_t)			
DSGE-VAR	3.24	4.79*	5.44*	6.35**	12.86***	8.35**
DSGE	1.32	0.92	1.51	0.55	2.68	13.17***
VAR	3.26	10.06***	11.00***	6.43**	8.29**	8.81**
BVAR	5.58*	6.53**	5.75*	7.67**	13.48***	22.47***
AR	3.50	3.70	21.22***	10.74***	8.34**	6.14**
			INFLATION(π_t)			
DSGE-VAR	8.35**	8.88**	8.35**	23.25***	42.68***	120.29***
DSGE	55.93***	108.21***	304.53***	541.18***	326.82***	359.73***
VAR	25.01***	26.63***	16.97***	29.48***	36.51***	351.09***
BVAR	4.32	9.11**	21.47***	40.50***	26.77***	170.40***
AR	17.07***	40.69***	85.04***	81.67***	80.43***	378.83***

Note: Absolute forecasting rationality performance of considered models (DSGE–VAR, DSGE, VAR, BVAR, AR) for GDP (ΔY_t), consumption (ΔC_t), investment (ΔI_t) growth rates and inflation (π_t) over different forecast horizons. Significance levels are: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

5.2.2 Rossi and Sekhposyan (2016) Rationality Fluctuation test

The MZ rationality test is subject to some caveats. In particular, the systematic biases can average out resulting in the non-rejection of the null of forecast rationality due to a systematic over- and underestimation of the predictions cast. As far as we aim at assessing our models forecasts optimality in the presence of instabilities, we make use of the Rossi and Sekhposyan (2016) fluctuation rationality test. Based a rolling window regression, let $\hat{\theta}_j$ be the parameter estimate in regression (22) but computed at time j over a rolling window of size m.[22] Let the Wald test in regression (22) be now defined as

$$\mathcal{W}_{j,m} = m\hat{\theta}_j'\hat{V}_\theta^{-1}\hat{\theta}_j, \text{ for } j = R+m, ..., T \tag{24}$$

where $\hat{V}_\theta$ is a HAC robust estimator for the asymptotic variance of the parameter estimates in the rolling window. Then, the fluctuation rationality test is $\max_{j\in\{R+m,...T\}} \mathcal{W}_{j,m}$ which is used to test the corresponding null

$$H_0 : \theta_j = \theta_0 \text{ vs } H_A : \theta_j \neq \theta_0, \quad \forall j = R+m, ...T$$

where $\theta_0 = 0$ and θ_j is the true parameter value. Figures 5 and 6 report fluctuation rationality tests for GDP, consumption, investment and inflation over increasing forecast horizons ($h = 1, 2, 4, 6, 8, 12$)[23], where the rolling window size is m=20.[24]

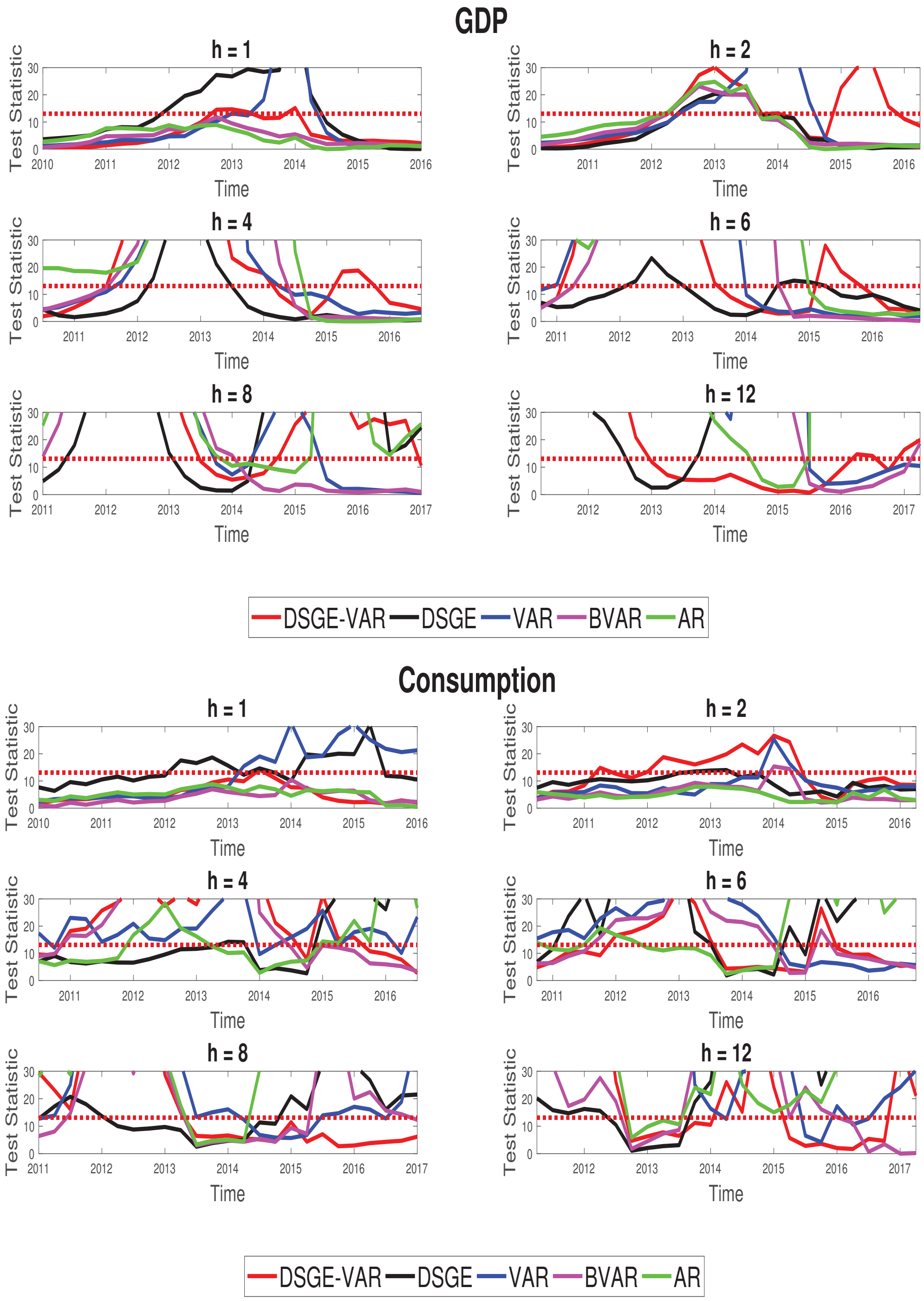
For GDP forecasts (Figure 5, upper panel), our results suggest that the non-rejection of the rationality hypothesis in Table 2 was actually due to the averaging out of the forecast bias in the case of DSGE (black) and VAR (blue) models at the one-step-ahead horizon. Indeed both models do not provide rational forecasts in the central part of the time span. In this regard, the DSGE–VAR (red) accuracy improvement is clear as it turns out to be marginally non-rational only for episodes when the financial crises weighted the most. By contrast, BVAR (magenta) and AR (green) forecasts are always assumed to be rational. All the models are locally non-rational in the central part of the evaluation samples at step ahead h=2 and the DSGE–VAR is also locally non–rational between the second half of 2014 and the first half of 2016. Then, as the forecast horizons increase, the forecasts non-rationality slowly shifts back. For step ahead h=8, only BVAR and VAR are rational from mid-2015 onward. Finally, for step ahead h=12, the DSGE-VAR is rational from 2013 to 2016, whilst from mid-2015 onward only the VAR does not reject the null hypothesis. Turning to consumption (Figure 5, lower panel), the statistical pattern is overall aligned to GDP, apart from the first step ahead where DSGE and VAR predictions are significantly non-rational mainly in the second half of the evaluation samples, and step ahead h=2 where the DSGE-VAR forecasts rationally also in most of the second half of the time span.

[22] $\hat{\theta}_j$ is estimated sequentially for $j = R+m, ...T$ using the m most recent observations, where m is the same for all $\hat{\theta}_j$.

[23] For sake of tractability, results inherent to A(B)VAR models are reported in Section A.4.1 in Appendix

[24] The dashed red line indicates the critical value at the 5% confidence level. However, according to Rossi and Sekhposyan (2016), given the relatively short length of m, the rationality fluctuation test is slightly subject to over-rejection and non-rejections of the null hypothesis are therefore particularly robust. The end of each evaluation sample corresponding to the associated statistic is reported on the horizontal axis.

Figure 5. Rossi and Sekhposyan (2016) Rationality Test.

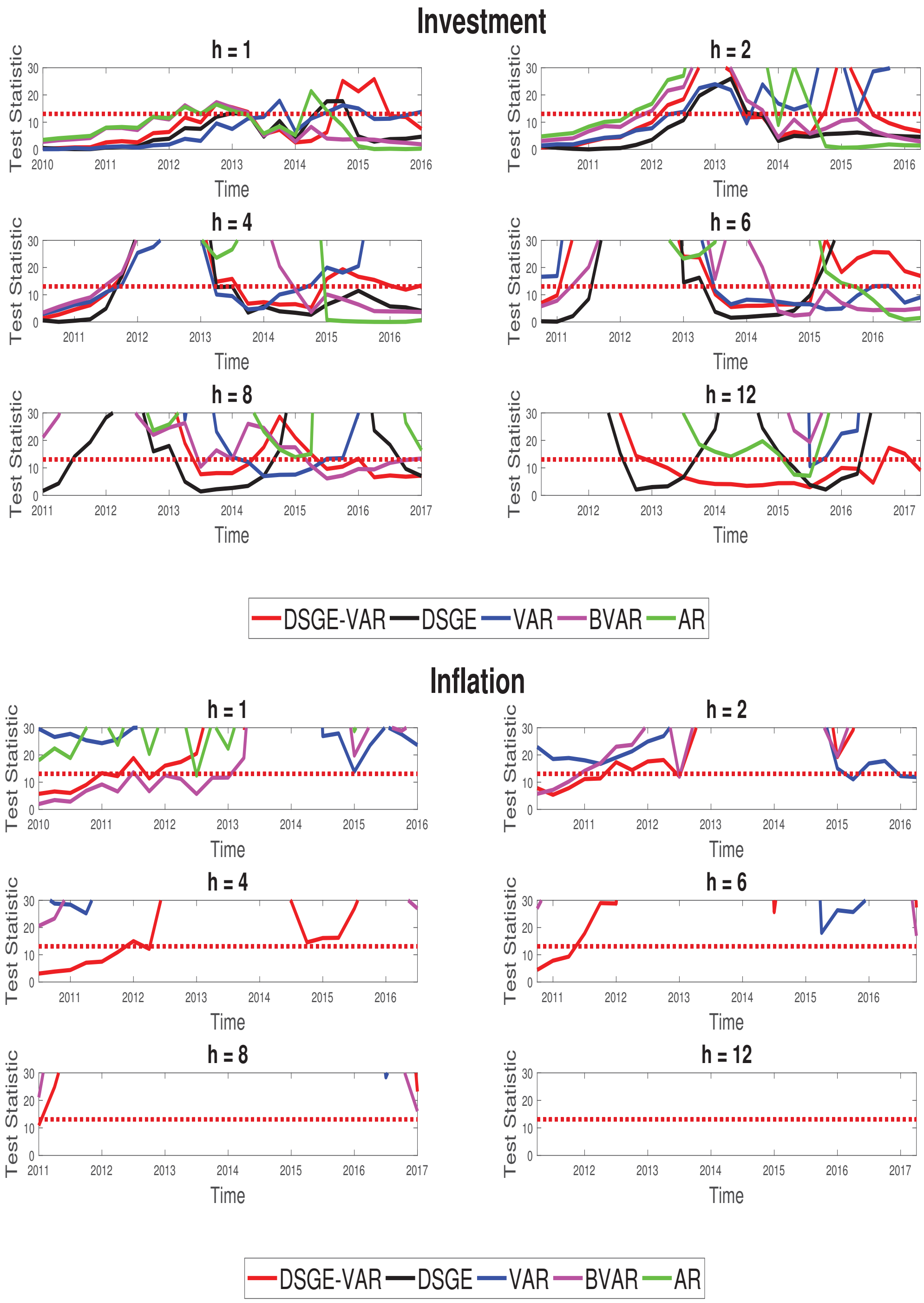


Note: Absolute local rationality forecasting performance of considered models (DSGE–VAR, DSGE, VAR, BVAR, AR) for GDP (ΔY_t) and consumption (ΔC_t) growth rates along different forecast horizons. When the statistics is above the critical value, the null hypothesis of forecast rationality is rejected. The statistics are reported for several steps ahead (h). Red dotted lines indicate the critical values at the 5% of confidence level.

