University of Pretoria
Department of Economics Working Paper Series

Forecasting Macroeconomic Variables Using Large Datasets: Dynamic Factor Model versus Large-Scale BVARs

Rangan Gupta
University of Pretoria
Alain Kabundi
University of Johannseburg

Working Paper: 2008-16
June 2008

Department of Economics
University of Pretoria
0002, Pretoria
South Africa
Tel: +27 12 420 2413
Fax: +27 12 362 5207

Forecasting Macroeconomic Variables Using Large Datasets: Dynamic Factor Model versus Large-Scale BVARs

Rangan Gupta* and Alain Kabundi#

Abstract

This paper uses two-types of large-scale models, namely the Dynamic Factor Model (DFM) and Bayesian Vector Autoregressive (BVAR) Models based on alternative hyperparameters specifying the prior, which accommodates 267 macroeconomic time series, to forecast key macroeconomic variables of a small open economy. Using South Africa as a case study and per capita growth rate, inflation rate, and the short-term nominal interest rate as our variables of interest, we estimate the two-types of models over the period 1980Q1 to 2006Q4, and forecast one- to four-quarters-ahead over the 24-quarters out-of-sample horizon of 2001Q1 to 2006Q4. The forecast performances of the two large-scale models are compared with each other, and also with an unrestricted three-variable Vector Autoregressive (VAR) and BVAR models, with identical hyperparameter values as the large-scale BVARs. The results, based on the average Root Mean Squared Errors (RMSEs), indicate that the large-scale models are better-suited for forecasting the three macroeconomic variables of our choice, and amongst the two types of large-scale models, the DFM holds the edge.

Journal of Economic Literature Classification: C11, C13, C33, C53.
Keywords: Dynamic Factor Model, BVAR, Forecast Accuracy.

*Associate Professor, University of Pretoria, Department of Economics, Pretoria, 0002, South Africa, Email: Rangan.Gupta@up.ac.za. Phone: +27 12 420 3460, Fax: +27 12 362 5207.

#To whom correspondence should be addressed. Senior Lecturer, University of Johannesburg, Department of Economics, Johannesburg, 2006, South Africa, Email: akabundi@uj.ac.za. Phone: +27 11 559 2061, Fax: +27 11 559 3039.

1. Introduction

This paper exploits information contained in a large cross-section of time series, 267 to be specific, to forecast three key macroeconomic variables, namely, per capita growth rate, the Consumer Price Index (CPI) inflation, and the 91 days Treasury Bill rate for the South African economy, using a Dynamic Factor Model (DFM) and alternative Bayesian Vector Autoregressive (BVAR) models, based on alternative values of the hyperparameters specifying the prior. The two types of model are first estimated over the period of 1980:01 to 2000:04 using quarterly data, and are then used to generate one- to four-quarters-ahead out-of-sample forecasts over a 24 quarter forecasting horizon of 2001:01 to 2006:04. The performance of these two large-scale models are also compared with an unrestricted Vector Autoregressive (VAR) model and BVAR models, with identical hyperparameter values as the large-scale BVARs, but including only the three variables we are concerned of. At this stage, it must be emphasized that the choice of South Africa, as our country of interest, emerges purely from the motivation of this study discussed below, and, hence, there is no reason that the current work cannot be conducted for any other economy, especially given the general nature of the econometric models we use here.

The main motivation for this current piece of work emanates from two recent studies carried out for the South African economy by Gupta and Kabundi (2008) and Das *et al.* (2008). Gupta and Kabundi (2008) used a DFM to forecast growth per capita, inflation based on Gross Domestic Product deflator and the 91 days Treasury Bill rate. When the forecast performance of the model was compared with an unrestricted VAR, alternative BVARs and a New Keynesian Dynamic Stochastic General Equilibrium (DSGE) model, estimated using the three variables of interest, the authors found the DFM to outperform all the models in terms of forecasting the interest rate, while it did no worse than the VAR and the BVARs in forecasting the other two variables. On the other hand, Das *et al.* (2008), forecasted regional house price inflation in five major metropolitan areas of South Africa based on the same DFM developed by Gupta and Kabundi (2008), which now also included house price inflation of different house size category[2] for the five metropolitans under consideration. When the authors compared the forecasting performance of the DFM with spatial and non-spatial BVARs, besides, an unrestricted VAR, based on only the house price inflation, the DFM was found to outperform the other models in 10 of the 15 cases. Both Gupta and Kabundi (2008) and Das *et al.* (2008) attribute the better performance of the DFM to its ability to efficiently handle large amounts of information, and, hence, its capability to forecast more accurately.

