Forecasting Ireland's Macroeconomy: A Large Bayesian VAR Approach

Killian Carroll

Working Paper No. 12

30th September 2020

Irish Fiscal Advisory Council

Forecasting Ireland's Macroeconomy: A Large Bayesian VAR Approach

Killian Carroll[1]

Version: September, 2020

Abstract

This paper develops a forecasting model of the Irish macroeconomy using a large Bayesian Vector Auto-Regression (VAR) model. We show that large Bayesian VARs can be a useful tool in forecasting Ireland's macroeconomy. We analyse the performance of the large Bayesian VAR at forecasting 20 macroeconomic variables for Ireland and find that it performs relatively well versus naïve models, ARIMA models, smaller VARs, and Factor-Augmented VARs. In particular, we find that the model performs well at forecasting variables that capture the underlying performance of the Irish macroeconomy.

Keywords: Bayesian VAR, Forecast, Macroeconomics, Small open economy
JEL No. C11, C32, C53, E27, E37

This report can be downloaded at www.FiscalCouncil.ie

[1] The author is a Research Assistant at the Irish Fiscal Advisory Council. Email: admin@fiscalcouncil.ie. The opinions expressed and arguments employed in this paper do not necessarily reflect the official views of the Irish Fiscal Advisory Council. The author would like to acknowledge the assistance of members of the Council and Secretariat of the Irish Fiscal Advisory Council, in particular, Eddie Casey and Sebastian Barnes, as well as excellent comments received by Michael O'Grady (Central Bank of Ireland).

1. Introduction

Formulating appropriate macroeconomic and fiscal policy requires “knowledge” of the future state of the economy. Yet, a wide range of dynamics can affect the economy in the short run. Forecasting the macroeconomy accurately may therefore require us to draw on information from a large number of variables.

There are challenges when using more variables. Often, historical data availability is limited. Modelling the macroeconomy in a system gives rise to a large number of parameters, which need to be estimated with only a limited data set. This problem is often referred to as the “curse of dimensionality”. This can in turn lead to in-sample overfitting and large out-of-sample forecasting errors. A common means of forecasting the macroeconomy is by using a Vector Auto-regression (VAR) model. VAR models have a long history of being a useful tool for the macroeconomist in forecasting and policy analysis.

To avoid the “curse of dimensionality”, conventional VARs usually only consist of a small number of variables. However, estimating VARs with a relatively small number of variables can create an omitted variable problem, leaving out potentially useful information, which can contribute to poor out-of-sample forecasting performance.

This paper builds a large Bayesian VAR model for forecasting Ireland’s underlying macroeconomy. We use the technique of “Bayesian shrinkage” to handle a large number of variables in a VAR framework. That is, we increase the tightness of the prior distribution around its central estimate as the number of variables increases. This reduces overfitting that occurs in larger conventional VARs and reducing the impact of omitted variable problems that smaller VARs are prone to. As Bańbura *et al*. (2010) have shown, large Bayesian VARs offer an alternative to factor models for analysis and forecasting of large dynamic systems. Using Bayesian shrinkage allows the modeller to include a larger information set in modelling the macroeconomy. For instance, information on business and consumer sentiment can be included alongside more disaggregated variables that may provide vital information for providing accurate forecasts of key macroeconomic variables.

We evaluate the forecasting accuracy of the large Bayesian VAR using a number of metrics. First, the forecasting performance of the large Bayesian VAR is compared to that of a small VAR with only a few key variables. Secondly, we compare the forecasting performance of the large Bayesian VAR to a Factor-Augmented VAR. We extend on previous literature by assessing the forecasting performance of the large Bayesian VAR for a wider set of variables than is common practice. Typical in the literature for large Bayesian VARs is to investigate the performance in forecasting only a small number of variables—usually only three. We investigate whether a large Bayesian VAR can be useful in forecasting as many as 20 variables. We compare the forecasting performance of the large Bayesian VAR to a naïve forecast, and to an ARMA forecast for each of the 20 variables. We also report the root mean squared forecast error for these variables. Based on these metrics, the Large Bayesian VAR performs favourably for a number of underlying variables for the Irish macroeconomy. For instance, the model performs relatively well at forecasting the annual growth rates of underlying variables such as Underlying Domestic Demand, Employment, Government Consumption, Personal Consumption, Disposable Income, and also performs relatively well at forecasting growth rates of headline variables such as GDP, GNP, Exports and Imports.

Large Bayesian VARs have been used for forecasting, policy analysis, and analysis of shock transmission. Bańbura *et al*. (2010), studying the US economy, show that large Bayesian VARs have a superior forecasting performance to that of smaller VARs and Factor Augmented VARs. This is a finding that is repeated in the literature. Gupt & Kabundi (2010) find that, in general, large Bayesian VARs outperform smaller conventional VARs, smaller Bayesian VARs, dynamic factor models and a small open economy DSGE model in forecasting the South African macroeconomy. Koop (2013) investigate how the forecasting performance of large Bayesian VARs varies with different priors and finds that large Bayesian VARs tend to have a better forecasting performance than that of factor models and find that the simple Minnesota prior performs relatively well in forecasting. Large Bayesian VARs have been used to estimate the effects of policy on outcomes. De Menezes Barboz & Vasconcelos (2019) show, using a large Bayesian VAR, that the Brazilian Development Bank had a positive impact on Brazilian aggregate investment. Bloor & Matheson use a large Bayesian VAR to analyse the transmission mechanism of shocks to monetary policy, net migration and climate for the New Zealand economy.

In an Irish setting, several different approaches to forecasting the macroeconomy, exist or have been attempted. Short-run models for key macroeconomic variables have been developed using a suite of models approach (Conroy & Casey, 2017). Medium term forecasts have been developed using the HERMES model (Bergin *et al*., 2013), and more recently using the COSMO structural model (Bergin *et al*., 2017). In the past, Bayesian VAR models have been used to forecast inflation in Ireland (Kenny *et al*., 1998), although, the Bayesian VARs estimated in this case were of a more conventional size (3-5 variables). In addition, estimates and forecasts of the supply side of the Irish economy have been developed in recent years (Casey, 2019; Murphy *et al*. 2018).

The approach of this paper is not necessarily to create competing forecasts, but instead, to create complementary forecasts of the Irish economy. Large scale Bayesian Vector Autoregressions are largely a "black box" and it can be difficult to disentangle why a given variable is displaying certain dynamics. To that end, models such as those in Conroy & Casey (2017) can provide a useful complement to the model outlined in this paper, by providing more granular information about the drivers of different variables. To the extent that different sets of models diverge, this divergence can also provide useful information. This will allow the practitioner to highlight potential problems in their modelling tool kit and—one would hope—eventually improve their forecasts.

2. Methodology

The model we employ follows that of Bańbura *et al*. (2010) closely. Using a standard VAR setup, let $Y_t = (y_{1,t}, y_{2,t}, \dots y_{n,t})$ be a large vector of random variables. Then, a VAR(p) model is given by:

$$Y_t = c + A_1 Y_{t-1} + \cdots + A_p Y_{t-p} + u_t \tag{1}$$

Where $A_{1,} \dots, A_p$ are matrices of dimension $n \times n$; u_t is an *n*-dimensional normally distributed white noise process with covariance matrix $\mathbb{E}u_t u_t' = \Psi$; $c = (c_1, \dots c_n)$ is an *n*-dimensional vector of constants.

The model is estimated using the Bayesian VAR approach, and following the literature, a Minnesota type prior is used (Litterman, 1986). In particular, we use a natural conjugate extension of the Minnesota prior, which has been used extensively in estimation of large Bayesian VARs. Koop (2013) find that this Minnesota prior forecasts relatively well, in comparison to other priors, in medium and large VARs and has the added benefit of being computationally straightforward to implement.[2] Under a Minnesota type prior, each equation is centred around a random walk with drift:

$$Y_t = c + Y_{t-1} + u_t \tag{2}$$

which essentially shrinks the diagonal coefficients in A_1 towards one, and all other coefficients towards zero. Under this prior, more recent lags provide more information than previous lags, and the "own" lags explain more variation than the lags of other variables. The priors are imposed by setting the moments of the prior distribution of the coefficients:

$$\mathbb{E}[(A_k)_{i,j}] = \begin{cases} \delta_i, & j = i, k = 1 \\ 0, & otherwise \end{cases}, \qquad \mathbb{V}[(A_k)_{i,j}] = \begin{cases} \dfrac{\lambda^2}{k^2}, & j = i \\ \vartheta \dfrac{\lambda^2}{k^2} \dfrac{\sigma_i^2}{\sigma_j^2}, & otherwise \end{cases} \tag{3}$$

[2] For instance, the direct forecasting method can be used and predictive simulation is not required in this case. See Appendix B for further details.