For investment (Figure 6 upper panel) instead, the picture is more mixed. At steps ahead h=1,2 the local rationality assumption is rejected between the first half of 2012 and the second of half 2013 and toward the end of the time span for all the models (with the exception of the BVAR and AR in the latter case). The regions of the time span where the null is rejected gradually spread as the forecast horizons increase and the DSGE–VAR forecasts return to be rational in the final part of the evaluation sample. Finally, the results for inflation (Figure 6 lower panel) confirm what was found in Table 2, namely that its forecasts are hardly locally rational, especially as steps ahead increase. At shorter forecast horizons, when an inflation forecast is locally rational, it comes either from the DSGE–VAR or VAR models, or both.

A general consideration is that the estimated models' forecasts for real variables are overall both stable and efficient. Signs of local non-rationality arise only coincident with the occurrence of financial disruption episodes and more consistently at longer forecast horizons. Moreover, the Rossi and Sekhposyan (2016) rationality fluctuation test reveals that for inflation forecasts, the null hypothesis is seldom non-rejected, particularly for DSGE–VAR and BVAR short run forecasts. Overall, these two models display the most accurate forecasting performance along the several dimensions explored.

Figure 6. Rossi and Sekhposyan (2016) Rationality Test.



Note: Absolute local rationality forecasting performance of considered models (DSGE–VAR, DSGE, VAR, BVAR, AR) for GDP (ΔY_t) and consumption (ΔC_t) growth rates along different forecast horizons. When the statistics is above the critical value, the null hypothesis of forecast rationality is rejected. The statistics are reported for several steps ahead (h). Red dotted lines indicate the critical values at the 5% of confidence level.

5.3 Empirical results based on test of relative forecasting performance

5.3.1 Diebold and Mariano's (1995) Tests

In general, looking at the RMSFEs gives an overall view of the models' forecasting accuracy. However, such a cross-model comparison exercise is limited as it is silent on whether those differences in the forecasting scores are statistically significant. To address such a shortcoming, we make use of the DM (Diebold and Mariano, 1995) test. Given its overall satisfying forecasting performance (see in Figure 4), we choose the DSGE–VAR model as benchmark for the relative (pairwise) forecast evaluation exercise. The benchmark forecasts are compared with those of the DSGE, VAR, BVAR and AR models, respectively.[25] Table 3 reports the p-value of the DM test statistics for comparing the specified forecasts over different forecasting horizons. When the test statistic is below (above) 0, the DSGE–VAR forecasts relatively better (worse). For GDP forecasts, the predictive accuracy is statistically different only for step ahead h=1, when the benchmark model forecasts beat those of the structural model (DSGE), but are in turn overcome by the BVAR and AR. For consumption, no model is worse (better) than the DSGE–VAR in a statistical sense until step ahead h = 6, where the VAR is again outperformed, then AR is outrun at h=12. The evidence is mixed for investment. The benchmark forecasts marginally defeat most of model alternatives except for the (B)VAR models, of which only VAR is then outrun at longer forecast horizons (h= 8,12, respectively). By contrast, the DSGE model forecasts outperform those of the benchmark at h=1,6,8,12. Finally, coming to inflation forecasts, the DSGE–VAR model beats all the alternatives by a large extent, especially at longer forecast horizons, except for the VAR model, against which it is never statistically preferred.

5.3.2 Giacomini and Rossi's (2010) Fluctuation tests

Although considered a useful evaluation instrument, the DM test faces some caveats. It relies on the models' average forecasting performance over the whole (out–of–sample) period. In this regard, it neglects the potential time–varying nature of each models' predictive accuracy. In other words, how the relative forecasting performance may change over time. To address the potential presence of parameter instabilities in predictive regressions, we put our model's forecasts under the lens of the Giacomini and Rossi (2010) fluctuation test.[26] In a nutshell, this test boils down to a sequence of estimated (HAC robust) DM statistics evolving over time following either a recursive or a rolling window scheme. For the sake of clarity, let $\Delta L_{t,h}$ be the difference between the squared forecast

[25] For sake of tractability we exclude A(B)VARs from the pool of selected models as their forecasting performance is in line with the (B)VARs.

[26] For the time evolution of DSGE and DSGE–VAR estimated parameters see Figures 14 and 15 in Section A.3 (Appendix).

Table 3 – (HAC Robust) Diebold and Mariano (1995) Test Statistics

	Steps Ahead (h)					
	1	2	4	6	8	12
			$GDP(\Delta Y_t)$			
DSGE	-2.83***	-0.87	-0.01	0.02	0.02	0.15
VAR	0.56	0.35	-0.23	-0.68	-1.53	-1.46
BVAR	3.43***	1.32	-0.36	-1.02	-1.22	-0.15
AR	2.65***	0.33	-0.90	-1.28	-1.43	-1.02
			$CONSUMPTION(\Delta C_t)$			
DSGE	-0.92	-0.81	-0.43	-1.25	-1.22	-1.60
VAR	-0.48	-0.02	0.03	-2.06**	-1.45	-1.11
BVAR	0.76	0.31	0.27	-1.50	-0.42	-0.74
AR	-0.64	-0.62	-0.29	-1.31	-1.22	-1.70*
			$INVESTMENT(\Delta I_t)$			
DSGE	-1.93*	0.31	0.97	1.65*	2.17**	1.94*
VAR	0.94	0.19	-0.28	-1.52	-2.43**	-1.68*
BVAR	-1.29	-0.32	-1.08	-1.56	-1.21	-0.48
AR	-2.00**	-1.59	-1.39	-1.43	-1.20	-0.80
			$INFLATION(\pi_t)$			
DSGE	-2.71***	-4.79***	-5.82***	-5.17***	-4.76***	-5.18***
VAR	-1.13	-0.58	-0.73	-1.22	-1.28	-1.22
BVAR	0.58	-2.49**	-2.00**	-2.08**	-1.96**	-0.91
AR	-0.90	-2.38**	-5.04***	-3.96***	-3.48***	-2.87***

Note: Pairwise relative forecasting performance of DSGE–VAR against a pool of selected models (DSGE, VAR, BVAR, AR) for GDP (ΔY_t), consumption (ΔC_t), investment (ΔI_t) growth rates and inflation (π_t) along different forecast horizons. When the statistics is below (above) 0, the DSGE–VAR forecasts relatively better (worse). Significance levels are: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

error of two competing models – in our case the Giacomini and Rossi (2010) test is implemented against the DSGE-VAR benchmark. Thus, a negative value of $\Delta L_{t,h}$ indicates that the model under question shows a worse predictive accuracy than the DSGE-VAR benchmark. More precisely, let the local relative loss for the two models be the sequence of out-of-sample loss differences computed over windows of size m (i.e., $m = 20$ quarters):

$$\frac{1}{m}\sum_{j=t-m+1}^{t} \Delta L_{j,h}, \quad t = m, m+1, ..., P. \tag{25}$$

The null hypothesis of equal predictive accuracy at each point in time is:

$$H_0 : E\left[\Delta L_{j,h}\right] = 0, \quad \forall t \tag{26}$$

and the alternative is $E[\Delta L_{j,h}] \neq 0$. The fluctuation test statistic is the highest value over the sequence of relative forecast error losses defined in (25): $\max_t \mathcal{F}^{OOS}_{t,m}$, where

$$\mathcal{F}^{OOS}_{t,m} = \frac{1}{\hat{\sigma}\sqrt{m}} \sum_{j=t-m+1}^{t} \Delta L_{j,h}, \quad t = m, m+1, ..., P.$$

where $\hat{\sigma}^2$ is a heteroskedasticity and autocorrelation consistent (HAC) estimator of the long-run variance of the loss differences.[27]

This time-varying investigation is particularly valuable in the context of several episodes of financial tensions embedded in the evaluation sample. It allows to track the evolution of the relative forecasting performance over time. In this regard, Edge et al. (2010) and Edge and Gürkaynak (2010) found that their benchmark estimated medium-scale DSGE model forecasts inflation and GDP growth very poorly, although statistical and judgmental forecasts do equally poorly. As a result, focusing on forecasting ability only during the Great Moderation is not a good metric by which to judge models.

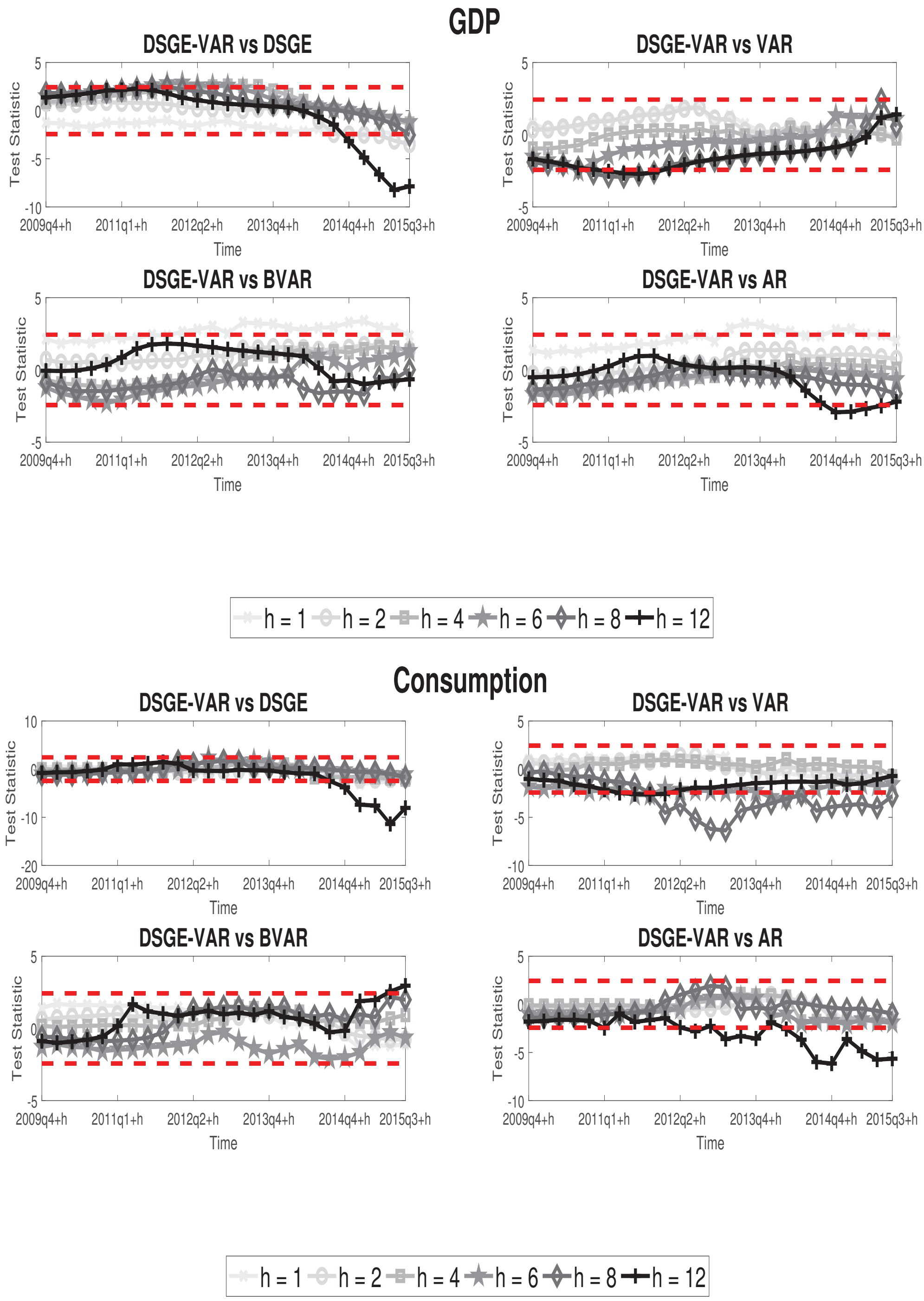
As mentioned before, we select the DSGE–VAR model estimated for Spain as the benchmark for the relative (pairwise) forecast stability evaluation. The predictive loss function of the benchmark model is compared against the DSGE, VAR, BVAR and AR, respectively.

Figures 7 and 9 below report Fluctuation test statistics for GDP, consumption and investment growth rates and inflation, respectively, for $h = [1, 2, 4, 6, 8, 12]$.[28] For each variable of interest we summarize four sets of statistics, each of them comparing the benchmark (i.e. DSGE–VAR model) against: (i) DSGE model (top–left); (ii), VAR(p) model (top–right); (iii), Bayesian VAR(p) model (bottom–left); and (iv) AR(1) model (bottom–right). Whenever the statistic falls below (above) the critical value, the relative forecasting performance of the benchmark model are statistically better (worse) than that of the alternative model at that given point in time.

[27] The only strict requirement is the use of a rolling or fixed estimation scheme in the production of the out–of–sample forecasts.

[28] Increasing steps ahead are associated to progressively darker shades of gray and dashed red lines indicate the critical values at the 5% confidence level.

Figure 7. Giacomini and Rossi (2010) Fluctuation Test.