In such a backdrop, this paper, using the same panel of 267 time series as used by Gupta and Kabundi (2008), tries to check for the validity of such a claim, by comparing the forecasting performance of a DFM, for three variables, with that of BVAR models, which based on their estimation method[3] can also accommodate a panel as large as the one used in the DFM. Our analysis is quite similar in spirit, but markedly different in structure, to the work of De Mol *et al.* (2006). This paper compared the forecasting performance of BVAR models, based on normal and double exponential priors, with that of a DFM model for the US economy. The authors show that the forecasts generated from the BVAR models based on 131 variables for industrial production and CPI perform equally well as that of a DFM.[4] Though our study, when compared to De Mol *et al.* (2006), is limited in the sense it only considers normal priors, namely the Minnesota prior, we by allowing for different degree of interaction amongst domestic and foreign variables[5], contained in our data, account for the small open economy structure of South Africa to play a role in the forecasting results, and, hence, allow for a bit of theory, in these otherwise

[2] The houses were categorized as small, medium and large based on the square metres of area covered. See Burger and van Rensburg (2008), Gupta and Das (2008) and Das et al. (2008) for further details regarding the housing market in South Africa.

[3] See Section 4 laying out the basics of the BVAR models for further details.

[4] Note De Mol et al. (2007) also looked into conditions for consistency of the Bayesian forecasts as the cross-section and the sample size became large.

[5] See Section 3 containing the discussion of data for further details.

atheoretical models.[6] To our knowledge, this is the first attempt to use large-scale BVAR models, as an alternative to a DFM, to forecast key macroeconomic variables in an emerging market economy.

In general, the motivation to use a large data set to forecast an economy originates not only from the fact that such data is now available at lower cost, but also because, and, perhaps, more importantly, the increased power of computation has facilitated in using such huge amount of information to estimate and forecast with econometric models. Where tradition forecasting models such VAR, known for their better predictability, are unable to handle information contained in a large dataset without facing the degrees of freedom problems, dynamic factor models (DFM) on the other hand can efficiently handle large amounts of information and therefore help improve the forecasting performance of econometric models. VAR that uses information contained in few fundamental variables seems limited and insufficient to mimic complex economic relations and forecast the future. Today modern econometricians have the ability to extract important information from a large dataset and accurately forecast the future. In addition, central bankers, policymakers, and academics agree that economic agents monitor hundreds of economic variables in their decision-making process (Bernanke and Boivin, 2003). As Stock and Watson (2005) agreeably put it, the DFM transforms the *curse of dimensionality* into *blessing of dimensionality*.

The original dynamic factor models of Sargent and Sims (1977), Geweke (1977), Chamberlain (1983) and Chamberlain, and Rothschild (1983) have been improved recently through advances in estimation techniques proposed by Stock and Watson (2002), Forni et al. (2005), and Kapetanios and Marcellino (2004). The success of DFM is due to the fact that only few extracted common factors can explain to a large extent the variation of variables in a large cross-section of time series. Several empirical researches provide evidence of improvement in forecasting performance of macroeconomic variables using factor analysis (Giannone and Matheson, 2007; Van Nieuwenhuyze, 2007; Cristadoro et al., 2005; Forni et al., 2005; Schneider and Spitzer, 2004, Kabundi, 2004; Forni et al., 2001; and Stock and Waston, 2002a, 2002b, 1999, 1991, and 1989).