The coefficients $A_{1,} \dots, A_p$ are assumed to be independent and normally distributed. The prior on the intercept is diffuse, while the covariance matrix of residuals is assumed to be diagonal, fixed and known ($\Psi = \Sigma$, where $\Sigma = (\sigma_i^2, \dots, \sigma_n^2)$).

Setting $\delta_i = 1$ is equivalent to a random walk prior, which is appropriate for variables that have high persistence and are not characterised by mean reversion. However, for variables that are characterised by mean reversion, a more appropriate prior would be setting $\delta_i = 0$. This is the equivalent of setting a white noise prior. In other words, $\delta_i = 0$ is appropriate for variables that are stationary, or $I(0)$, whereas, setting $\delta_i = 1$ would be appropriate for variables that are stationary after taking the first difference, i.e. $I(1)$.[3]

The hyperparameter λ controls the overall tightness of the prior distribution around the central estimate for the prior. The relative weight attached to the prior beliefs with respect to the data is determined by λ. Setting $\lambda = 0$, imposes the prior exactly: that is, the data does not influence the parameter estimates, and our prior then becomes our estimate. Setting $\lambda = \infty$, negates the influence of the prior, and the estimates of the parameters are equivalent to ordinary least squares (OLS) estimates. As argued by Bańbura *et al*. (2010) and shown formally by De Mol *et al*. (2008), the overall tightness of the parameters, λ, should be set relative to the size of the system. As the size of the system becomes larger, more shrinkage should be applied to the parameters in order to avoid over-fitting.

As the lag length increases, the prior variance decreases by the factor $1/k^2$. The factor σ_i^2/σ_j^2 is used to account for the differences in scales and variance of the different time series. Finally, the coefficient $\vartheta \in (0,1)$ determines how important lags of other variables are relative to own lags. A normal Wishart prior is imposed in order to take into account the possibility that there is correlation among the residuals of different variables. This prior retains the principles of a Minnesota prior under the condition that $\vartheta = 1$. See Appendix B for a mathematical description of the implantation of the Minnesota prior.

[3] $I(d)$ represents the order of integration, that is the number of times, d, that a variable must be differenced in order to achieve stationarity.

Up to now, the model outlined here assumes a symmetric form, that is, each variable is assumed to be endogenous. However, as Ireland is a small open economy, foreign variables will have an effect on domestic variables and the corresponding effect of Irish variables on foreign variables will be relatively limited. As such, the approach we take here is to treat the foreign variables as exogenous to the domestic variables. This is a similar approach taken to estimates of large Bayesian VARs for other small open economies such as for South Africa (Gupta & Kabundi, 2010) and for New Zealand (Bloor & Matheson, 2009). Both Gupta & Kabundi (2010), and Bloor & Matheson (2009) estimate the Bayesian VAR by dividing the variables into blocks, a domestic block, which depends on foreign and domestic variables, and a foreign block which depends on foreign variables only. However, we differ from these approaches in that we do not explicitly model the foreign variables within the Bayesian VAR framework. That is, the foreign variables are treated as entirely exogenous to the Bayesian VAR and are forecasted in a separate exercise to the model presented here.[4] The implementation of the foreign variables as exogenous to the rest of the model is discussed further in Appendix B.

[4] The exogenous variables are derived separately. For example, the external demand variables are built up based on weighted imports growth for Ireland's main trading partners. The forecasts of these exogenous variables are derived using forecasts for imports growth from international forecasting bodies such as the OECD or the IMF. When calibrating λ and evaluating the models forecasting performance, realised values of these foreign variables are used, as opposed to real time values. That is, when forecasting out-of-sample, the realised values for these exogenous variables are used as inputs in estimating the forecasts of the endogenous variables. This may improve the forecasting performance of the model, but it will also isolate the forecasting error that is solely attributed to the large Bayesian VAR, as opposed to an error that is also partly a result of an error in forecasting the exogenous variables. As such, this approach provides a cleaner evaluation of the models forecasting performance.

3. Data and Estimation

We use a dataset of 46 variables, drawing from a broad category of variables that are relevant to the Irish macroeconomy, such as, national account data, labour market data, financial data and some variables with an international dimension. The data spans from Q1 2000 to Q4 2018. Data are seasonally adjusted, where needed, using the US Census Bureau's X-13 ARIMA SEATS procedure.[5] Data are transformed to logarithms for all variables, except those that are in rates. Where needed, we difference variables to achieve stationarity. Details on the variables used and the transformations applied are in Appendix A.

As all variables are transformed to stationarity, the white noise prior is used. That is, $\delta_i = 0$ for all variables. As is convention with quarterly models of this type in the literature, a lag length of 4 was chosen.[6]

As the choice of λ is largely arbitrary, in calibrating λ, we follow the procedure of Bańbura *et al.* (2010). That is, we choose the value of λ that achieves a similar fit, for key variables in the large BVAR, as that of a small VAR with the same key variables.

The standard choice of variables in the literature for the small VAR is to have a three variable VAR, usually with GDP/employment, consumer price index (CPI), and the interest rate. For instance, Bańbura *et al.* (2010) calibrate λ using a three variable VAR of employment, CPI and Federal Funds Rate. However, there are a number of reasons why both the size, and the choice of variables for the small VAR might not be entirely appropriate for Ireland. First, as a small open economy, Ireland's macroeconomy is significantly influenced by external demand. Second, Ireland's national account statistics are heavily distorted by the presence of multinationals, and as such, variables such as GDP are not representative of the underlying production in the macroeconomy. In the same vein as Bloor & Matheson (2009), to address some of these issues, we use a slightly larger VAR, to calibrate λ. We use a five variable VAR, with Underlying Domestic Demand (UDD), wages, employment,

[5] Where seasonally adjusted variables are already readily available these are used. This is often the case for variables from the CSO. The CSO seasonally adjust variables using the same US Census Bureau seasonal adjustment process.

[6] Arbitrarily setting the lag length in this fashion is not uncommon in the literature for large Bayesian VARs, see for instance, Domit *et al.* (2019). A similar approach to estimating a large Bayesian VAR for New Zealand using quarterly data was taken by Bloor & Matheson (2009).

the ECB deposit rate, and one external demand variable which is treated as exogenous.[7,8] However, Ireland is a member of a monetary union. Therefore, while the ECB deposit rate is important in determining the supply and the demand for credit in Ireland, the degree to which the Irish macroeconomy has a bearing on how the deposit rate is set is minimal. Forecasting the deposit rate solely on the basis of Irish variables would therefore not be appropriate. Additionally, the goal of this paper is not to forecast the deposit rate. Therefore, when determining λ we minimise the difference in the fit of the three key variables (UDD, wages and employment) between the smaller five variable VAR and the larger 46 variable VAR.

More formally, λ is chosen using the following equation:

$$\lambda_m(Fit) = \arg Min_\lambda \left| Fit - \frac{1}{3}\sum_{i \in W} \frac{MSFE_i^{(\lambda\,,\mathrm{m})}}{MSFE_i^{(0)}} \right| \tag{4}$$

Where MSFE is the one step-ahead mean squared forecast error, for horizon of length H, and is given by:

$$MSFE_{i,h}^{(\lambda\,,\mathrm{m})} = \frac{1}{T_1 - T_0 - H + 1} \sum_{T=T_0+H-h}^{T_1-h} (y_{i,T+h|T}^{(\lambda,m)} - y_{i,T+h})^2 \tag{5}$$

Where, for a given model, m, and tightness given by λ, the h-step ahead forecast for variable y_i is given by $y_{i,T+h|T}^{(\lambda,m)}$. $MSFE_i^{(0)}$ is the h-step ahead mean square forecast error for the model where the prior is imposed exactly, that is, $\lambda = 0$. The purpose of this term is to account for the differences in the relative scales of the variables.

Finally, Fit is given by:

[7] Underlying Domestic Demand is an indicator that contains Personal Consumption Expenditure, Government Consumption and Underlying Investment. Underlying investment strips out intangibles and aircraft investment as these are, largely, imported and are therefore largely GDP neutral.