Note: Pairwise local relative forecasting performance of DSGE–VAR against a pool of selected models (DSGE, VAR, BVAR, AR) for GDP (ΔY_t) and consumption (ΔC_t) growth rates along different forecast horizons. When the statistic is below (above) the lower (upper) critical value, the DSGE–VAR forecasts significantly better (worse) at that point in time. Higher steps ahead (h) are associated with increasingly darker lines. Red dashed lines indicate the critical values at the 5% of confidence level.

5.3.3 Empirical results on forecasting GDP growth

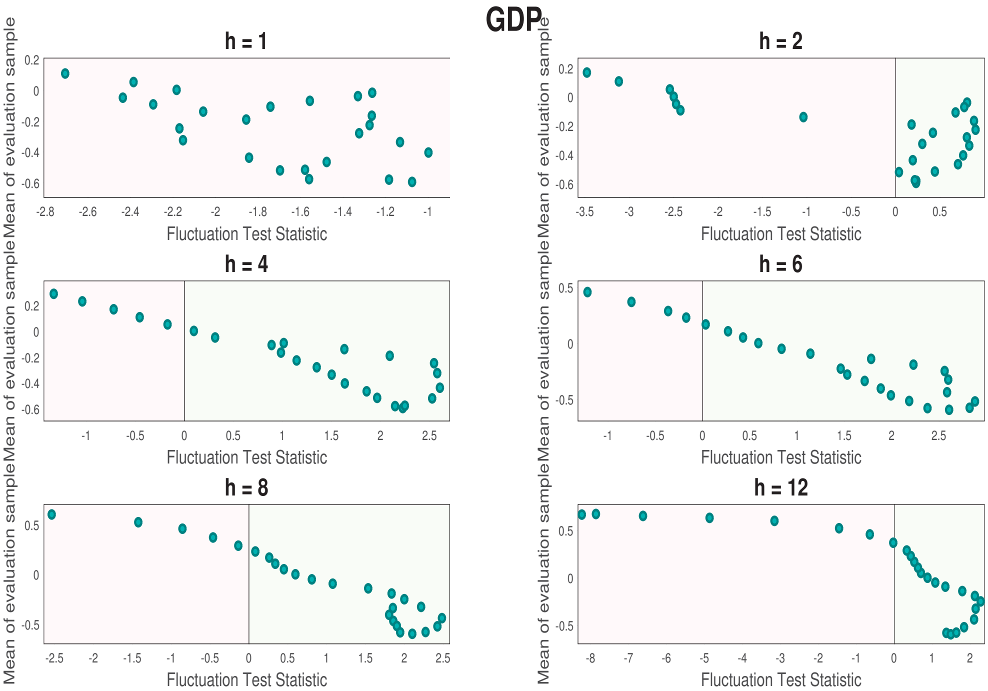
The upper panel of Figure 7 shows that, when considering output growth, the DSGE–VAR relative forecasting performance does change over time with respect to the DSGE.[29] First, we can observe how the DSGE–VAR constantly outperforms at the one-quarter-ahead horizon the DSGE model's forecasts, even significantly as the financial crises weight in the evaluation samples increases. However, for longer horizons, the fluctuation test statistics point to less stable relative forecasting performance following an inverse U-pattern. Therefore, for steps ahead h=4, 6 and 8 the DSGE model turns out to deliver more accurate forecasts around the centered time span under consideration. Then at step ahead h=8 the DSGE–VAR again outperforms the DSGE in the last evaluation sample, and for step ahead $h = 12$ it is statistically better from 2014Q4+h. The picture looks quite different from the static relative forecasting performance summarized in Table 3, as the forecast stability assumption is clearly violated.

This result needs to be rationalised. First, the medium-run forecasting properties of the DSGE model seem to be uniformly better as the time span embeds the financial crises, and just the opposite holds true for the one-step-ahead horizon. Second, when forecasting output growth, the DSGE model's forecasts have proven to perform better than those of non-structural models at longer horizons but slightly worse at shorter horizons (Gürkaynak et al., 2013). Moreover, it has also been evidenced that DSGE forecasting properties struggle to deal with financial crises periods, especially when financial frictions are not considered (Cai et al., 2018). The solution to this puzzle evidenced by our empirical results points to the fact that the DSGE model tends to underestimate the deterministic growth rate of real variables compared to the DSGE–VAR (i.e. γ, see Table 1A and also Figure 14 in section A.3 in the Appendix), and this translates into better medium-long run forecasts when the evaluation samples embed the financial crisis episodes.

To confirm this hypothesis, Figure 8 shows the scatter plot of the Giacomini and Rossi (2010) test statistic against the mean of the associated evaluation sample. Lower evaluation sample means proxy higher weights of financial crisis episodes in the evaluation sample. Against this background, where the test statistic is positive (green shaded areas), the DSGE model's predictive ability is relatively more accurate than that of the benchmark (DSGE–VAR model); in contrast, in those cases where the test statistic moves into negative territory (red shaded areas), the DSGE-VAR model's forecasts are relatively more accurate. These results suggest that, the DSGE model's forecasts become relatively more accurate as the mean of the associated evaluation sample turns negative, from step ahead h=4 onward. This is a pattern completely absent for step ahead $h = 1$.

[29]This finding relates to Stock and Watson (2003) proving the existence of widespread instabilities in the parameters of models describing output growth and inflation in the U.S and to Rossi and Sekhposyan (2010) who show that most predictors for output growth lost their predictive ability in the mid-1970s.

FIGURE 8. Fluctuation test statistic (DSGE–VAR vs DSGE) against the mean of the associated evaluation sample, GDP growth.



Note: Scatter plot of the DSGE–VAR vs DSGE Giacomini and Rossi (2010) fluctuation tests statistic (horizontal axis) against the mean of the associated evaluation sample (vertical axis) along several steps ahead: $h = [1, 2, 4, 6, 8, 12]$ for GDP growth (ΔY_t). Green shaded areas identify regions of the fluctuation test statistic where the DSGE (predictive accuracy) prevails; red shaded areas identify regions of the test statistic where the DSGE–VAR (predictive accuracy) prevails.

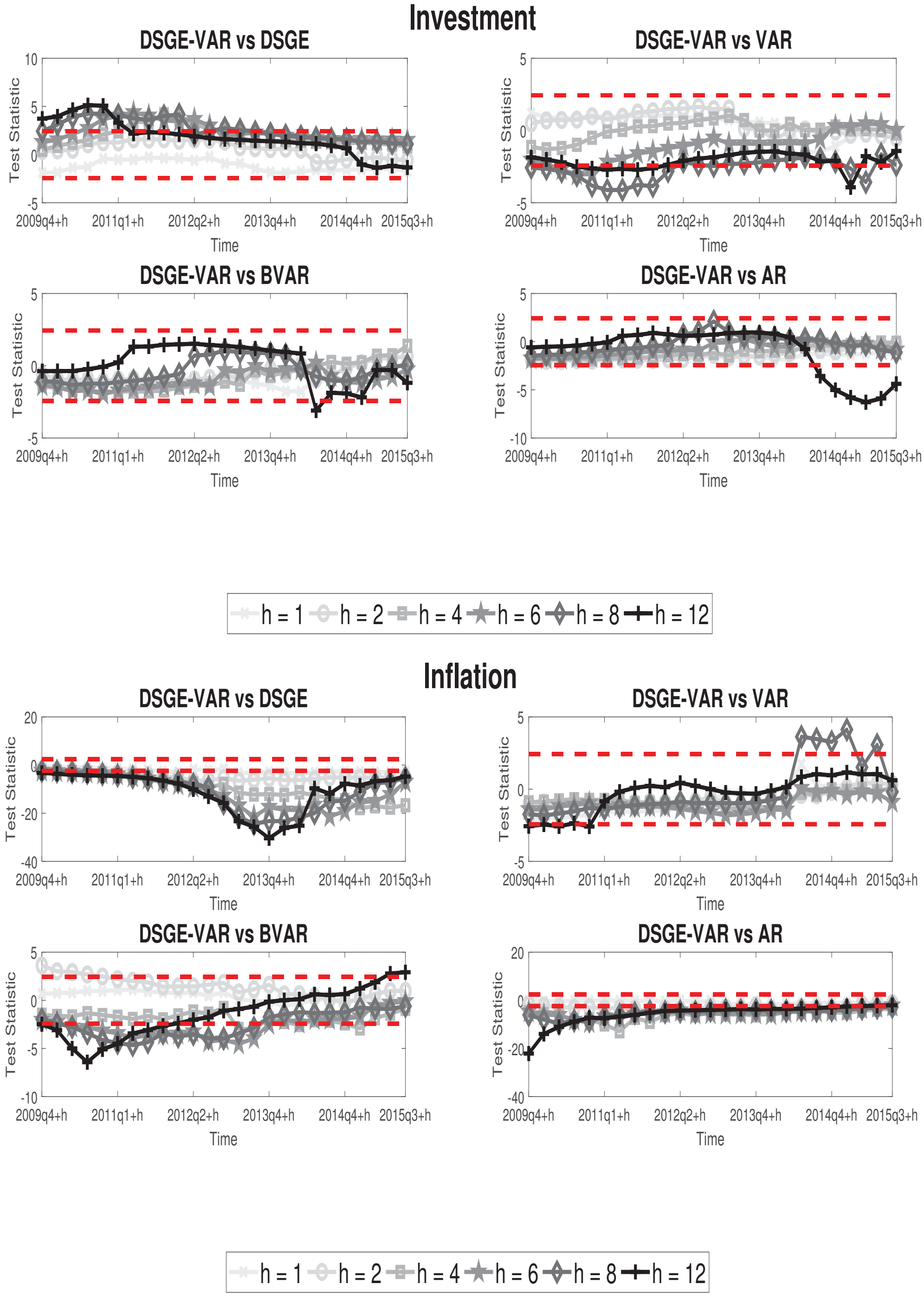
Finally, when confronting the benchmark with the BVAR and AR models, our results point to the fact that both models' forecasts are significantly better at the shortest forecast horizon. By contrast, the benchmark and VAR model's forecasts are relatively stable and almost never statistically different.

5.3.4 Empirical results on forecasting consumption growth

The lower panel of Figure 7 displays results for consumption growth. As comparing the benchmark with DSGE (top–left corner), it emerges that as long as the steps ahead increase, the DSGE–VAR becomes progressively better, up to becoming significant in the last data points for h=1,2 and 12. The top–right corner shows that the benchmark (DSGE–VAR) outruns the VAR as the steps ahead increase, in particular for h = 6,8. For the comparison of the DSGE–VAR with the Bayesian-VAR (bottom–left corner), it is possible to draw conclusions similar to GDP. The only exception is that now statistically significant differences arise only in the last part of the evaluation sample for h=12.

Finally, with regard to the AR model, the picture resembles the DSGE–VAR vs DSGE comparison. This is not surprising since the RMSFE dynamics of both DSGE and AR are quite similar along increasing steps ahead, as shown in Figure 4.

Figure 9. Giacomini and Rossi (2010) Fluctuation Test.



Note: Pairwise local relative forecasting performance of DSGE–VAR against a pool of selected models (DSGE, VAR, BVAR, AR) for Investment (ΔI_t) growth rates and inflation (π_t) along different forecast horizons. When the statistics is below (above) the lower (upper) critical value, the DSGE–VAR forecasts significantly better (worse) at that point in time. Higher steps ahead (h) are associated with increasingly darker lines. Red dashed lines indicate the critical values at the 5% of confidence level.

5.3.5 Empirical results on forecasting investment growth

The upper panel of Figure 9 reports the fluctuation test results for investment growth. From the comparison of the benchmark with the DSGE model (top–left corner) it emerges that the relative forecasting accuracy of the two models is overall in line with that seen for GDP. The DSGE model tends to beat the benchmark model from $h = 4$ onward as the weight of financial crises in the evaluation samples increases. Moving to the VAR (top–right corner), the DSGE–VAR is almost always strictly preferred along the whole time span for step ahead $h = 8$. Finally, for the BVAR (bottom–left) and the AR (bottom–rights) models, the forecasts are relatively stable and none of the alternatives is statistically better, except for the last step ahead where the DSGE–VAR beats both the alternatives at least once.

5.3.6 Empirical results on forecasting inflation

The lower panel of Figure 9 reports results for inflation. Typically, simple AR models are mostly accurate at long horizons and DSGE models are mostly accurate at short horizons when forecasting inflation. However, the DSGE–VAR benchmark performance is outstanding. As usual, the top–left corner reports results for the pairwise comparison with the DSGE. Looking at the statistics dynamics for all the steps ahead, it becomes obvious how the DSGE–VAR improves its relative forecasting accuracy as the time span approaches the end of each evaluation sample. Moreover, the benchmark forecasts are statistically more accurate at least from 2013Q1 onward, and for $h > 2$, they are always statistically better. From Table 1A we can see that the deterministic mean of the inflation rate, $\bar{\pi}$, is closer to the actual sample mean for the DSGE than it is for the DSGE–VAR. In other words, the benchmark tends to underestimate the inflation sample mean. However, a quick look at the inflation dynamics (see Figure 10 in Section A.1 in Appendix) reveals that the inflation mean is far higher in the first half of the sample than in the second, and this fact favors the benchmark inflation forecasting accuracy (to confirm this view look at the evolution of $\bar{\pi}$ marginal posterior along all the evaluation samples as depicted in Figure 14 in section A.3 in Appendix).