Unlike the unrestricted VAR, the Bayesian VAR (BVAR) can be seen as a valid alternative to the DFM as it can equally accommodate large number of predictors without facing the risk of losing degrees of freedom associated with unrestricted VAR. By imposing restrictions related to the distribution of coefficients, BVAR models avoid the overparametrization and overfitting of the model that are inherent in a VAR framework. Hence, a BVAR is based on fast exploration of the model space and, as part of the Bayesian methodology, it does not need to rely on asymptotic theory as in the case of an unrestricted VAR. Furthermore, given their estimation procedures, the DFM and the BVAR are capable of forecasting simultaneously a large number of time series other than the key variables of interest. In other words, the two models are ideal in using large amount of information to forecast the economy, and, hence, could be considered to be on an even ground.

The remainder of the paper is organized as follows: The following section briefly discusses the DFM. Section 3 outlines the basics of the VAR and Minnesota-type BVARs. Section 4 discusses the data used to estimate the DFM and BVARs, while Section 5 presents the results from the forecasting exercise. Finally, section 6 concludes.

2. The Basics of a DFM

This study uses the Dynamic Factor Model (DFM) developed by Forni et al. (2005) to extract common components between macroeconomics series, and then these common components are used to forecast the three key macroeconomic variables. In the VAR models, since all variables are used in forecasting, the number of parameters to estimate depend on the number of variables n. With such a large information set, n, the estimation of a large number of parameters leads to a curse of dimensionality. The DFM uses information set accounted by few factors $q << n$, which transforms the curse of dimensionality into a blessing of dimensionality.

The DFM expresses individual times series as the sum of two unobserved components: a common component driven by a small number of common factors and an idiosyncratic

[6] See Section 4 for further details.

component, which are specific to each variable. The relevance of the method is that the DFM is able to extract the few factors that explain the comovement of all South African macroeconomic variables. Forni et al. (2005) demonstrated that when the number of factors is small relative to the number of variables and the panel is heterogeneous, the factors can be recovered from the present and past observations.

Consider an $n \times 1$ covariance stationary process $Y_t = (y_{1t}, ..., y_{nt})'$. Suppose that X_t is the standardized version of Y_t, i.e. X_t has a mean zero and a variance equal to one. Under DFM proposed by Forni et al. (2005) X_t is described by a factor model, it can be written as the sum of two orthogonal components:

$$x_{it} = b_i(L) f_t + \xi_{it} = \lambda_i F_t + \xi_{it} = \chi_{it} + \xi_{it} \quad (1)$$

or, in vector notation:

$$X_t = B(L) f_t + \xi_{it} = \Lambda F_t + \xi_{it} = \chi_{it} + \xi_{it} \quad (2)$$

where f_t is a $q \times 1$ vector of dynamic factors, $B(L) = B_0 + B_1 L + ... + B_s L^s$ is an $n \times q$ matrix of factor loadings of order s, ξ_{it} is an $n \times 1$ vector of idiosyncratic components, F_t is $r \times 1$ vector of factors, with $r = q(s+1)$. However, in more general framework $r \geq q$, instead of the more restrictive $r = q(s+1)$. In a DFM, f_t and ξ_{it} are mutually orthogonal stationary process, while χ_{it} is the common component.

In factor analysis jargon, $X_t = B(L) f_t + \xi_{it}$ is referred to as dynamic factor model, and $X_t = \Lambda F_t + \xi_{it}$ is the static factor model. Similarly, f_t is regarded as vector of dynamic factors while F_t is the vector of static factors. Since dynamic common factors are latent, they need to be estimated. Forni et al. (2005) estimate dynamic factors through the use of dynamic principal component analysis. It involves the estimating the eigenvalues and eigenvactors decomposition of spectral density matrix of X_t, which is a generalization of orthogonalization process in case of static principal components.[7]

3. Alternative Forecasting Models: VAR and BVAR[8]

The Vector Autoregressive (VAR) model, though 'atheoretical', is particularly useful for forecasting purposes. A VAR model can be visualized as an approximation of the reduced-form simultaneous equation structural model.

An unrestricted VAR model, as suggested by Sims (1980), can be written as follows:

$$y_t = A_0 + A(L) y_t + \varepsilon_t \quad (3)$$

where y is a $(n \times 1)$ vector of variables being forecasted; $A(L)$ is a $(n \times n)$ polynomial matrix in the backshift operator L with lag length p, *i.e.,* $A(L) = A_1 L + A_2 L^2 + ... + A_p L^p$; A_0 is a $(n \times 1)$ vector of constant terms, and ε is a $(n \times 1)$ vector of error terms. In our case, we assume that $\varepsilon \sim N(0, \sigma^2 I_n)$, where I_n is a $n \times n$ identity matrix.