[8] The external demand variable is the simple average of an index for external goods demanded and an index of external services demanded The indices for goods and for services demanded are compiled by taking the share of each trading partner's exports of goods and of services from Ireland, and growing these exports forward based on the forecasts of the trading partner's demand for imports of goods (for goods) and of combined goods and services (for services). This is then aggregated into world-demand indices for goods and for services.

$$Fit = \frac{1}{3} \sum_{i \in W} \frac{MSFE_i^{(\lambda ,\mathrm{m})}}{MSFE_i^{(0)}} \Bigg|_{\lambda = \infty, m=SMALL} \tag{6}$$

Where $w = \{UDD,\ wages,\ employment\}$ and the model SMALL consists of variables UDD, wages, employment, the ECB's deposit rate, and an external demand variable.

The overall tightness, λ, was estimated using a training sample period from Q1 2000 to Q4 2009. The tightness parameter, λ, is then kept constant for evaluating the out-of-sample performance. The plot of the absolute difference of the MSFE for the small VAR and the large Bayesian VAR is shown for various values for λ is shown in Figure C1 in Appendix C.

4. Forecasting Performance

To evaluate the model performance, the out-of-sample forecasting accuracy was investigated across several metrics and over a number of different horizons.[9] We assess out-of-sample forecast performance using four-quarter ahead year-on-year growth rates and the four-quarter ahead annualised growth rates.[10] The four-quarter-ahead year-on-year forecast can be thought of as a means of assessing the performance at a particular future point in time. Whereas the four-quarter-ahead annualised forecast is an accumulation of forecasts over a period of time and can therefore, be thought of as a means of assessing the cumulative forecast performance over that horizon.

One metric by which to assess the forecasting performance against, is to compare the errors of the large Bayesian VAR with that of other models. The standard approach in the literature to assess the performance of large Bayesian VAR models is to compare it to that of the smaller VAR of key variables that is used in calibrating λ. This is one of the approaches taken here. Another approach taken in the literature is to evaluate the forecasting performance of the large Bayesian VAR for key variables against that of a Factor-Augmented VAR. The Factor-Augmented VAR makes use of all the same data as the large Bayesian VAR, but instead of using shrinkage on the parameters, extracts a number of factors from the data, to reduce the dimensions and overcome the "curse of dimensionality".

Unlike most of the previous studies in this literature—where the focus is typically only on a few variables—we investigate the forecasting performance for a wide range of variables. We assess the performance of the large Bayesian VAR in forecasting 20 variables of the Irish macroeconomy. We assess these based on metrics such as Theil's U2, the Root Mean Squared Forecast error (RMSFE), and against the forecasts of ARMA models for each of the variables.

A qualitive analysis is also carried out to assess the model's performance in forecasting recent turning points in several key variables for Irish economic data.

[9] Real time data is not used in evaluating the out-of-sample forecasting performance.

[10] For variables in rates, such as the Unemployment rate, the 4 quarter ahead forecasted change in the rate is assessed.

4.1 Relative Forecasting performance with the Small VAR

Following the convention in the literature, the relative out-of-sample forecasting performance of the large Bayesian VAR versus the small VAR is assessed using the following formula:

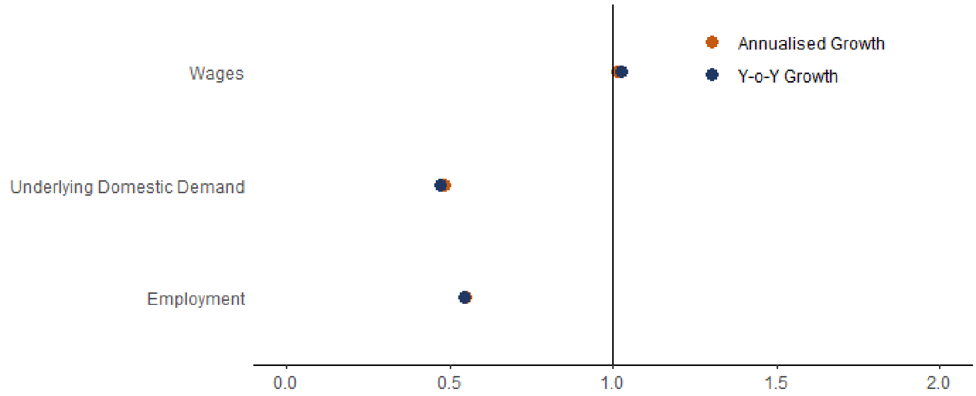
$$Relative\ MSFE_i = \frac{MSFE_i^{(\lambda,\mathrm{m})}}{MSFE_i^{(Model)}} \tag{7}$$

where, in this case $Model = SMALL$. Given that the metric outlined in equation 7 is non-linear, care should be taken when interpreting the relative MSFE. While a value below (above) one indicates that the Large Bayesian VAR overperformed (underperformed) the small VAR in forecasting, the magnitude of the overperformance/underperformance is not linearly comparable. For instance, a value of 2 for the relative MSFE, would indicate that the small VAR was twice as good as the large Bayesian VAR at forecasting a particular variable, while in the other direction, a value of 0.5 would indicate that the large Bayesian VAR was twice as good as the small VAR at forecasting the same variable.

Figure 1 shows a plot of the relative MSFE for the four quarter ahead year-on-year growth rate and for the one year ahead annualised growth rate for the key variables. The large Bayesian VAR outperforms the forecasting performance of the small VAR for both Underlying Domestic Demand and Employment. The large Bayesian VAR performs roughly twice as well as that of the small VAR in forecasting these variables. It performs on a par with the forecasts for wages, (the small VAR being marginally better, see Table D1).

While the sample size is small (only 36 observations), we also use the Diebold-Mariano test to compare the predictive accuracy and test whether the two forecasts are statistically different from each other (see Diebold & Mariano, 2002). The large Bayesian VARs forecasts for the annualised growth rate of Employment are statistically different from those of the small VAR at a 10 per cent significance level (see Table D1). All other forecasts in this case are not statistically different from each other according to the Diebold-Mariano test.

Figure 1: Relative MSFE of the large Bayesian VAR versus the Small VAR for Q1 2010 to Q4 2018



Sources: Author's calculations.
Note: The Vertical black line marks the relative MSFE value of one. Values below one indicates that the large Bayesian VAR has a superior forecasting performance to the small VAR. Corresponding values are presented in Appendix Table D1.

4.2 Relative Forecasting performance with the Factor-Augmented VAR

Next, we turn to comparing the large Bayesian VAR with a Factor-Augmented VAR. Factor models have been used to extract relevant information from a large amount of data and summarise the data in a small number of factors. This greatly reduces the dimensions of the data. This approach to handling large datasets is another way of overcoming the "curse of dimensionality". The approach we take here is to extract principle components from the data and then include these principle components in a VAR framework. More specifically, we use the same three key variables in the small VAR (wages, UDD, employment) and augment the VAR with principle components. This is known as a Factor-Augmented VAR.[11] Comparing the performance of the large Bayesian VAR with that of the Factor-Augmented VAR has the added benefit that both models are based on the exact same information set.

As with the large Bayesian VAR, all variables are transformed to stationarity. Variables are then standardised before principle components are extracted.[12] We experimented with using several different principle components and using a number of different lags for the Factor-Augmented VAR. However, using one

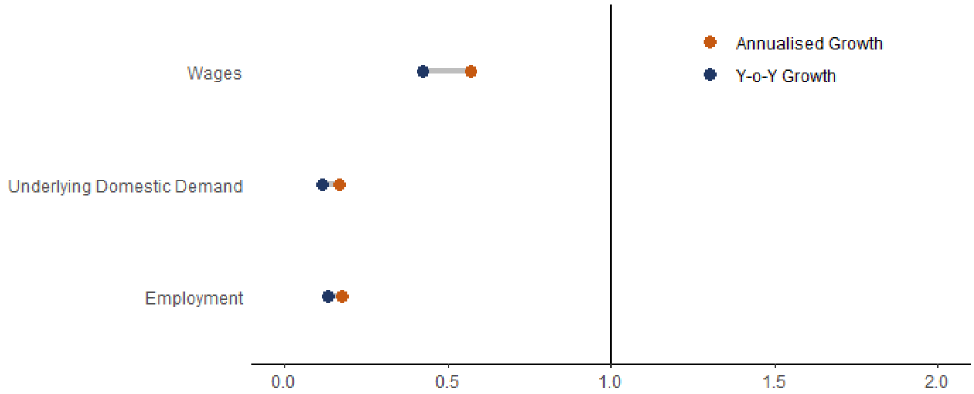
[11] For a detailed outline of the Factor-Augmented VAR modelling approach, see, amongst others, Bernanke et al. (2005).