Considering the VAR (top–right corner), for $h = 12$ the benchmark forecasts' accuracy initially dominates to then quickly revert. Then, apart from h=8 where the VAR outruns the benchmark (DSGE–VAR) at some points in the second half of the time span, none of the statistics is significant.

For the BVAR (bottom–left corner), and considering the first step ahead, the picture is essentially in line with the DSGE case. Then, abstracting from step ahead $h = 2$ (where the BVAR is better until the end of 2010), the benchmark is always significantly better at some point in time, especially in the first half of the horizontal axis. Finally, the AR (bottom–right corner) is constantly dominated by the DSGE–VAR, except for $h = 1$ where the benchmark prevails from the second half of 2013.

6 Concluding Remarks

In this paper, we exploit several combinations of the Smets and Wouters (2003) structural model with standard reduced–form methods such as (B)VAR (i.e. DSGE–VAR and Augmented–(B)VAR–DSGE methods) and place them in a forecasting competition for the Spanish economy. We show that the forecasts for real variables produced by combinations of the structural and non–structural models, especially the DSGE–VAR and the BVAR models, are generally optimal and rational (i.e. unbiased and efficient) in absolute terms. We find that, in the context of a pseudo real-time out-of-sample forecasting exercise, our large scale New Keynesian DSGE model underestimates the growth of real variables due to its mean reverting properties. However, given the DSGE–VAR model predictive accuracy, we demonstrate that reduced-form VARs largely benefit from the imposition of economic priors from such structural models. As for the Augmented-(B)VAR, we show that pooling fragmented information extracted from more complex underlying structural models does not give rise to any particularly relevant gain in terms of forecasting accuracy. In particular, the benefit of incorporating the main business cycle drivers (also conditional on the imposition of a non-economic prior) as additional observables in reduced–form models is quite limited. Finally, looking across the sample, the relaxation of the DSGE prior is particularly effective during specific episodes, coherent with the dynamics of the observables' first moments.

Besides the fact that results may be data-driven, one avenue for future work could include testing for improvements by conditioning the model forecasts on either Consensus Forecasts or nowcasting estimates. Another would be to estimate the DSGE model in its non-linear form, as there is considerable evidence that taking into account the inherent non-linearities both in the solution (nonlinear approximation of policy rules as opposed to log-linearization) and the estimation (particle filter as opposed to the Kalman filter) considerably increases the in-sample fit (Fernández-Villaverde and Rubio-Ramírez, 2005).

References

[1] Adolfson, Malin, Stefan Laséen, Jesper Lindé, and Mattias Villani (2007). "Bayesian estimation of an open economy DSGE model with incomplete pass-through". In: *Journal of International Economics* 72.2, pp. 481–511.

[2] Andrés, Javier, Pablo Burriel, and Ángel Estrada (2006). "BEMOD: A DSGE Model for the Spanish Economy and the Rest of the Euro Area". In: *Banco de Espana Research Paper No. WP-0631*.

[3] Andrés, Javier, Samuel Hurtado, Eva Ortega, and Carlos Thomas (2010). "Spain in the euro: a general equilibrium analysis". In: *SERIEs* 1.1-2, pp. 67–95.

[4] Andrés, Javier, Óscar Arce, and Carlos Thomas (2017). "Structural reforms in a debt overhang". In: *Journal of Monetary Economics* 88, pp. 15–34.

[5] Boscá, JE, A Díaz, R Doménech, E Pérez, and L Puch (2007). "The REMSDB macroeconomic database of the Spanish economy". In: *Documentos de Trabajo del Ministerio de Economía y Hacienda, No. WP-2007-04*.

[6] Boscá, JE, A Díaz, Rafael Domenech, Javier Ferri, E Pérez, and Luis Puch (2010). "A rational expectations model for simulation and policy evaluation of the Spanish economy". In: *SERIEs* 1.1-2, pp. 135–169.

[7] Boscá, JE, A Díaz, Rafael Domenech, Javier Ferri, and J F Rubio-Ramírez (2018). "Financial and Fiscal Shocks in the Great Recession and Recovery of the Spanish Economy". In: *Working Papers 2018-05, FEDEA*.

[8] Burriel, Pablo, Jesús Fernández-Villaverde, and Juan F Rubio-Ramírez (2010). "MEDEA: a DSGE model for the Spanish economy". In: *SERIEs* 1.1-2, pp. 175–243.

[9] Cai, Michael, Marco Del Negro, Marc P Giannoni, Abhi Gupta, Pearl Li, and Erica Moszkowski (2018). "DSGE forecasts of the lost recovery". In: *Federal Reserve Bank of New York Staff Reports, no. 844*.

[10] Christiano, Lawrence J, Martin Eichenbaum, and Charles L Evans (2005). "Nominal rigidities and the dynamic effects of a shock to monetary policy". In: *Journal of Political Economy* 113.1, pp. 1–45.

[11] Del Negro, Marco and Frank Schorfheide (2004). "Priors from general equilibrium models for VARs". In: *International Economic Review* 45.2, pp. 643–673.

[12] — (2006). "How good is what you've got? DGSE-VAR as a toolkit for evaluating DSGE models". In: *Economic Review-Federal Reserve Bank of Atlanta* 91.2, p. 21.

[13] Del Negro, Marco, Frank Schorfheide, Frank Smets, and Rafael Wouters (2007). "On the fit of new Keynesian models". In: *Journal of Business & Economic Statistics* 25.2, pp. 123–143.

[14] Diebold, Francis X (1998). "The past, present, and future of macroeconomic forecasting". In: *Journal of Economic Perspectives* 12.2, pp. 175–192.

[15] Diebold, Francis X and Robert S Mariano (1995). "Comparing predictive accuracy". In: *Journal of Business & Economic Statistics* 13.3, pp. 253–263.

[16] Edge, Rochelle M and Refet S Gürkaynak (2010). *How useful are estimated DSGE model forecasts for central bankers?* Centre for Economic Policy Research.

[17] Edge, Rochelle M, Michael T Kiley, and Jean-Philippe Laforte (2010). "A comparison of forecast performance between federal reserve staff forecasts, simple reduced-form models, and a DSGE model". In: *Journal of Applied Econometrics* 25.4, pp. 720–754.

[18] Erceg, Christopher J, Luca Guerrieri, and Christopher J Gust (2006). "SIGMA: a new open economy model for policy analysis". In: *International Journal of Central banking*, 2(1):1–50.

[19] Fernández-de-Córdoba, Gonzalo and José L Torres (2011). "Forecasting the Spanish economy with an augmented VAR–DSGE model". In: *SERIEs* 2.3, pp. 379–399.

[20] Fernández-Villaverde, Jesús and Juan F Rubio-Ramírez (2005). "Estimating dynamic equilibrium economies: linear versus nonlinear likelihood". In: *Journal of Applied Econometrics* 20.7, pp. 891–910.

[21] Giacomini, Raffaella and Barbara Rossi (2010). "Forecast comparisons in unstable environments". In: *Journal of Applied Econometrics* 25.4, pp. 595–620.

[22] Giacomini, Raffaella and Halbert White (2006). "Tests of conditional predictive ability". In: *Econometrica* 74.6, pp. 1545–1578.

[23] Goméz-González, Patricia and Daniel M Rees (2018). "Same Spain, less pain?" In: *European Economic Review* 110(C), pp. 78–2107.

[24] Gürkaynak, Refet S, Burçin Kısacıkoğlu, and Barbara Rossi (2013). "Do DSGE models forecast more accurately out-of-sample than VAR models?" In: *CEPR Discussion Paper* No. DP9576.

[25] Harrison, Richard, Kalin Nikolov, Meghan Quinn, Gareth Ramsay, Alasdair Scott, and Ryland Thomas (2005). "The Bank of England quarterly model". In:

[26] Kilponen, Juha and Antti Ripatti (2006). "Introduction to AINO". In: *Mimeo, Bank of Finland*.

[27] King, Robert G and Sergio T Rebelo (1999). "Resuscitating real business cycles". In: *Handbook of Macroeconomics* 1, pp. 927–1007.

[28] Kortelainen, Mika (2002). "EDGE: a model of the euro area with applications to monetary policy". In: *Scientific Monographs, Bank of Finland*.

[29] Lees, Kirdan, Troy Matheson, and Christie Smith (2011). "Open economy forecastsing with a DSGE-VAR: Head to head with the RBNZ published forecasts". In: *International Journal of Forecasting* 27.1, pp. 512–528.

[30] Litterman, Robert B (1986a). "A statistical approach to economic forecasting". In: *Journal of Business & Economic Statistics* 4.1, pp. 1–4.

[31] — (1986b). "Forecasting with Bayesian vector autoregressions—five years of experience". In: *Journal of Business & Economic Statistics* 4.1, pp. 25–38.

[32] Mincer, Jacob A and Victor Zarnowitz (1969). "The evaluation of economic forecasts". In: *Economic forecasts and expectations: Analysis of forecasting behavior and performance*. NBER, pp. 3–46.

[33] Murchison, Stephen, Andrew Rennison, et al. (2006). *ToTEM: The Bank of Canada's new quarterly projection model*. Bank of Canada Ottawa, Ontario, Canada.

[34] Rossi, Barbara and Tatevik Sekhposyan (2010). “Have economic models’ forecasting performance for US output growth and inflation changed over time, and when?” In: *International Journal of Forecasting* 26.4, pp. 808–835.

[35] — (2016). “Forecast rationality tests in the presence of instabilities, with applications to Federal Reserve and survey forecasts”. In: *Journal of Applied Econometrics* 31.3, pp. 507–532.

[36] Sims, Christopher A (1980). “Macroeconomics and reality”. In: *Econometrica: Journal of the Econometric Society*, pp. 1–48.

[37] Smets, Frank and Raf Wouters (2003). “An estimated dynamic stochastic general equilibrium model of the euro area”. In: *Journal of the European Economic Association* 1.5, pp. 1123–1175.

[38] Stock, James H and Mark W Watson (2003). “Forecasting output and inflation: The role of asset prices”. In: *Journal of Economic Literature* 41.3, pp. 788–829.

[39] Warne, Anders, Günter Coenen, Kai Christoffel, et al. (2008). “The new area-wide model of the euro area: a micro-founded open-economy model for forecasting and policy analysis”. In: *ECB Working Paper No. 944*.

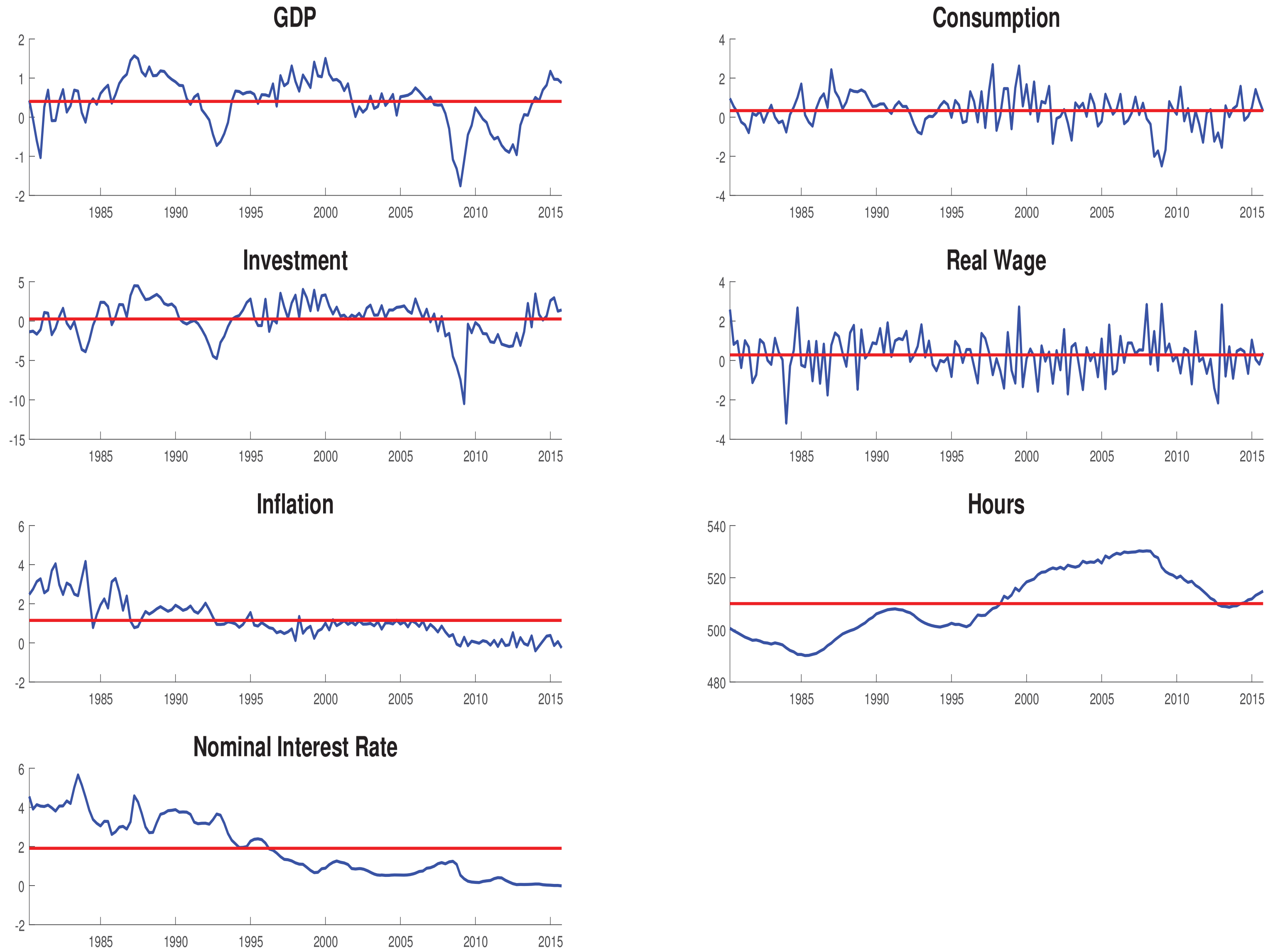
[40] West, Kenneth D and Michael W McCracken (1998). “Regression-based tests of predictive ability”. In: *National Bureau of Economic Research*.