Note the VAR model, generally uses equal lag length for all the variables of the model. One drawback of VAR models is that many parameters need to be estimated, some of which may be insignificant. This problem of overparameterization, resulting in multicollinearity and a loss of degrees of freedom, leads to inefficient estimates and possibly large out-of-sample forecasting errors. One solution, often adapted, is simply to exclude the insignificant lags based on statistical

[7] See Gupta and Kabundi (2008) for a detailed description of the model.

[8] This section relies heavily on the discussion available on VAR and BVAR in Dua and Ray (1995), LeSage (1999), Gupta and Sichei (2006), Gupta (2006, 2007a,b) and Gupta and Das (2008).

tests. Another approach is to use a near VAR, which specifies an unequal number of lags for the different equations.

However, an alternative approach to overcoming this overparameterization, as described in Litterman (1981), Doan *et al.* (1984), Todd (1984), Litterman (1986), and Spencer (1993), is to use a BVAR model. Instead of eliminating longer lags, the Bayesian method imposes restrictions on these coefficients by assuming that they are more likely to be near zero than the coefficients on shorter lags. However, if there are strong effects from less important variables, the data can override this assumption. The restrictions are imposed by specifying normal prior distributions with zero means and small standard deviations for all coefficients with the standard deviation decreasing as the lags increase. The exception to this is that the coefficient on the first own lag of a variable has a mean of unity. Litterman (1981) used a diffuse prior for the constant. This is popularly referred to as the 'Minnesota prior' due to its development at the University of Minnesota and the Federal Reserve Bank at Minneapolis.

Formally, as discussed above, the means and variances of the Minnesota prior take the following form:

$$\beta_i \sim N(1, \sigma_{\beta_i}^2) \text{ and } \beta_j \sim N(0, \sigma_{\beta_j}^2) \tag{4}$$

where β_i denotes the coefficients associated with the lagged dependent variables in each equation of the VAR, while β_j represents any other coefficient. In the belief that lagged dependent variables are important explanatory variables, the prior means corresponding to them are set to unity. However, for all the other coefficients, β_j's, in a particular equation of the VAR, a prior mean of zero is assigned to suggest that these variables are less important to the model.

The prior variances $\sigma_{\beta_i}^2$ and $\sigma_{\beta_j}^2$, specify uncertainty about the prior means $\overline{\beta}_i$ = 1, and $\overline{\beta}_j$ = 0, respectively. Because of the overparameterization of the VAR, Doan *et al.* (1984) suggested a formula to generate standard deviations as a function of small numbers of hyperparameters: *w, d*, and a weighting matrix *f(i, j)*. This approach allows the forecaster to specify individual prior variances for a large number of coefficients based on only a few hyperparameters. The specification of the standard deviation of the distribution of the prior imposed on variable *j* in equation *i* at lag *m*, for all *i, j* and *m*, defined as $S_1(i, j, m)$, can be specified as follows:

$$S_1(i, j, m) = [w \times g(m) \times f(i, j)] \frac{\hat{\sigma}_j}{\hat{\sigma}_i} \tag{5}$$

with $f(i, j) = 1$, if $i = j$ and k_{ij} otherwise, with $(0 \le k_{ij} \le 1)$, $g(m) = m^{-d}, d > 0$. Note that $\hat{\sigma}_i$ is the estimated standard error of the univariate autoregression for variable *i*. The ratio $\hat{\sigma}_i / \hat{\sigma}_j$ scales the variables to account for differences in the units of measurement and, hence, causes specification of the prior without consideration of the magnitudes of the variables. The term *w* indicates the overall tightness and is also the standard deviation on the first own lag, with the prior getting tighter as we reduce the value. The parameter *g(m)* measures the tightness on lag *m* with respect to lag 1, and is assumed to have a harmonic shape with a decay factor of *d*, which tightens the prior on increasing lags. The parameter *f(i, j)* represents the tightness of variable *j* in equation *i* relative to variable *i*, and by increasing the interaction, i.e., the value of k_{ij}, we can loosen the prior.[9] Note, in the standard Minnesota-type prior, the overall tightness (*w*) takes the values of 0.1, 0.2 and 0.3, while, the lag decay (*d*) is generally chosen to be equal to 0.5, 1.0 and 2.0. The interaction parameter (k_{ij}) is traditionally set at = 0.5. We will call the BVARs estimated with this set of parameterization of the priors as symmetric BVARs.