[12] This is the standard approach to extracting principle components because principle components are not scale invariant.

principle component and one lag achieved the best out-of-sample forecasting performance.[13]

Using the relative MSFE to compare the forecasting performances of the two models, Figure 2 shows a plot of the four-quarter-ahead year-on-year growth rate and for the four-quarter-ahead annualised growth rate for the three key variables. The large Bayesian VAR performs better than the Factor-Augmented VAR at forecasting each of the three key variables. The large Bayesian VAR performs approximately twice as well as the Factor-Augmented VAR in forecasting Wages, and roughly five to eight times better at forecasting UDD and Employment. However, using the Diebold-Marino test, the forecasts of the large Bayesian VAR and the Factor-Augmented VAR were not statistically different from each other (Table D.2).

Figure 2: Relative MSFE of the large Bayesian VAR versus the Factor-Augmented VAR



Sources: Author's calculations.
Note: The Vertical black line marks the relative MSFE value of one. Values below one indicates that the large Bayesian VAR has a superior forecasting performance to the Factor-Augmented VAR. Corresponding values are presented in Appendix Table D2.

The usefulness of interpreting Figures 1 and 2 in evaluating the forecasting performance of the large Bayesian VAR is entirely dependent on how good the small VAR and the Factor-Augmented VAR are in forecasting these variables. If either the small VAR or the Factor-Augmented VAR are not good at forecasting the variables in question, then Figure 1 and 2 will tell us little about the forecasting performance of the large Bayesian VAR. Figures 1 and 2 will simply only tells us that the large Bayesian VAR is better at forecasting than the small VAR or the Factor-Augmented

[13] The first principle component accounted for 22.7 per cent of the variance in the data.

VAR. As a result, other metrics are needed to evaluate the forecasting performance of the large Bayesian VAR. These are discussed in the following section.

4.3 Evaluating the models forecasting performance of a wider range of variables

This section assesses a wider set of variables for the models' out-of-sample forecasting performance. We do so for four-quarters-ahead year-on-year growth rates and the four-quarter-ahead annualised forecast. In particular, we evaluate the forecasting performance of the variables based on Theil's U2, which compares the forecast of the large Bayesian VAR with that of a naïve forecast. The naïve forecast is simply a prediction that the forecast value will be the same as the current value. That is, for a four-quarter-ahead year-on-year growth rate forecast, the forecasted value will be the same as the year-on-year growth rate for the current quarter. Formally, Theil's U2 for each variable for a four quarter ahead forecast, is calculated as:

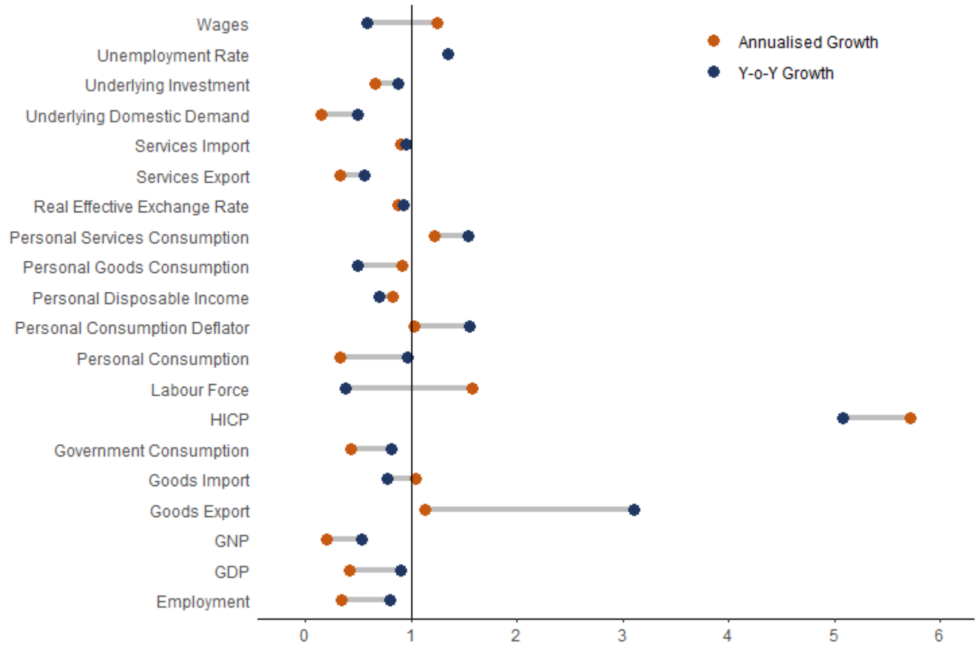
$$U2 = \sqrt{\frac{\frac{1}{H}\sum_{t=1}^{H}\left(\frac{\hat{y}_{t+4} - y_{t+4}}{y_t}\right)^2}{\frac{1}{H}\sum_{t=1}^{H}\left(\frac{y_t - y_{t+4}}{y_t}\right)^2}} \tag{8}$$

Similar to the Relative MSFE, Theil's U2, values below one indicates a superior forecast performance of the large Bayesian VAR to the naïve forecast for that variable.

A second metric used is the Root Mean Squared Forecast error. This is similar to that the Mean Squared Forecast error in equation 5, but the square root of this value is taken as the metric. The Root Mean Squared Forecast error is a measure of how concentrated the errors are around the actual value.

Finally, the performance of the forecast of each variable is compared to the forecast of an ARMA model for that variable. The ARMA model for each variable is selected using a stepwise algorithm of Hyndman & Khandakar (2008), based on the AIC for each model.[14] The relative performance for each variable of the large Bayesian VAR is assessed using the same formula as in equation 7 but with $MSFE_i^{(ARMA)}$ as the denominator.

Figure 3: Theil's U2 Q1 2010 to Q4 2018



Sources: Author's calculations.
Note: The Vertical black line marks the Theil's U2 value of one. Values below one indicates that the large Bayesian VAR has a superior forecasting performance to the naïve forecast. Corresponding values are presented in Appendix Tables D3 and D4. For the unemployment rate, the forecast four-quarter ahead change in the unemployment rate is assessed, not the growth rate.

Shown in Figure 3 are the Theil's U2 statistics for 20 of the main variables included in the model for Q1 2010 to Q4 2018. Theil's U2 is a non-linear metric in a similar vein as the relative MSFE in equation 7. The same caveat applies to interpreting this figure as it does to the two previous figures.

Overall, the large Bayesian VAR performs relatively well. The large Bayesian VAR performs better than the naïve forecast as measured by Theil's U2 for both the annualised growth rate and the year-on-year growth rate in 12 cases. The naïve forecast performs better in 5 cases (goods exports, HICP (Harmonised Index of Consumer prices), personal consumption deflator, personal services consumption, and the unemployment rate). A further three cases straddle the Theil's U2 value of one (wages, goods imports, and the labour force). In each of these three cases, the

[14] As the forecasts are evaluated using an expanding window, the ARMA model selection is carried out at each iteration and as a result the model selected for an individual variable may not necessarily be the same across the entire horizon evaluated.

large Bayesian VAR performs better than the naïve forecast at forecasting the year-on-year growth rate, and worse at forecasting the annualised growth rate.

Using the Diebold-Marino test to see if the forecasts from the large Bayesian VAR are statistically different from those of the naïve forecasts, the year-on-year forecast for government consumption was statistically different from the naïve forecast at the 5 per cent significance level. and the large Bayesian VAR's forecasts of goods imports and GNP were also statistically different from the naïve forecast at the 10 per cent significance level (Table D3). The naïve forecast performed better at forecasting the unemployment rate, and this was statistically different from the forecast from the large Bayesian VAR at the 5 per cent level. Other forecasts of the year-on-year growth rate from the large Bayesian VAR were not statistically different from the naïve forecast.