A Appendix

A.1 Stationarity Analysis

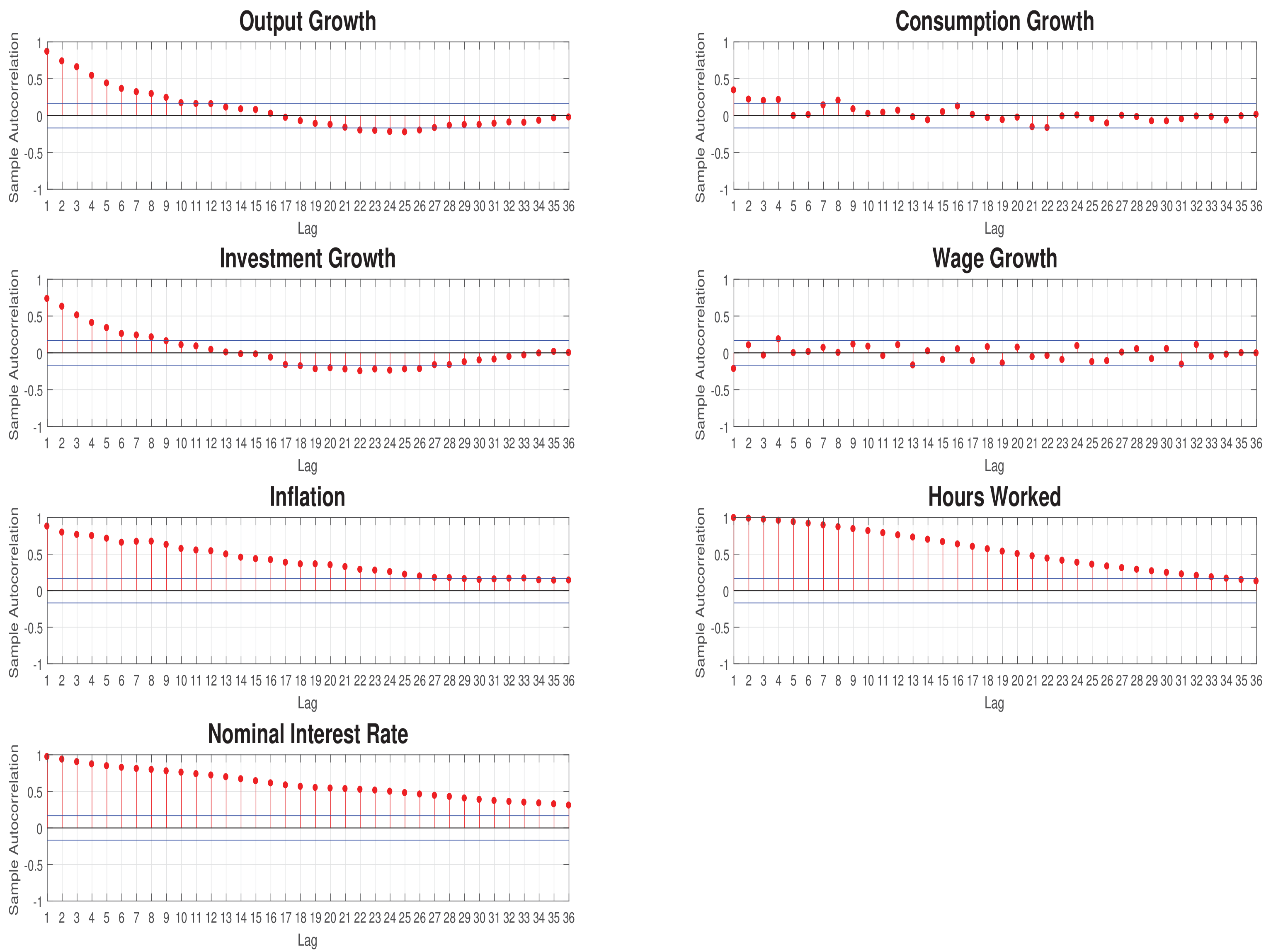
In this section, the stationarity properties of the observables employed for the modelling strategy are analysed. To begin with, observables are plotted in Figure 10. At first sight, since the 2008-09 financial crisis, the GDP and investment series display quite different persistence. Differently, the consumption and wage series show higher levels of volatility over the whole time span. Turning to inflation and the nominal interest rate, there exists sound grounds of instability in the form of at least a couple structural breaks, the first occurring in the early nineties as Spain joins the ERM and the other following the financial crisis. Hours worked dynamics appear to be quite smooth, persistent and clearly follow an oscillatory path suggesting the presence of at least a complex root. In Figure 11, observables correlograms up to 36 lags are displayed. The results suggest that all of the real variables are likely to be stationary, even if investment shows marginally significant autocorrelations between lag 18 and 27. Concerning inflation, nominal interest and hours worked, all of them display clear signs of non-stationarity.

Figure 10. Observables.



Note: Time series displayed are output growth (ΔY_t), consumption (ΔC_t), investment (ΔI_t), real wage (ΔW_t), inflation (π_t), nominal interest rate (R_t) and hours worked (N_t). Data treatment is described is Section 2.3.1.

Figure 11. Observables Autocorrelograms.



Note: Correlograms displayed refer to output growth (ΔY_t), consumption (ΔC_t), investment (ΔI_t), real wage (ΔW_t), inflation (π_t), nominal interest rate (R_t) and hours worked (N_t). Confidence bands are at 5%.

Table 4 – ADF and KPSS tests.

	ADF (Unit Root test)			KPSS (Stationarity Test)		
	Statistic	C.V. 5%	C.V. 10%	Statistic	C.V. 5%	C.V. 10%
ΔY_t	-3.14	-2.89	-2.58	0.28	0.46	0.35
ΔC_t	-8.34	-2.89	-2.58	0.30	0.46	0.35
ΔI_t	-3.57	-2.89	-2.58	0.16	0.46	0.35
ΔW_t	-15.07	-2.89	-2.58	0.21	0.46	0.35
π_t	-2.49	-2.89	-2.58	1.29	0.46	0.35
N_t	-1.81	-2.89	-2.58	0.95	0.46	0.35
R_t	-1.11	-2.89	-2.58	1.33	0.46	0.35

Note: In all unit root and stationarity tests an intercept has been included. Lags selection follows the Schwarz Info Criterion, and the maximum number of lags is set to 11.

Finally, Table 4 summarizes the results for the Augmented Dickey-Fuller (ADF) unit root test (Column 1) and, on a confirmatory ground, for the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) stationarity test (Column 4) and the respective critical values at 5 and 10% (Columns 2:3 and 5:6). For GDP, consumption, investment and wage it is clear that the null hypothesis of a unit root has to be rejected. Further, the KPSS does not reject the null of stationarity. For inflation, the null hypothesis of a unit root cannot be rejected, whilst the KPSS test largely rejects the null of stationarity. In this regard we conclude that inflation is non-stationary. Then, unit root tests for hours worked and nominal interest rate clearly report the presence of a unit root. Finally, some considerations on the implications of the stationarity analysis. First, given that three out of seven series show non-stationarity, it is not surprising that OLS VAR forecasts may face an explosive path, especially for inflation, interest rate amd hours worked.

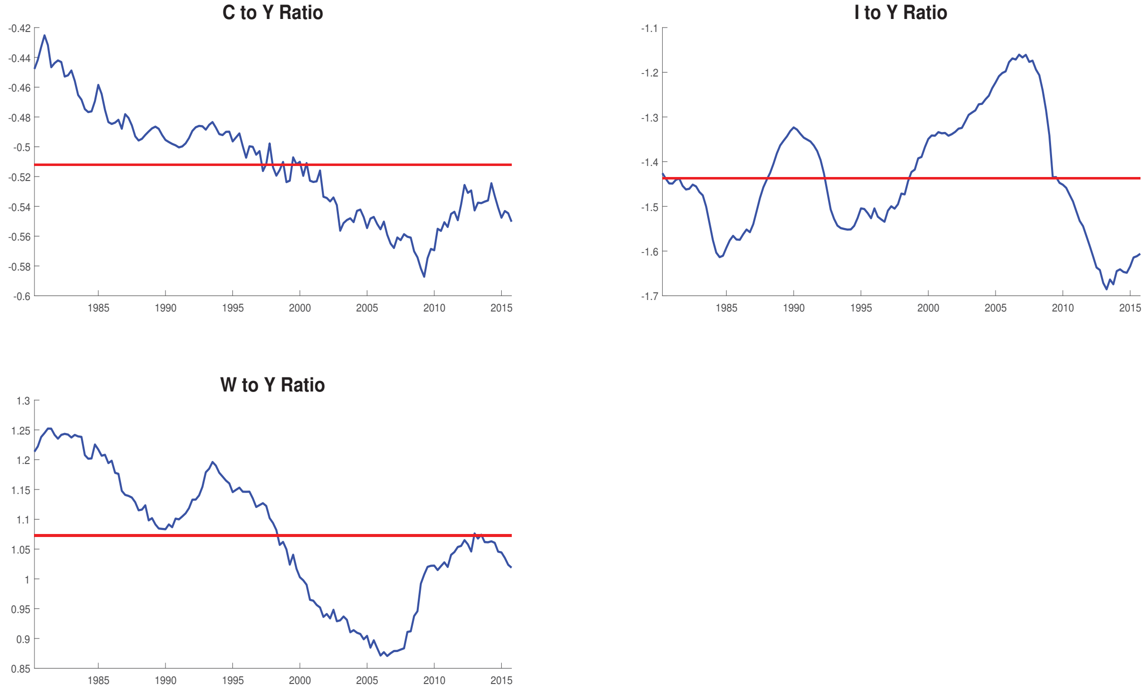
A.2 Investigating the Existence of Cointegrating Relationships

In this section we motivate our modelling choice of abstracting from the existence of a unit root in the technology progress as in the original formulation of SW. Such an assumption nests cointegrating relationships of consumption, investment and real wage with output. However, as already made clear by Del Negro et al. (2007), the data reject these restriction for the US. In this regard we show that this is even more critically so for the case of Spain.

In Figure 12 we show the log-ratio of consumption, investment and real wage to output. It is evident at first glance that both consumption and the wage to output ratio display a clear decreasing trend. This is different to the investment to output ratio whose trend is instead quite flat. All in all the three series show clearly divergent paths. This evidence is clearly at odds with the cointegrating assumption implied by the original formulation of Del Negro et al. (2007).

In Figure 13 the correlograms for the cointegrating ratios are reported. As already hinted, all the series display clear signs of non-stationarity.

Figure 12. Cointegrating Relationships.



Note: Time series displayed are the log of consumption to output ratio, investment to output ratio, real wage to output ratio.

Finally, in Figure 5 we report, as before, results for both unit root and stationarity tests. As expected, for all the cointegrating relationships under analysis clearly emerges the presence of a unit root which clearly rules out the presence of cointegration. The stationarity tests confirm our finding except for the investment to output ratio for which the stationarity hypothesis cannot be rejected [30]. These results are virtually unchanged when a deterministic trend is included in the test specification.