Given that, we have domestic as well as foreign and world variables in the DFM, and realizing

[9] For an illustration, see Dua and Ray (1995).

the South Africa is a small open economy, and, hence, domestic variables would have minimal, if any, effect on foreign and world variables, while the latter set of variables is sure to have an influence on the South African variables, setting k_{ij} = 0.5 could be a quite far fetched from reality. Hence, borrowing from the BVAR models used for regional forecasting, involving both regional and national variables, and following Kinal and Ratner (1986) and Shoesmith (1992), the weight of a foreign or world variable in a foreign or world equation, as well as a domestic equation, is set at 0.6. The weight of a domestic variable in other domestic equation is fixed at 0.1 and that in a foreign or world equation at 0.01. Finally, the weight of the domestic variable in its own equation is 1.0. These weights are in line with Litterman's circle-star structure. Star (foreign or world) variables affect both star and circle (domestic) variables, while circle variables primarily influence only other circle variables.[10] We will call the BVARs estimated with this set of parameterization of the priors as asymmetric BVARs.

Finally, once the priors have been specified, the alternative BVARs, whether based on the 3 variables or all the 267 variables (symmetric or asymmetric), are estimated using Theil's (1971) mixed estimation technique. Specifically, suppose we denote a single equation of the VAR model as: $y_1 = X\beta + \varepsilon_1$, with $Var(\varepsilon_1) = \sigma^2 I$, then the stochastic prior restrictions for this single equation can be written as:

$$\begin{bmatrix} M_{111} \\ M_{112} \\ . \\ . \\ . \\ M_{nnm} \end{bmatrix} = \begin{bmatrix} \sigma/\sigma_{111} & 0 & . & . & . & 0 \\ 0 & \sigma/\sigma_{112} & 0 & . & . & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & . & . & . & . & 0 \\ 0 & 0 & . & . & 0 & \sigma/\sigma_{nnm} \end{bmatrix} \begin{bmatrix} a_{111} \\ a_{112} \\ . \\ . \\ . \\ a_{nnm} \end{bmatrix} + \begin{bmatrix} u_{111} \\ u_{112} \\ . \\ . \\ . \\ u_{nnm} \end{bmatrix} \tag{6}$$

Note, $Var(u) = \sigma^2 I$ and the prior means M_{ijm} and σ_{ijm} take the forms shown in (4) and (5). With (6) written as:

$$r = R\beta + u \tag{7}$$

and the estimates for a typical equation are derived as follows:

$$\hat{\beta} = (X'X + R'R)^{-1}(X'y_1 + R'r) \tag{8}$$

Essentially then, the method involves supplementing the data with prior information on the distribution of the coefficients. The number of observations and degrees of freedom are increased by one in an artificial way, for each restriction imposed on the parameter estimates. The loss of degrees of freedom due to over- parameterization associated with a classical VAR model is, therefore, not a concern in the BVARs.

4. Data

It is imperative in factor analysis framework to extract common components from a data rich environment. After extracting common components of per capita growth rate, inflation, and nominal interest rates of South Africa, we make out-of-sample forecast for one, two, three, and four quarters ahead.