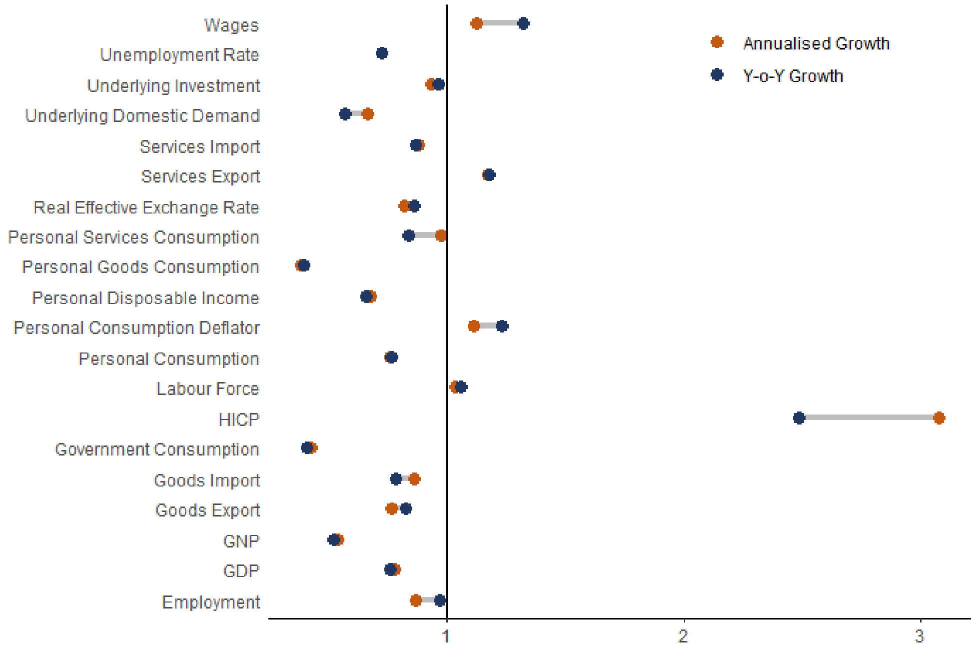
For the annualised growth rate, forecasts for personal consumption, personal goods consumption, government consumption, and GDP were statistically different from the naïve forecast at the 10 per cent significance level (Table D4). While forecast for services exports were significantly different at the 5 per cent level.

Figure 4 shows the plot of the relative MSFE between the large Bayesian VAR and the ARMA models for 20 variables, based on a sample from Q1 2010 to Q4 2018. In all but 5 cases, the large Bayesian VAR performs better than the ARMA models (wages, services export, personal consumption deflator, the labour force and HICP). For the labour force, services export and the personal consumption deflator, the relative MSFE is only marginally above one. Whereas, for the HICP forecast of the ARMA model perform twice as well as the forecast of the large Bayesian VAR.

In terms of whether the forecasts are statistically different from each other, the forecasts of UDD and GNP are statistically different from each other at the 1 per cent significance level for both the year-on-year growth and the annualised growth (Table D3 & Table D4). It is particularly encouraging that UDD—a key measure of the performance of the domestic economy—performs better than the ARMA model, and that the forecasts are statistically different. Finally, the year-on-year forecasts for personal disposable are statistically different from that of the ARMA models at the 10 per cent significance level (Table D3).

Looking at the RMSFE, the forecasts for the annualised growth rate of UDD, personal consumption, personal goods consumption, personal disposable income, employment, labour force, HICP and the personal consumption deflator have an RMSFE below 2 percentage points (Table D4).

Figure 4: Relative MSFE of the large Bayesian VAR versus ARMA models



Sources: Author's calculations.
Note: The Vertical black line marks the relative MSFE value of one. Values below one indicates that the large Bayesian VAR has a superior forecasting performance to the ARMA model for that variable. For the unemployment rate, the forecast four-quarter ahead change in the unemployment rate is assessed, not the growth rate. Corresponding values are presented in Appendix Tables D3 and D4.

4.4 Turning points and periods of turbulence

Evaluating the performance of forecasting models also warrants a discussion on the model's forecasting around turning points and during periods of extreme volatility. Some forecasting models can have a good performance over calm periods, i.e. periods not characterised by large swings in the economy, and some models may accurately predict turning points in macroeconomic variables. Relatively few models are able to do both. This section provides an overview of the large Bayesian VAR's performance in forecasting the macroeconomy around the crisis years from 2008-2010. [15]

[15] Real-time data has not been used for this exercise.

While Bayesian shrinkage mitigates to some extent the problems associated with having a small sample to estimate parameters from, it is by no means a panacea. Given the short sample from which to draw estimates from, forecasts are the start of the horizon should be treated with caution and may not provide an accurate indication of the Model's ability to forecast when there is more data available. For instance, forecasts of the growth rates for 2008 are based on only 32 quarters of data (2000 Q1 to 2007 Q4).

Figure 5 shows the forecast one-year ahead growth rate versus the corresponding actual year-end growth rate for several key variables relating to the domestic economy for 2005-2018. We take 2007-2008 as the first turning point and assess the model's performance in forecasting 2008 outturns, given data up to 2007. The forecasts errors for 2008 are quite large for some variables shown. For instance, the forecast growth for UDD is 5.6 percentage points off the actual growth rate. This is hardly surprising given that the model is only estimated with 32 quarters of data at this point. The forecast for personal services consumption is the most accurate at this point, only 0.9 percentage points from the actual growth rate. However, personal consumption growth did not experience as severe a drop off in growth as the other variables and was still a robust growth rate of 5.1 per cent for 2008. The forecast accuracy for 2009 improved for all variables, except for government consumption. In all cases the change in the growth rate between 2008 and 2009 was in the right direction.

For most of the variables selected here, growth rates were at their lowest in 2009. The model forecasts reflect that relatively well, with forecasted growth rates for 2010 picking up from the levels of growth in 2009. The model slightly overshoots in forecasting the pick–up in the level of the growth rate for 2010 for all variables, except for personal goods consumption. The forecast errors for 2010 are in general relatively small, given the short sample to that point and the volatility of actual data up over that period. In particular, the forecast errors for UDD, personal goods consumption, personal services consumption, and personal disposable income were all less than 0.6 percentage points in 2010.

Taking a more general view of the models forecasting ability over this period, the model performs relatively well, despite the turbulence in data at the earlier periods in the sample and given the short sample at the start of the period from which forecasts are derived. The average absolute error for all variables, except underlying investment, over this period are all below 2 percentage points. Perhaps unsurprisingly given the volatility of the series, the forecast errors for underlying investment are the largest. The average absolute error, in this case was 7.4 percentage points.

Table D5 shows how the large Bayesian VAR preforms relative to a naïve forecast and to ARMA forecast for this period. The large Bayesian VAR outperforms both the naïve forecast and the ARMA models in forecasting the year-end growth rates for 2005-2018 for each variable. The Root Mean Square Errors are, for the most part, relatively small, with the notable exception of underlying investment.

Figure 5: Forecast annual growth rates versus actual growth rates for 2005 to 2018

Source: Author's calculations.

Note: Figures show the forecast annual growth rate for the year *t+1*, using data up to year *t*, with the corresponding actual growth rate for year *t+1.*

5. Discussion of results

In general, the model has a good forecasting performance. In particular, for the key variables UDD and employment, the forecasting performance of the large Bayesian VAR is better than that of all other models/forecasts. In some cases, the forecasts for these variables are statically different than those of the other models. The performance of the large Bayesian VAR in forecasting wages is slightly more mixed. Some models perform better than the large Bayesian VAR in forecasting wages (for instance, the small VAR), however, the degree to which the other models outperform the large Bayesian VAR is low in this regard.

More broadly, the large Bayesian VAR has a relatively good performance at forecasting other variables of the underlying performance of the Irish economy. These variables are less affected by the distortionary effects that the globalisation activities of small number of large multinational firms have on more headline national accounts data (such as GDP, GNP, investment etc.) in Ireland. In particular the forecasts of UDD, employment, personal consumption, personal goods consumption, government consumption, underlying investment and personal disposable income are all relatively better than other models.

In terms of the headline variables, GDP, GNP, exports and imports, the large Bayesian VAR performs well. However, these variables are notoriously difficult to forecast for Ireland and while the large Bayesian VAR performs relatively well, the RMSFE can be quite large for these variables.