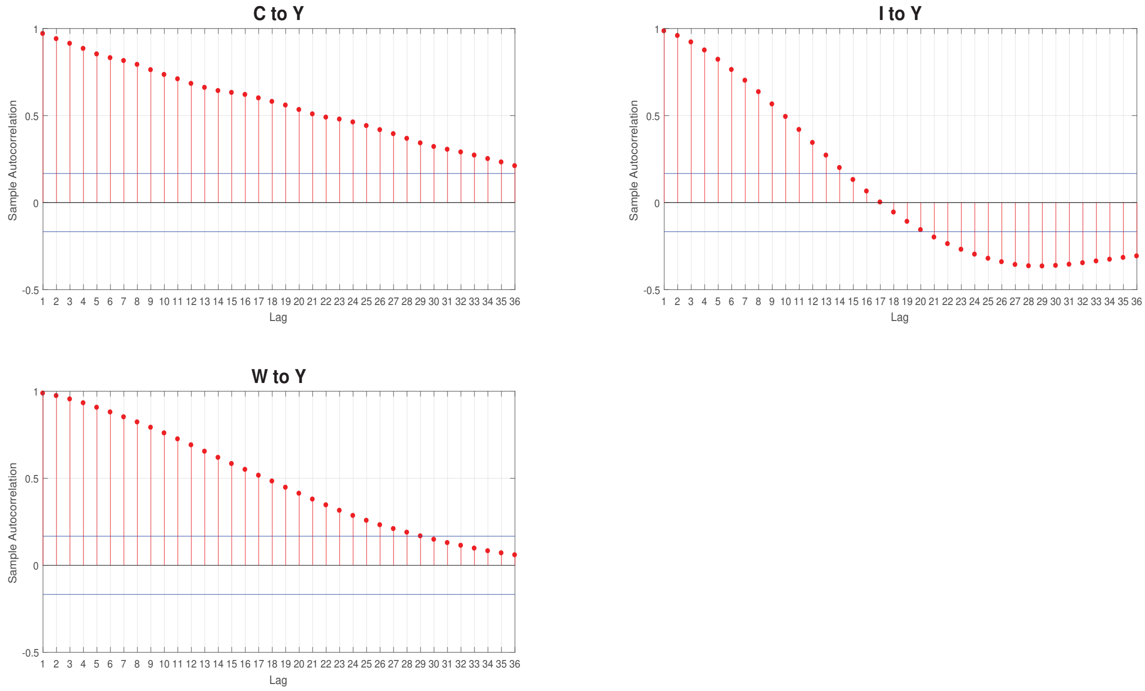
Table 5 – ADF and KPSS tests.

	ADF (Unit Root test)			KPSS (Stationarity Test)		
	Statistic	C.V. 5%	C.V. 10%	Statistic	C.V. 5%	C.V. 10%
C_t/Y_t	-1.35	-2.89	-2.58	1.27	0.46	0.35
I_t/Y_t	-2.00	-2.89	-2.58	0.17	0.46	0.35
W_t/Y_t	-1.76	-2.89	-2.58	0.94	0.46	0.35

Note: In all unit root and stationarity tests an intercept has been included. Lags selection follows the Schwarz Info Criterion, and the maximum number of lags is set to 11.

[30] However, the Phillips-Perron test confirms the presence of a unit root.

Figure 13. Cointegrating Relationships Autocorrelograms.

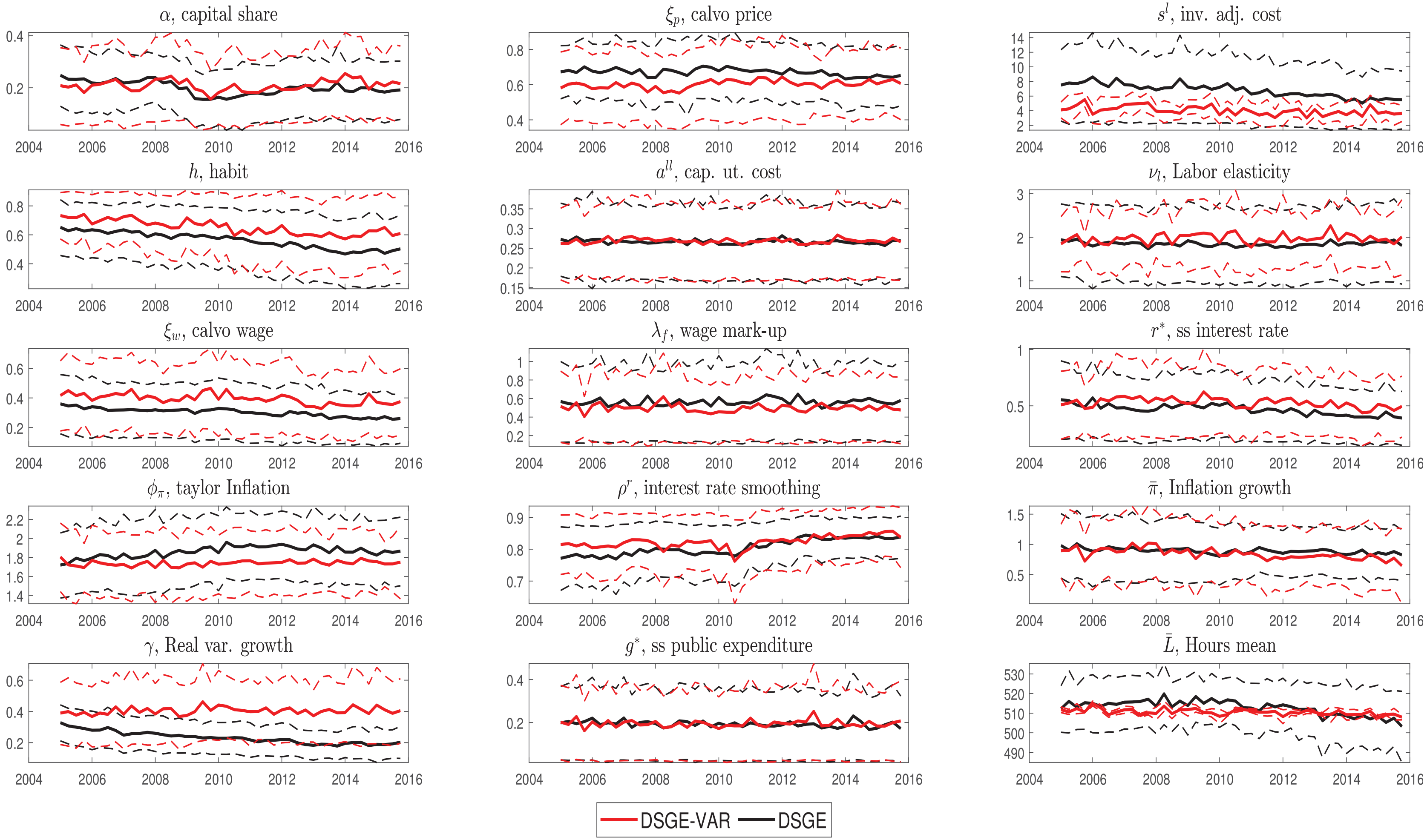


Note: Correlograms displayed are the log of consumption to output ratio, investment to output ratio, real wage to output ratio. Confidence bands are at 5%.

A.3 DSGE and DSGE-VAR parameters stability

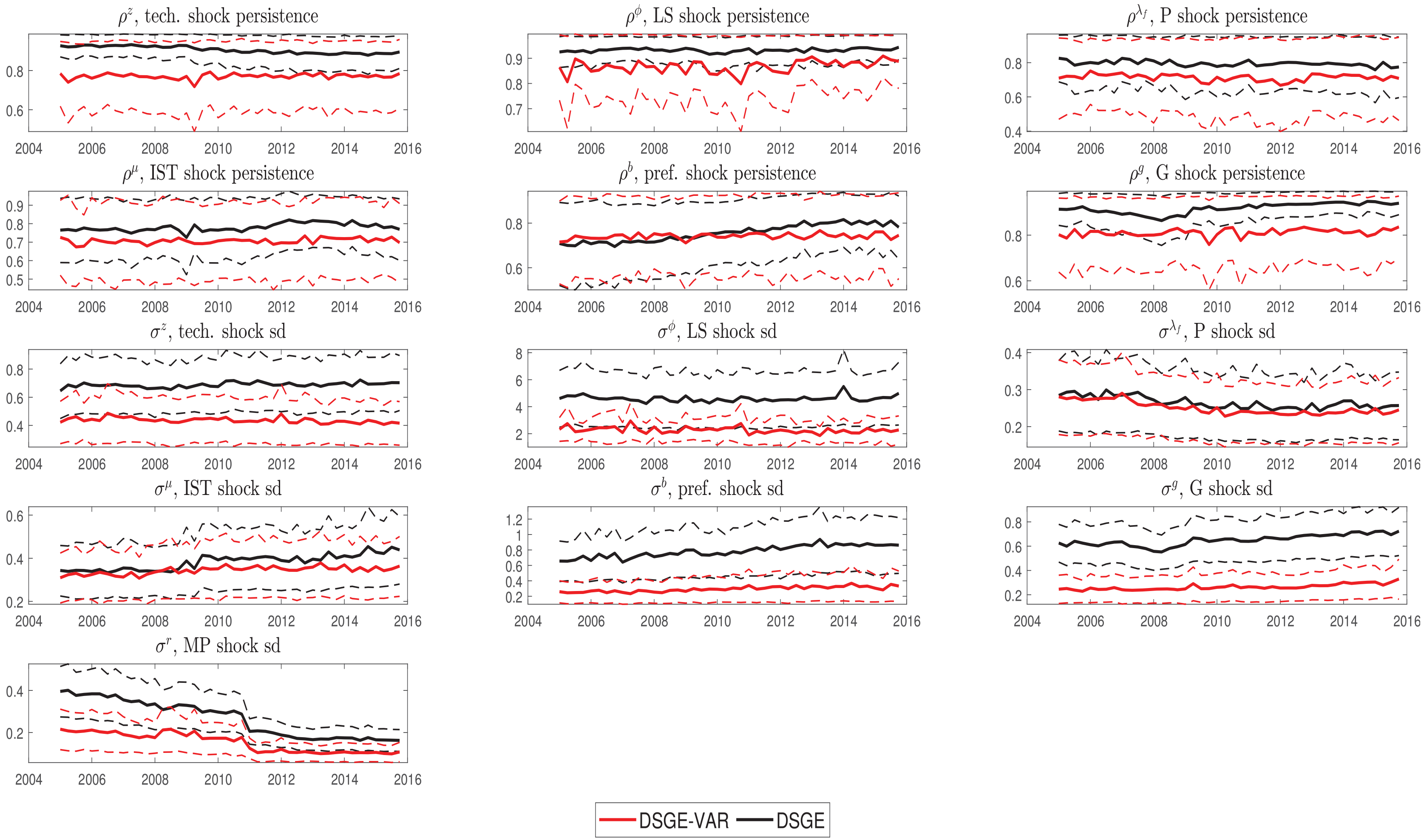
In this section we look at the evolution over time of the estimated parameters considered in the estimation of the structural (DSGE) and hybrid (DSGE–VAR) models. The evolution over time of the estimated posteriors is reported in Figures 14 and 15 below. Continuous black and red lines track the evoulution of the DSGE and DSGE–VAR posterior median, dashed black and red lines track the dynamics of the 5th and 95th posterior percentiles, respectively.

Figure 14. Estimated structural parameters stability.



Note: Estimated posterior densities for all structural parameters along all the evaluation samples. Black: DSGE; Red: DSGE–VAR; continuous lines indicate the median, dashed lines indicate the 5th and 95th percentile, respectively.

FIGURE 15. Estimated exogenous processes parameters stability.



Note: Estimated posterior densities for all exogenous processes parameters along all the evaluation samples. Black: DSGE; Red: DSGE–VAR; continuous lines indicate the median, dashed lines indicate the 5th and 95th percentile, respectively.

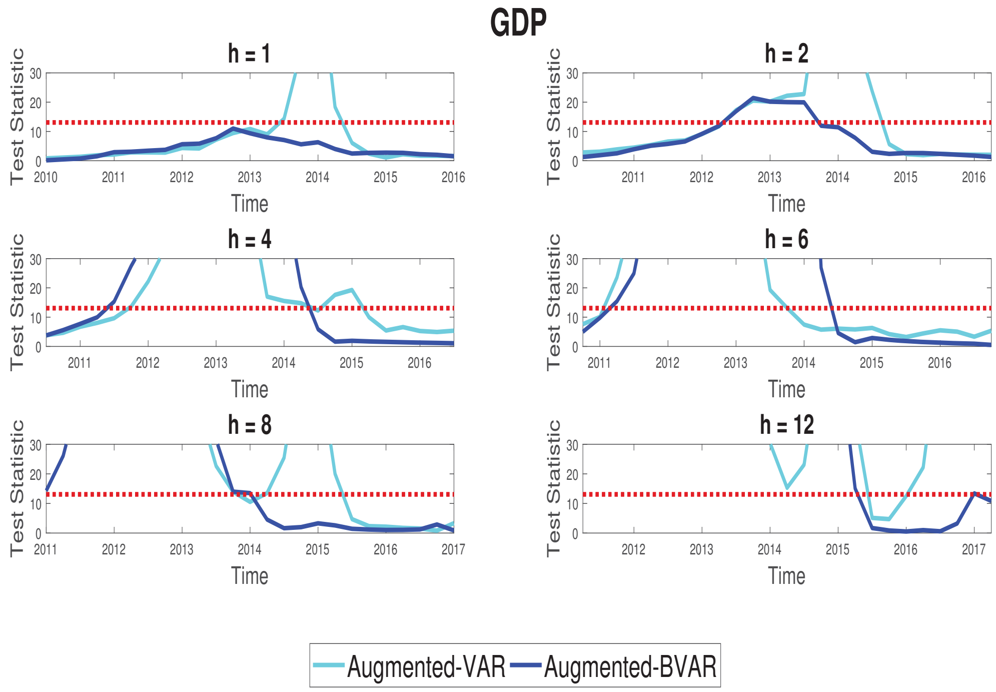
A.4 Augmented–(B)VAR Relative Forecast absolute and relative accuracy

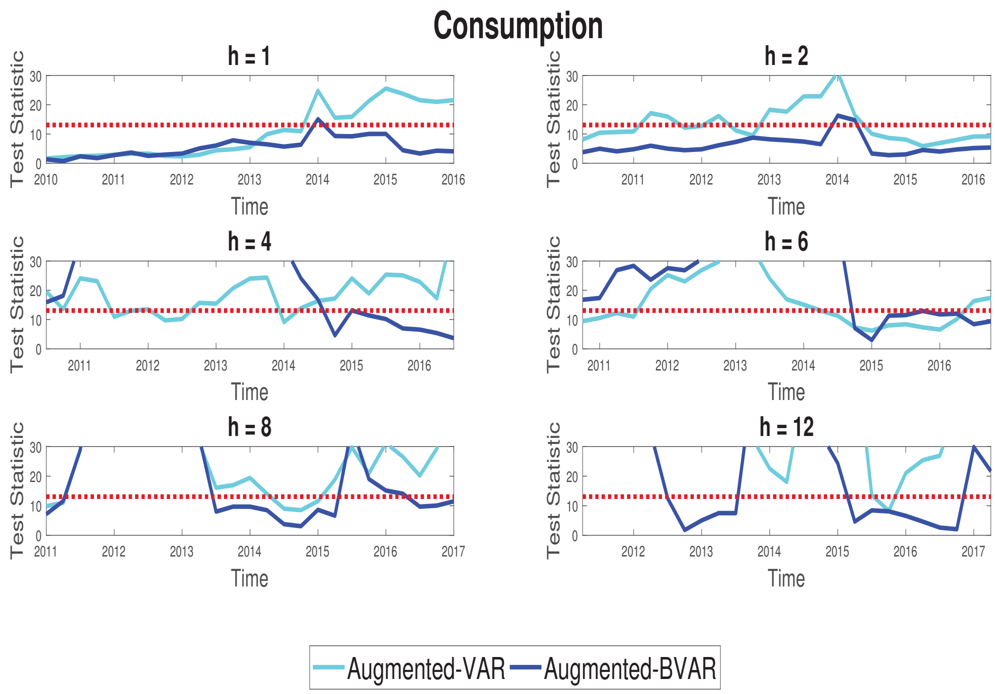
In this section we explicitly confront the Augmented–(B)VAR models forecast absolute and relative stability. For the relative forecast stability test, we compare A(B)VARs with their reduced–form counterparts. The aim of this comparison is that of learning where the contribution of the DSGE model extracted exogenous process matters the most.

A.4.1 Augmented–(B)VAR absolute Forecast Stability Analysis

In this section we explore the absolute forecast stability of A(B)VARs. Figures 16 and 17 report results for the fluctuation rationality test. As it is possible to see, results are broadly in line with those of (B)VARs.

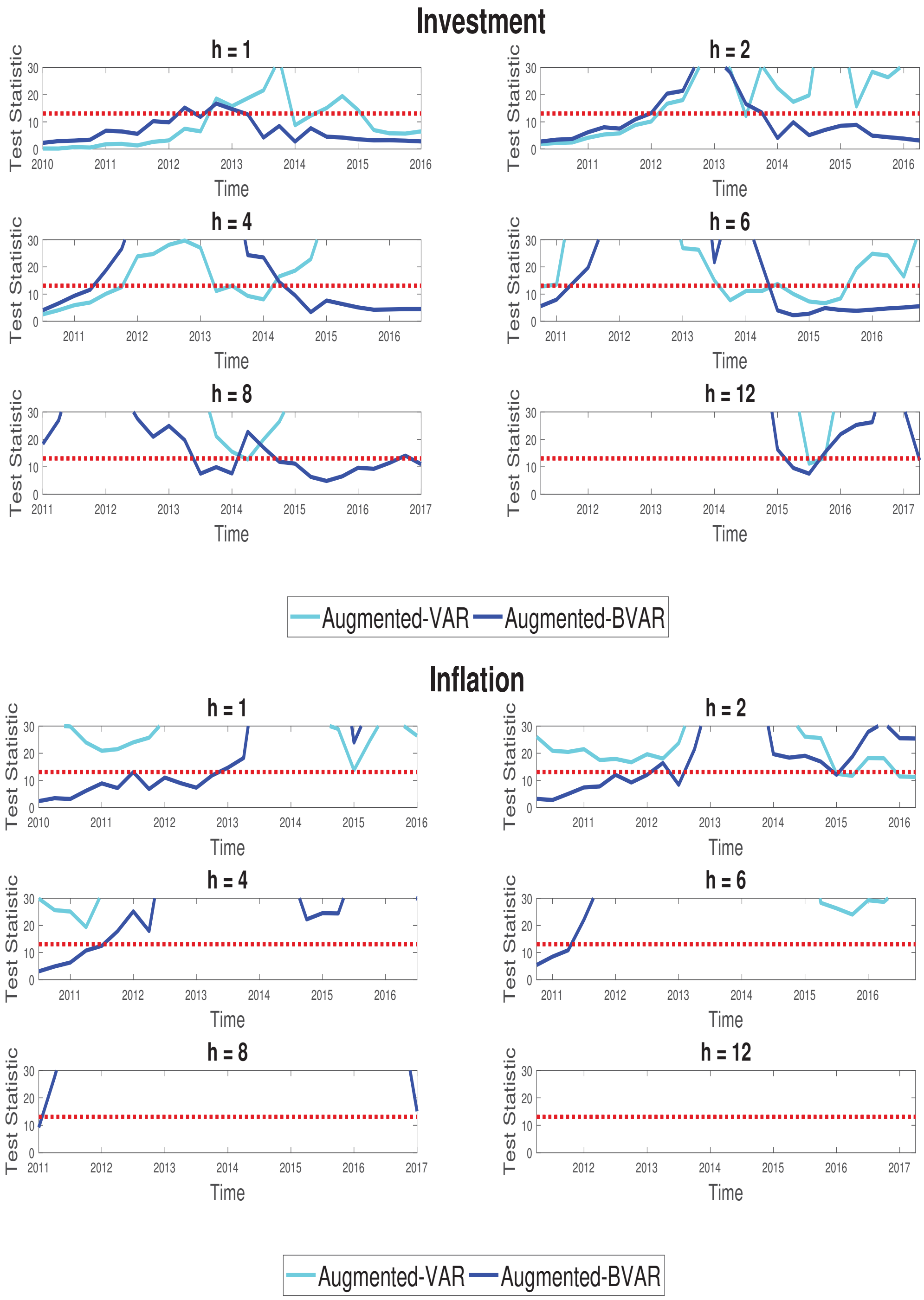
FIGURE 16. Rossi and Sekhposyan (2016) Rationality Test.





Note: Absolute local rationality forecasting performance of A(B)VAR for GDP (ΔY_t) and consumption (ΔC_t) growth rates along different forecast horizons. When the statistics is above the critical value, the null hypothesis of forecast rationality is rejected. The statistics are reported for several steps ahead (h). Red dotted lines indicate the critical values at the 5% of confidence level.

FIGURE 17. Rossi and Sekhposyan (2016) Rationality Test.

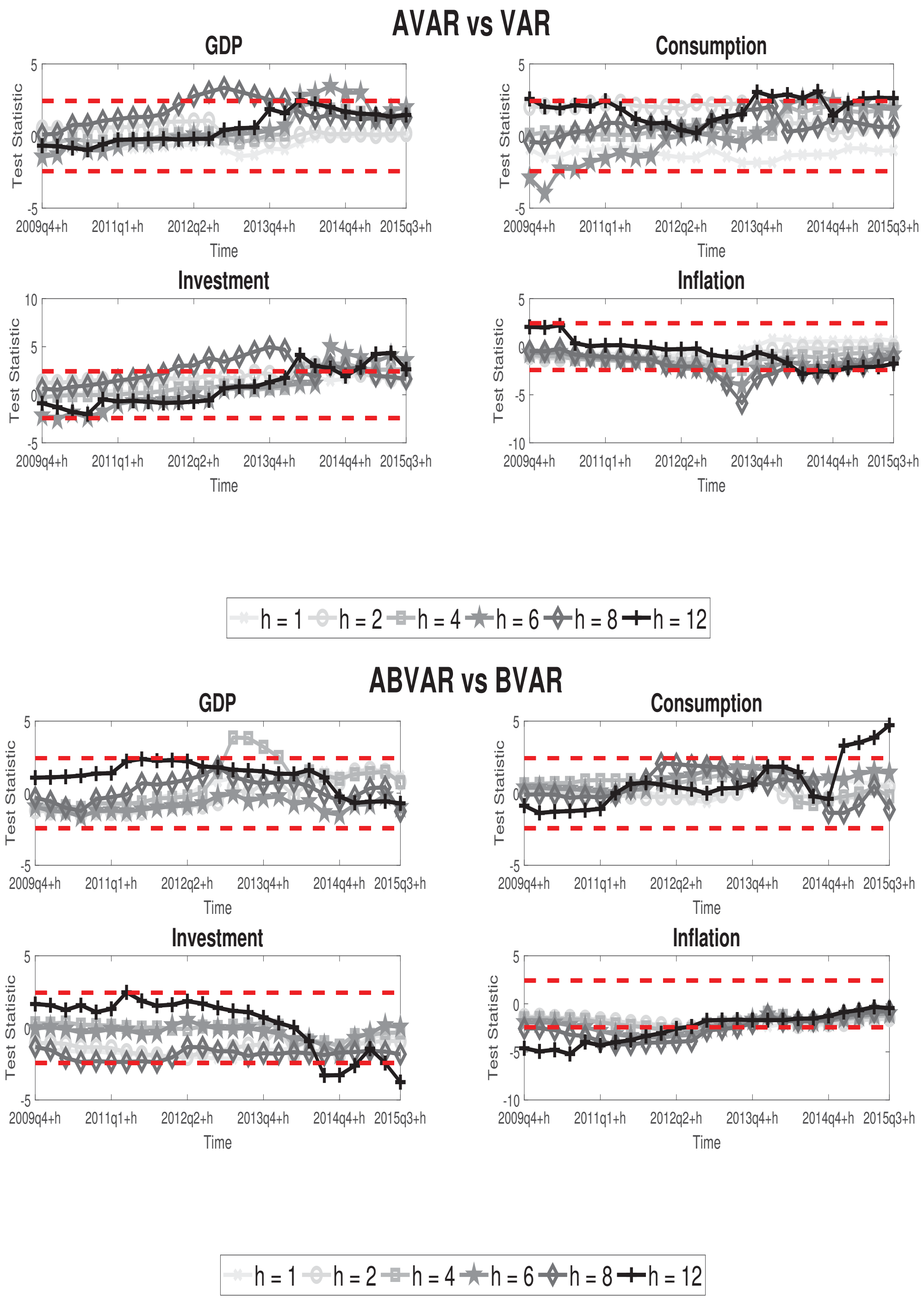


Note: Absolute local rationality forecasting performance of A(B)VAR for investment growth rates (ΔI_t) and inflation (π_t) along different forecast horizons. When the statistics is above the critical value, the null hypothesis of forecast rationality is rejected. The statistics are reported for several steps ahead (h). Red dotted lines indicate the critical values at the 5% of confidence level.

A.4.2 Augmented–(B)VAR relative forecast stability analysis

Figure 18 reports the fluctuation test statistics. As to consider both the AVAR vs VAR and the ABVAR vs BVAR comparison figuring out a clear pattern is a difficult task. Overall it emerges a contrast between the set of real variables and inflation. The former tends to favour reduced–form models (with some exceptions for investment at some specific forecast horizon), the latter reveals a better performance of Augmented–(B)VARs (especially conditional on the imposition of a statistical prior). An explanation to this fact, can be found looking at Figure 3, where the here considered labour supply shock is the one yielding the best relative forecast for inflation along (almost) all the steps ahead for both Augmented–VAR and –BVAR. This polarisation is not so evident for real variables and thus, in a sense, the relatively better forecasting performance for inflation is not surprising. In addition, we observe that Augmented–(B)VAR relative forecast statistics follow a similar pattern for different steps ahead. This implies a tendency of Augmented–(B)VAR to have relatively better short forecasting performances as the inclusion of the financial crises in the evaluation sample increases. This is also true for longer run relative forecasts.