On the flipside, the large Bayesian VAR performs worse than other models at forecasting the price variables. Both the naïve forecasts and the ARMA forecasts are better at forecasting HICP and the personal consumption deflator. In some instances, the naïve forecast of HICP are five times better than the forecasts from the large Bayesian VAR. However, according to the Diebold-Marino tests, these forecasts are not statically different from those of the large Bayesian VAR. The

RMSFE are also relatively low for these variables. However, this partly reflects that fact that these variables are not vary volatile.

In terms of forecasting turning points, the model's performance in forecasting the downturn in 2008 was poor. However, this was based on only 32 quarters of data. Promisingly, it did however, forecast, to some degree, a pickup in the level of growth rates over 2010-2011. It is too early to draw definitive conclusions on the models ability to forecasting turning points and it remains to be seen whether the model's forecasting of downturns will improve as more data becomes available and more turning points—from which to assess the models ability to forecast these points—occur.

As a broader point, the inclusion of more variables may ultimately improve the forecasting performance of the model. However, as noted by Bańbura *et al.* (2010), the extent to which adding more variables improves the forecasting performance may face diminishing marginal returns. This may ultimately lead to a practical trade off in running the model.

The model is based on a VAR framework which requires a balanced dataset. The current model is based on 46 different time-series. Given the extent of the data requirement for the current model, this can potentially pose practical limitations on its uses. Should data series be discontinued, the modeler will have to determine whether an alternative data series should be used in its place, or should the series be dropped entirely. Other issues relating to data availability can affect the ability to run the current model. For instance, for privacy reasons or data licensing issues, data for particular variables for certain quarters may be suppressed by the data provider.[16] Running the model in this case may require the modeller to input a "best guess" estimate for the value of this variable in this particular quarter, given all other available data or alternatively, dropping the variable entirely, until such a time as the data becomes available. However, given priors are used, and Bayesian shrinkage around the distribution of the parameters are used, the effect of the error arising from the use of this best guess estimate will be reduced.

[16] For example, this was the case in the 2019 Q2 release of the Quarterly National Accounts by the CSO. Data on Machinery & Equipment, other Transport and Equipment, and Intangible Assets was suppressed by the CSO for confidentiality reasons.

A further consideration in the practical use of large Bayesian VAR models is that while they may be good at forecasting, the largely data drive nature of models such as these has its limitations. Due to their largely statistical and atheoretical nature, large Bayesian VAR models are largely a “black box” and it can be difficult to disentangle why a certain variable is displaying certain dynamics. Statistical models should often be used as a complement, alongside other models that have a greater foundation in economic theory.

6. Conclusion

This paper builds a dynamic model of the Irish macroeconomy using a large number of variables. Using quarterly data from Q1 2000 to Q4 2018, a Large Bayesian VAR is estimated for Ireland. Bayesian shrinkage is applied to the estimated parameters to overcome the “curse of dimensionality”. The Model is tailored for a small open economy and treats as exogenous the import demand of the rest of the world.

We assess the performance of the large Bayesian VAR in forecasting a wide set of variables. We show that the large Bayesian VAR can provide relatively good forecasts for a number of key variables of the domestic Irish economy. The model’s forecasting performance is assessed against the performance of other models, such as a small VAR, a Factor-Augmented VAR, ARMA models and against a naïve forecast.

Overall, the large Bayesian VAR performs favourably in forecasting most variables examined. In particular, the model performed well at forecasting variables that capture the underlying Irish macroeconomy, namely, UDD, employment, personal consumption, personal goods consumption, underlying investment and personal disposable income. However, there are some variables for which the model performs relatively worse than naïve models, ARMA models or smaller VARs. If anything, this highlights the need to use multiple models in forecasting the macroeconomy, and underscores the approach taken by Conroy & Casey (2017) in using a suite of model’s approach.

References

Bańbura, M., Giannone, D. and Reichlin, L., 2010. Large Bayesian vector auto regressions. *Journal of applied Econometrics*, *25*(1), pp.71-92.

Bergin, A., Conefrey, T., FitzGerald, J., Kearney, I., and Žnuderl, N. *The HERMES-13 macroeconomic model of the Irish economy*. No. 460. ESRI Working Paper, 2013.

Bergin, A., Conroy, N., Rodriguez, A.G., Holland, D., McInerney, N., Morgenroth, E.L. and Smith, D., 2017. COSMO: A new COre Structural MOdel for Ireland. *(No. 553) ESRI Working Paper.*

Bernanke, B.S., Boivin, J. and Eliasz, P., 2005. Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach. *The Quarterly journal of economics, 120(1), pp.387-422.*

Bloor, C. and Matheson, T., 2010. Analysing shock transmission in a data-rich environment: A large BVAR for New Zealand. *Empirical Economics*, *39*(2), pp.537-558.

Casey, E., 2019. Inside the "Upside Down": Estimating Ireland's Output Gap. *The Economic and Social Review*, *50*(1, Spring), pp.5-34.

Conroy, N. and Casey, E., 2017. Producing short-term forecasts of the Irish economy: a suite of models approach*. Irish Fiscal Advisory Council, Working Paper*, (4).

Carriero, A., Clark, T.E. and Marcellino, M., 2015. Bayesian VARs: specification choices and forecast accuracy. *Journal of Applied Econometrics*, *30*(1), pp.46-73.

Christiano, L.J., Eichenbaum, M. and Evans, C.L., 1999. Monetary policy shocks: What have we learned and to what end?. *Handbook of macroeconomics*, *1*, pp.65-148.

De Mol, C., Giannone, D. and Reichlin, L., 2008. Forecasting using a large number of predictors: Is Bayesian shrinkage a valid alternative to principal components?. *Journal of Econometrics*, *146*(2), pp.318-328.

De Menezes Barboza, R. and Vasconcelos, G.F., 2019. Measuring the aggregate effects of the Brazilian Development Bank on investment. *The North American Journal of Economics and Finance*, *47*, pp.223-236.

Diebold, F.X. and Mariano, R.S., 2002. Comparing predictive accuracy. *Journal of Business & economic statistics*, *20*(1), pp.134-144.

Domit, S., Monti, F. and Sokol, A., 2019. Forecasting the UK economy with a medium-scale Bayesian VAR. International Journal of Forecasting, 35(4), pp.1669-1678.

Gupta, R. and Kabundi, A., 2010. Forecasting macroeconomic variables in a small open economy: a comparison between small-and large-scale models. *Journal of Forecasting*, *29*(1-2), pp.168-185.

Hyndman, R.J. and Khandakar, Y., 2008. Automatic time series forecasting: The forecast Package for R. *Journal of Statistical Software*, *27*(3), pp.1-22.

Karlsson, S., 2013. Forecasting with Bayesian vector autoregression. In *Handbook of economic forecasting* (Vol. 2, pp. 791-897). Elsevier.

Kenny, G., Meyler, A. and Quinn, T., 1998. Bayesian VAR models for forecasting Irish inflation. *Central Bank of Ireland Technical Paper* 4/RT/98.

Koop, G.M., 2013. Forecasting with medium and large Bayesian VARs. *Journal of Applied Econometrics*, *28*(2), pp.177-203.

Litterman, R.B., 1986. Forecasting with Bayesian vector autoregressions—five years of experience. *Journal of Business & Economic Statistics*, *4*(1), pp.25-38.

Murphy, G., Nacheva, M. and L. Daly, (2018). Estimating Ireland's output gap; an analysis using selected statistical filters. December 2018, Dublin: Department of Finance.

Rusnák, M., Havranek, T. and Horváth, R., 2013. How to solve the price puzzle? A meta-analysis. *Journal of Money, Credit and Banking*, 45(1), pp.37-70.

Sims, C.A., 1992. Interpreting the macroeconomic time series facts: The effects of monetary policy. *European economic review*, *36*(5), pp.975-1000.

Song, S. and Bickel, P.J., 2011. Large vector auto regressions. *arXiv preprint arXiv:1106.3915*.

Appendix A: Description of Dataset

Table A2 gives a description of each variable used in the large Bayesian VAR and the source of the data. Transforms are given the following coding:

Table A.1: Transformation Key

Transform Code	Description
1	Differenced once
2	Seasonally adjusted then differenced once
3	Logged and then differenced
4	Seasonally adjusted then logged and differenced
5	Logged and then differenced twice

Note: some variables from the CSO have already been seasonally adjusted, in which case this has been indicated in the description column of table A2.