FIGURE 18. Giacomini and Rossi (2010) Fluctuation Test.

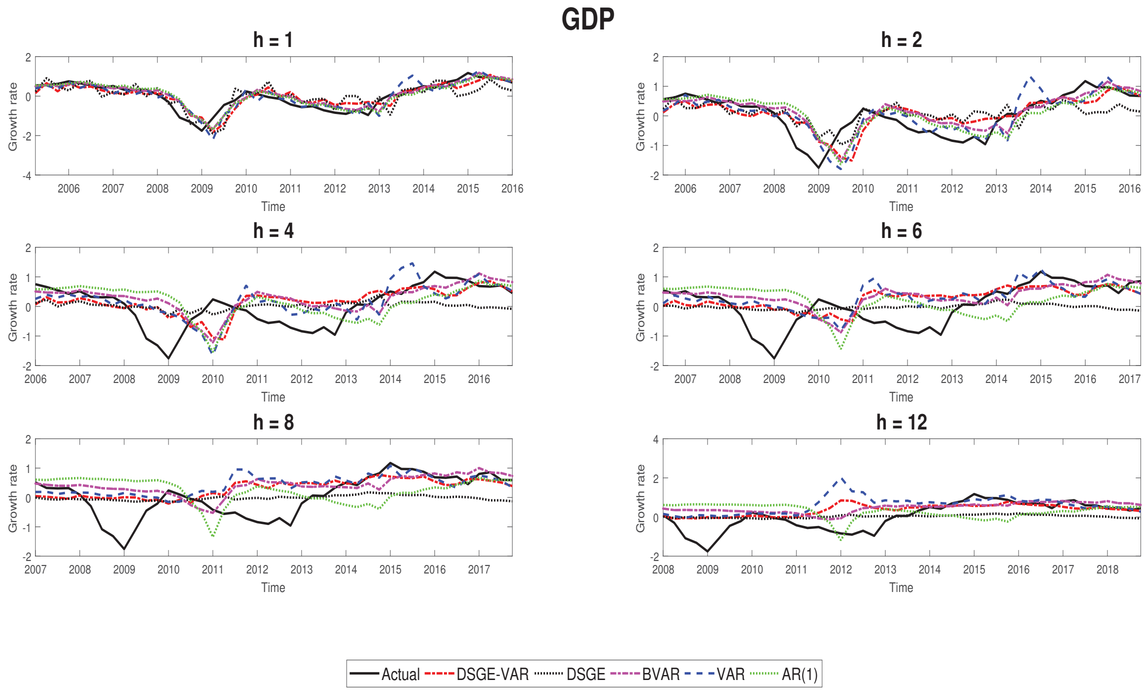


Note: Pairwise local relative forecasting performance of Augmented–(B)VAR against (B)VAR for GDP (ΔY_t) and consumption (ΔC_t) growth rates along different forecast horizons. When the statistic is below (above) the lower (upper) critical value, the Augmented–(B)VAR forecasts significantly better (worse) at that point in time.

A.5 Models Generated Forecasts

A.5.1 GDP Forecasts

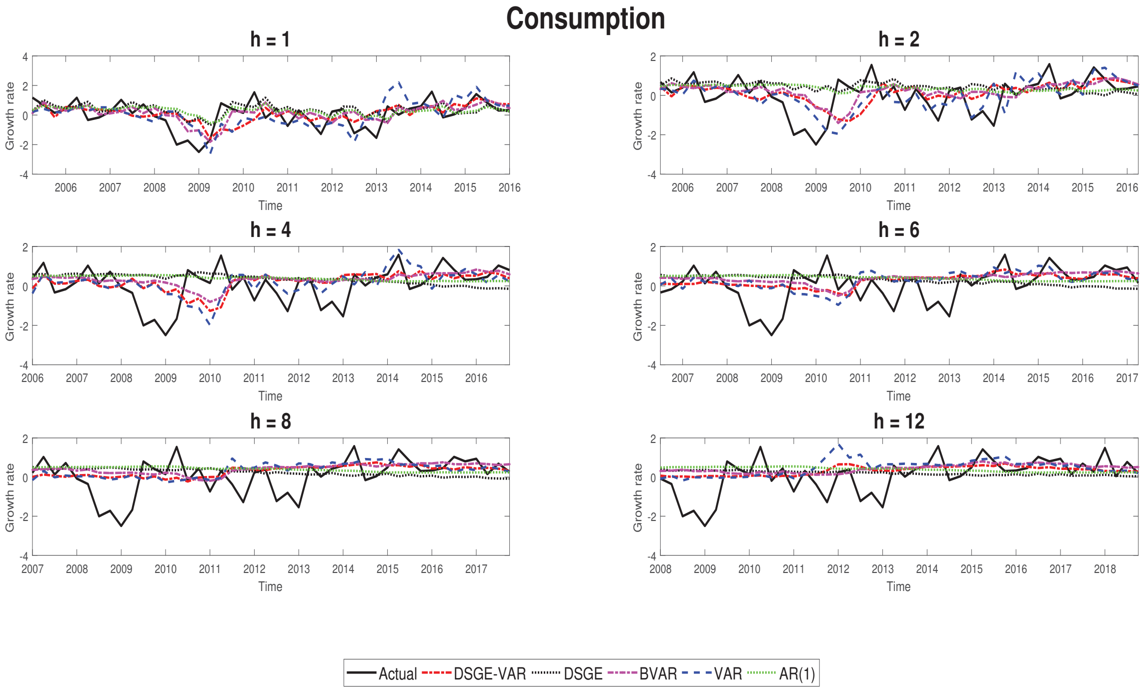
FIGURE 19. GDP growth generated forecasts.



Note: GDP growth forecasts along several steps ahead (h) for DSGE–VAR, DSGE, BVAR, VAR and AR.

A.5.2 Consumption Forecasts

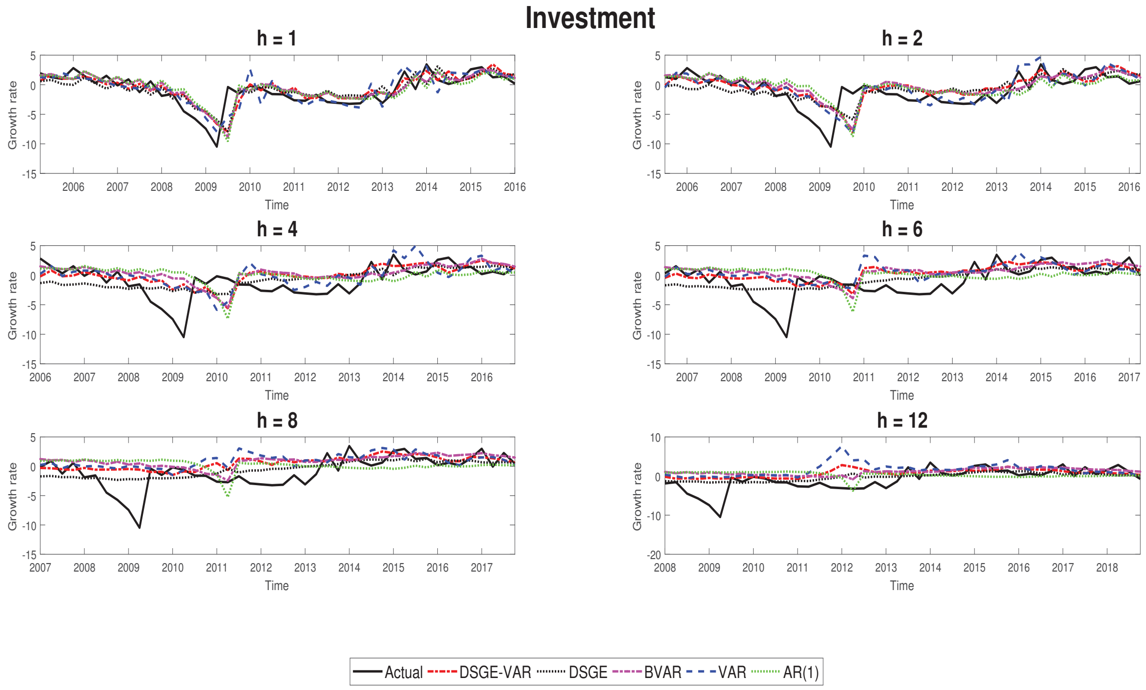
FIGURE 20. Consumption growth generated forecasts.



Note: Consumption growth forecasts along several steps ahead (h) for DSGE–VAR, DSGE, BVAR, VAR and AR.

A.5.3 Investment Forecasts

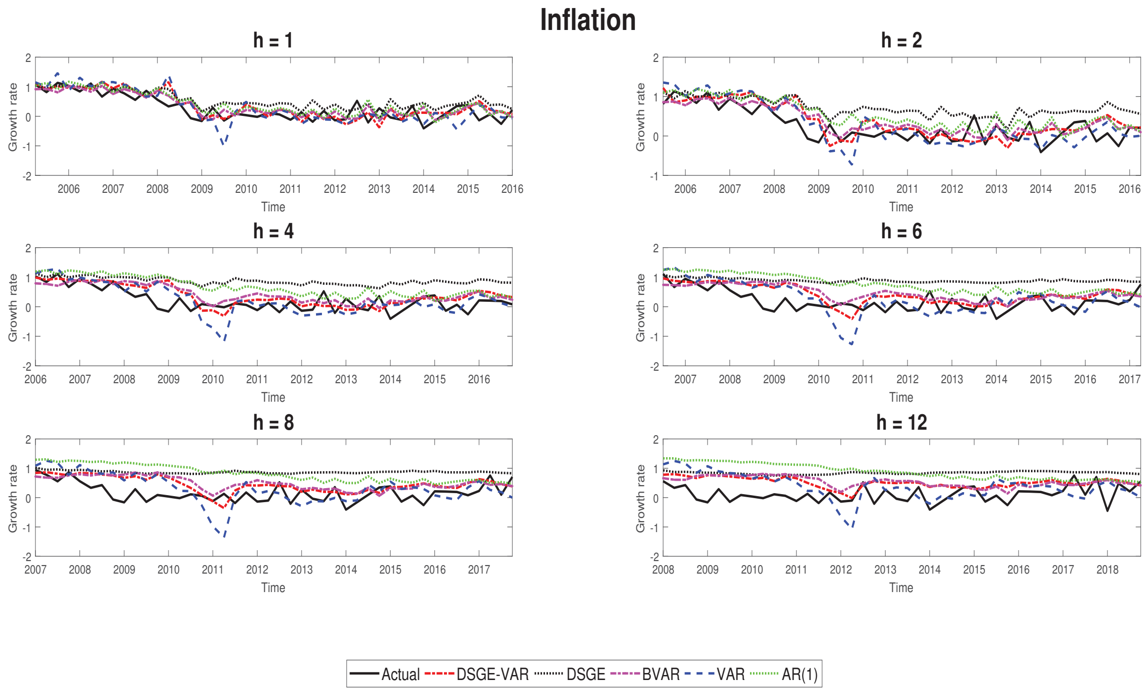
Figure 21. Investment growth generated forecasts.



Note: Investment growth forecasts along several steps ahead (h) for DSGE–VAR, DSGE, BVAR, VAR and AR.

A.5.4 Inflation Forecasts

Figure 22. Inflation growth generated forecasts.



Note: Inflation growth forecasts along several steps ahead (h) for DSGE–VAR, DSGE, BVAR, VAR and AR.

Acknowledgements

We are extremely grateful to Marco Del Negro, Jesús Fernández-Villaverde, Philip Lane, Diego Rodríguez-Palenzuela, Barbara Rossi, Juan F. Rubio-Ramírez and seminar participants at the European Central Bank, at the 2019 International Association for Applied Econometrics Annual Conference (Nicosia), and at the 3rd Forecasting at Central Banks Conference (Ottawa) for their stimulating and helpful comments. The views expressed in the paper are those of the authors and do not necessarily reflect those of the ECB, the Banco de España or the Eurosystem. No part of our compensation was, is or will be directly or indirectly related to the specific views expressed in this paper. All remaining errors are our own responsibility.

Jaime Martínez-Martín
European Central Bank, Frankfurt am Main, Germany; Banco de España; email: jaime.martinez-martin@ecb.europa.eu

Richard Morris
European Central Bank, Frankfurt am Main, Germany; email: richard.morris@ecb.europa.eu

Luca Onorante
European Central Bank, Frankfurt am Main, Germany; email: mailto:luca.onorante@ecb.europa.eu

Fabio M. Piersanti
European Central Bank, Frankfurt am Main, Germany; email: fabio_massimo.piersanti@ecb.europa.eu

PDF ISBN 978-92-899-3904-1 ISSN 1725-2806 doi:10.2866/939095 QB-AR-19-116-EN-N