Table A.2: Description of variables

Description	Source	Transform
Industry Confidence Index	EC	1
Building Confidence Index	EC	1
ECB Deposit Rate	ECB	1
Private Credit	CSO	5
Benchmark 10-year Government Bond yield	R	1
3-Month Euribor rate	R	1
Housing Stock	CSO	5
Index of External Goods Demand	FC	4
Index of External Services Demand	FC	4

Table A.2 Continued

Description	Source	Transform
Real GDP seasonally adjusted	CSO	3
Real GNP seasonally adjusted	CSO	3
Real Personal Consumption seasonally adjusted	CSO	3
Real Government Consumption seasonally adjusted	CSO	3
Real Underlying investment	FC	3
Real Investment seasonally adjusted	CSO	3
Real Goods Export seasonally adjusted	CSO	3
Real Goods Import seasonally adjusted	CSO	3
Real Services Export seasonally adjusted	CSO	3
Real Services Import seasonally adjusted	CSO	3
Real Underlying Domestic Demand	FC	4
Real Personal Goods consumption	CSO	4
Real Personal Services Consumption	CSO	4
Unemployment rate	CSO	2
Participation rate	CSO	2
Real Machinery and Equipment seasonally adjusted	CSO	3
Real Intangibles seasonally adjusted	CSO	3
Real Building and construction Seasonally adjusted	CSO	3
Construction share of Employment	FC	1
Improvements	CSO	3
Labour Force	CSO	2
Employment	CSO	4
Real Transaction Costs	CSO	4
Real Other Building and Construction	CSO	4
Real Dwellings	CSO	4
Real Personal Disposable Income	CSO	4
Real Compensation of Employees	CSO	4
Real Effective Exchange Rate	FC	4
Personal Consumption Deflator	CSO	4
Real House Prices	CSO	3
Oil prices	R	3
Energy prices	ES	4
Harmonised Index of Consumer Prices	CSO	4
Consumer Confidence Index	EC	1
Services Confidence Index	EC	1
Retail Confidence Index	EC	1

Sources: CSO = Central Statistics office, ECB = European Central Bank, EC = European Commission, ES = Eurostat, FC = Irish Fiscal Advisory Council, R = Reuters.

Appendix B: Implementing the Minnesota prior

Following Bańbura *et al*. (2010) we rewrite the VAR system as a system of multivariate regressions to show how the Minnesota prior is implemented:

$$Y_{T\times N} = X_{T\times N}B_{K\times N} + U_{T\times N} \tag{9}$$

Where $Y = (Y_1, \dots, Y_T)'$, $X = (X_1, \dots, X_T)'$,with $X_t = (Y'_{t-1}, \dots, Y'_{t-1}, 1)'$, $U = (u_1, \dots, u_T)'$, and $B = (A_1, \dots, A_P, c)'$ is the $k \times n$ matrix containing all coefficients with $k = np + 1$. The normal inverted Wishart prior then takes the form:

$$vec(B)|\Psi\sim N(vec(B_0), \Psi\otimes\Omega_0) \quad \text{and} \quad \psi \sim iW(S_0, \alpha_0) \tag{10}$$

With the prior parameters, B_0, Ω_0, S_0 and α_0 chosen so that the moments coincide with those in equation 3 and the expectation of Ψ is equal to the fixed residual covariance matrix Σ of the Minnesota prior.

The prior is implemented by adding dummy observations to the system. Adding T_d dummy observations, Y_d and X_d to the system; Equation 9 is then equivalent to imposing the normal inverted Wishart prior with $B_0 = (X'_d X_d)^{-1} X'_d Y_d$, $\Omega_0 = (X'_d X_d)^{-1}$, $S_0 = (Y_d - X_d B_0)'(Y_d - X_d B_0)$ and $\alpha_0 = T_d - k$. The following dummy observations are included in order to retain the principles of the Minnesota prior;

$$Y_d = \begin{pmatrix} diag(\delta_1\sigma_1, \dots, \delta_1\sigma_n)/\lambda \\ 0_{n(p-1)\times n} \\ \cdots \\ diag(\sigma_1, \dots, \sigma_n) \\ \cdots \\ 0_{1\times n} \end{pmatrix}, \quad X_d = \begin{pmatrix} J_p\otimes diag(\sigma_1, \dots, \sigma_n)/\lambda & 0_{np\times 1} \\ \cdots & \cdots \\ 0_{n\times np} & 0_{n\times 1} \\ \cdots & \cdots \\ 0_{1\times np} & \varepsilon \end{pmatrix} \tag{11}$$

where $J_p = diag(1,2, \dots, p)$. The first block of dummies, i.e. Y_d, implements the prior belief on the autoregressive coefficients, the second block of dummies (first column in X_d), implements the prior on the covariance matrix and the third block implements the diffuse prior on the intercept (with ε a small number, which we arbitrarily set to 0.1).[17] In principle, parameters should be set using only prior knowledge, but Bańbura *et al*. (2010) follow Litterman (1986) and set the scale of the

[17] This is also the approach taken by de Menezes Barboza & Vasconcelos (2019).

parameter σ_i^2 equal to the variance of the residuals of a univariate AR(p) model for the variable i. We differ from this approach, and use the approach taken by de Menezes Barboza & Vasconcelos (2019) and simply use σ_i^2 the variance of the *ith* variable. In circumstances where non-stationary time series are modelled, this may lead to very high variances, but this would only make the priors less informative for these variables.[18] Using the dummies above, we now have an augmented system for equation 9:

$$Y^*_{T^*\times N} = X^*_{T^*\times N} B_{K\times N} + U^*_{T^*\times N} \tag{12}$$

where $T^* = T + T_d$, $Y^* = (Y', Y_d')'$, $X^* = (X', X_d')'$ and $U^* = (U', U_d')'$. The posterior is then of the form:

$$vec(B)|\Psi, \mathrm{Y} \sim N(vec(\hat{B}), \Psi \otimes (X^{*'} X^*)^{-1}) \quad \text{and} \quad \psi|Y \sim iW(\hat{\Sigma}, T + T_d + 2 + k) \tag{13}$$

which has coefficients that coincide with those estimated by OLS regression of Y^* and X^*.

To add exogenous variables to the model, given the system in equation 12, and foreign variables X_f, X^* becomes $X^* = (X', X_d', X_f')'$ with the size of the matrix B adjusted accordingly for the dimension of X_f.

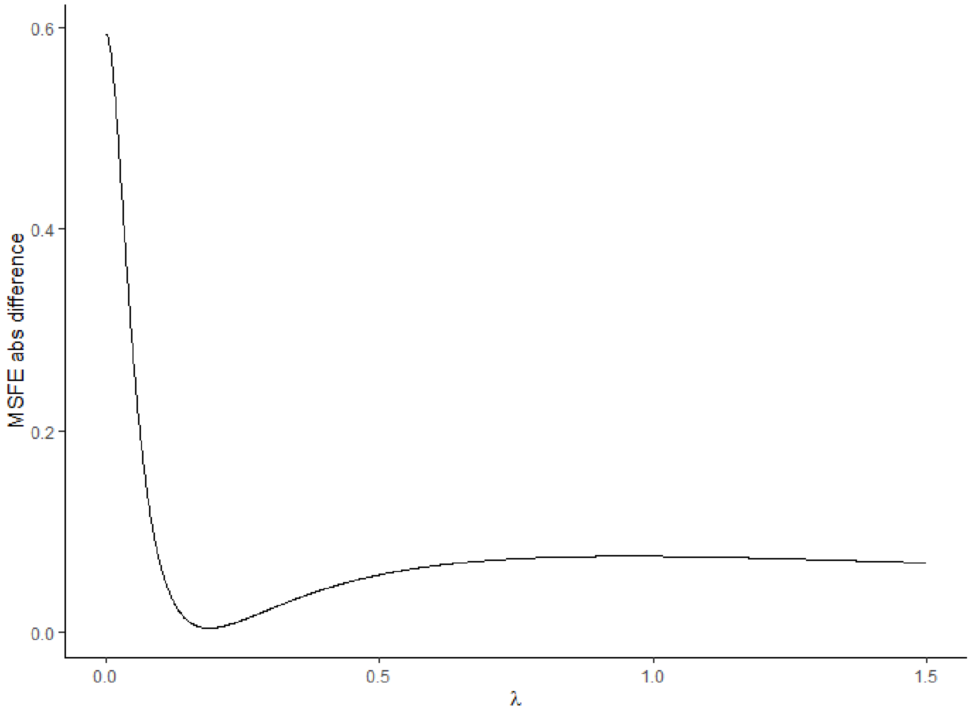
In terms of forecasting, obtaining the one-step-ahead predictive distribution is computationally easy as the natural conjugate prior has an analytical form, so no simulation is required. However, when forecasting more than one period ahead, an analytical formula does not exist. In this case, we use the direct forecasting method in order to avoid the computationally intensive predictive simulation that would be required otherwise. This approach does not systematically underperform (or indeed, overperform) other methods (see Koop, 2013).

[18] In any event, this is not an issue here, as all variables are transformed to stationarity prior to modelling.

Appendix C: Calibration of λ

The calibration was carried out using the procedure outlined in Section 3 with data from Q1 2000 to Q4 2009. The one-step ahead forecasting procedure was only carried out over the horizon of Q1 2005 to Q4 2009. This was to allow the small VAR to be estimated, and provide an additional window of data so as to reduce the possibility of over fitting the small VAR. The lag length for the small VAR was selected by analysing various metrics and comparing forecasting performances of the small VAR with different lag lengths. The lag length of one provided the best out-of-sample forecasting performance and was the lag length selected based on the Schwarz Criterion and the Hanna-Quinn Criterion. The simple average of the index for external goods demanded and services demanded was used instead of including both indices in order to further reduce the problem of overfitting for the small VAR.

Figure C.1: Calibration of λ



Sources: Author's calculations.
Note: The global minimum of the absolute difference of the Mean Squared Forecast Error between the Small VAR and the large Bayesian VAR for the key variables was achieved with $\lambda = 0.188$.

Appendix D: Forecast performance evaluation

Table D1: Forecasting performance relative to the small VAR for Q1 2010 to Q4 2018

Variable	Y-o-Y	Annualised Growth
Underlying Domestic Demand	0.47	0.49
Wage	1.02	1.01
Employment	0.54	0.55*

Sources: Author's calculations.
Note: This table corresponds to values presented in Figure 1. Values for the ARMA column here correspond to the relative MSFE in equation 12, where Model = small VAR. Values below one indicates superior forecasting performance relative to the small VAR. The Diebold-Mariano test was carried out to determine whether the forecasts for the large Bayesian VAR were significantly different from those of the small VAR. Values with *, ** or *** correspond to forecasts that are significantly different from each other at the 10%, 5% and 1% significance levels respectively. However, caution should be taken when interpreting these significance values, given the sample that there are only 36 observations from which to carry out the test.

Table D2: Forecasting performance relative to the Factor-Augmented VAR for Q1 2010 to Q4 2018

Variable	Y-o-Y	Annualised Growth
Underlying Domestic Demand	0.12	0.17
Wage	0.42	0.57
Employment	0.13	0.18

Sources: Author's calculations.
Note: This table corresponds to values presented in Figure 2. Values presented here correspond to the relative MSFE in equation 12, where Model = FAVAR. Values below one indicates superior forecasting performance relative to the FAVAR. The Diebold-Mariano test was carried out to determine whether the forecasts for the large Bayesian VAR were significantly different from those of the Factor Augmented VAR. Values with *, ** or *** correspond to forecasts that are significantly different from each other at the 10%, 5% and 1% significance levels respectively. However, caution should be taken when interpreting these significance values, given the sample that there are only 36 observations from which to carry out the test.

Table D3: Four-quarter ahead year-on-year forecasting performance for Q1 2010 to Q4 2018

Variable	Theil's U2	RMSFE	ARMA
Underlying Domestic Demand	0.49	0.022	0.57***
Wages	0.59	0.036	1.32
Unemployment Rate	1.35**	1.420	0.73
Personal Consumption	0.97	0.023	0.77
Personal Goods Consumption	0.49	0.026	0.40
Personal Services Consumption	1.54	0.031	0.84
Government Consumption	0.82**	0.033	0.41
Underlying Investment	0.88	0.135	0.97
Personal Disposable Income	0.70	0.031	0.66*
Employment	0.80	0.020	0.97
Labour Force	0.37	0.016	1.06
Gross Domestic Product	0.90	0.065	0.76
Gross National Product	0.53*	0.052	0.52***
Real Effective Exchange Rate	0.92	0.100	0.87
Goods Import	0.77*	0.094	0.79
Services Import	0.95	0.179	0.87
Goods Export	3.10	0.161	0.83
Services Export	0.56	0.064	1.18
HICP	5.07	0.017	2.48
Personal Consumption Deflator	1.55	0.021	1.23

Sources: Author's calculations

Note: This table corresponds to values presented in Figure 3. For the unemployment rate, the forecast four-quarter ahead change in the unemployment rate is assessed, not the growth rate. Values for the ARMA column here correspond to the relative MSFE in equation 12, where Model = ARMA. Values for Theil's U2 and ARMA below 1 indicate that the large Bayesian VAR has a superior forecasting performance to the that of the naïve forecast and the ARMA forecast respectively. The Diebold-Mariano test was carried out to determine whether the forecasts for the large Bayesian VAR were significantly different from those of the naïve forecast (Theil's U2 column) and the ARMA forecasts. Values with *, ** or *** correspond to forecasts that are significantly different from each other at the 10%, 5% and 1% significance levels respectively. However, caution should be taken when interpreting these significance values, given the sample that there are only 36 observations from which to carry out the test.

Table D4: Four-quarter ahead annualised growth forecasting performance for 2010 to 2018

Variable	Theil's U2	RMSFE	ARMA
Underlying Domestic Demand	0.15	0.013	0.66***
Wages	1.25	0.021	1.13
Unemployment Rate	1.35**	1.420	0.73
Personal Consumption	0.32*	0.015	0.76
Personal Goods Consumption	0.92*	0.016	0.39
Personal Services Consumption	1.22	0.020	0.98
Government Consumption	0.42*	0.023	0.43
Underlying Investment	0.66	0.084	0.94
Personal Disposable Income	0.82	0.019	0.68
Employment	0.34	0.012	0.87
Labour Force	1.58	0.010	1.04
Gross Domestic Product	0.41*	0.045	0.78
Gross National Product	0.20	0.036	0.54***
Real Effective Exchange Rate	0.87	0.064	0.82
Goods Import	1.03	0.063	0.87
Services Import	0.90	0.116	0.88
Goods Export	1.13	0.108	0.77
Services Export	0.33**	0.043	1.17
HICP	5.72	0.010	3.07
Personal Consumption Deflator	1.02	0.012	1.11

Sources: Author's calculations.
Note: This table corresponds to values presented in Figure 4. For the unemployment rate, the forecast four-quarter ahead change in the unemployment rate is assessed, not the growth rate. Values for the ARMA column here correspond to the relative MSFE in equation 12, where Model = ARMA. Values for Theil's U2 and ARMA below 1 indicate that the large Bayesian VAR has a superior forecasting performance to the that of the naïve forecast and the ARMA forecast respectively. The Diebold-Mariano test was carried out to determine whether the forecasts for the large Bayesian VAR were significantly different from those of the naïve forecast (Theil's U2 column) and the ARMA forecasts. Values with *, ** or *** correspond to forecasts that are significantly different from each other at the 10%, 5% and 1% significance levels respectively. However, caution should be taken when interpreting these significance values, given the sample that there are only 36 observations from which to carry out the test.

Table D5: 1-Year ahead forecasting performance for 2005 to 2018

Variable	Theil's U2	RMSFE	ARMA
Underlying Domestic Demand	0.33	0.019	0.92
Wages	0.48	0.023	0.47
Personal Consumption	0.18	0.019	0.52
Personal Goods Consumption	0.49	0.019	0.58
Personal Services Consumption	0.66	0.013	0.62
Government Consumption	0.58	0.024	0.52
Underlying Investment	0.72	0.087	0.87
Personal Disposable Income	0.62	0.017	0.68
Employment	0.28	0.012	0.80

Sources: Author's calculations.
Note: This table corresponds to values presented in Figure 5. Values for the ARMA column here correspond to the relative MSFE in equation 12, where Model = ARMA. Values correspond to forecasts shown in Figure 5 for the annual growth rates from 2005 to 2018. Values for Theil's U2 and ARMA below 1 indicate that the large Bayesian VAR has a superior forecasting performance to the that of the naïve forecast and the ARMA forecast respectively. The Diebold-Mariano test was not carried out on these values, given that short sample size (only 14 observations)