The data set contains 267 quarterly series of South Africa, ranging from real, nominal, and financial sectors. We also have intangible variables, such as confidence indices, and survey variables. In addition to national variables, the paper uses a set of global variables such as commodity industrial inputs price index and crude oil prices. The data also comprises series of major trading partners such as Germany, the United Kingdom (UK), and the United States (US)

[10] We also experimented by assigning higher and lower interaction values, in comparison to those specified above, to the star variables in both the star and circle equations, but, the rank ordering of the alternative forecasts remained the same.

of America. The in-sample period contains data from 1980Q1 to 2000Q4. All series are seasonally adjusted and covariance stationary. The more powerful DFGLS test of Elliott, Rothenberg, and Stock (1996), instead of the most popular ADF test, is used to assess the degree of integration of all series. All nonstationary series are made stationary through differencing. The Schwarz information criterion is used in the selecting the appropriate lag length in such a way that no serial correction is left in the stochastic error term. Where there were doubts about the presence of unit root, the KPSS test proposed by Kwiatowski, Phillips, Schmidt, and Shin (1992), with the null hypothesis of stationarity, was applied. All series are standardized to have a mean of zero and a constant variance. It must, however be pointed out that, non-stationarity is not an issue with the BVAR, since Sims *et al.* (1990) indicates that with the Bayesian approach entirely based on the likelihood function, the associated inference does not need to take special account of nonstationarity, since the likelihood function has the same Gaussian shape regardless of the presence of nonstationarity. Hence, for the sake of comparison amongst the VARs, both classical and Bayesian, we make no attempt o make the variables stationary, unlike in the DFM.[11] The in-sample period contains data from 1980Q1 to 2000Q4, while the out-of-sample set is 2001Q1-2006Q4.[12]

There are various statistical approaches in determining the number of factors in the DFM. For example, Bai and Ng (2002) developed some criteria guiding the selection of the number of factors in large dimensional panels. The principal component analysis (PCA) can also be used in establishing the number of factors in the DFM. The PCA suggests that the selection of a number of factors q be based on the first eigenvalues of the spectral density matrix of X_t. Then, the principal components are added until the increase in the explained variance is less than a specific $\alpha = 0.05$. The Bai and Ng (2002) approach proposes five static factors, while Bai and Ng (2007) suggests two primitive or dynamic factors. Similar to the latter method, the principal component technique, as proposed by Forni et al. (2000) suggests two dynamic factors. The first two dynamic principal components explain approximately 99 percent of variation, while the eigenvalue of the third component is $0.005 < 0.05$.

5. Evaluation of Forecast Accuracy

Given the specification of the priors above, we estimate a VAR and the small- and large-scale BVARs (both symmetric and asymmetric) over the period of 1980:01 to 2000:04, based on quarterly data. Then we compute the out-of-sample one- through four-quarters-ahead forecasts for the period of 2001:Q1 to 2006:Q4, and compare the forecast accuracy relative to that of the forecasts generated the benchmark DFM model. The different types of the VARs are estimated with 5 lags[13] of each variable. The VAR and the BVARs, for an initial prior, are estimated for the period of 1980:Q1 to 2000:Q4 and, then, we forecast from 2001:Q1 through to 2006:Q4. Since we use five lags, the initial five quarters of the sample, 1980:Q1 to 1981:Q4, are used to feed the lags. We generate dynamic forecasts, as would naturally be achieved in actual forecasting practice. The models are re-estimated each quarter over the out-of-sample forecast horizon in order to update the estimate of the coefficients, before producing the 4-quarters-ahead forecasts. This iterative estimation and 4-steps-ahead forecast procedure was carried out for 24 quarters, with the first forecast beginning in 2001:Q1. This experiment produced a total of 24 one-quarter-ahead forecasts, 24-two-quarters-ahead forecasts, and so on, up to 24 4-step-ahead forecasts. We use the algorithm in the Econometric Toolbox of MATLAB[14], for this purpose. The RMSEs[15]

[11] See Dua and Ray (1995) for further details.

[12]Details about data and their statistical treatment of the variables used to estimate the DFM are available upon request.

[13] The choice of 5 lags is based on the unanimity of the sequential modified LR test statistic, Akaike information criterion (AIC), the final prediction error (FPE) criterion, Schwarz information criterion and the Hannan-Quinn (HQ) information criterion applied to a stable VAR estimated with the three variables of concern. Note, stability, as usual, implies that no roots were found to lie outside the unit circle.

[14] All statistical analysis was performed using MATLAB, version R2006a.

for the 24, quarter 1 through quarter 4 forecasts are then calculated for the per capita growth, CPI inflation and the short-term nominal interest rate. The values of the RMSE statistic for one- to four-quarters-ahead forecasts for the period 2001:Q1 to 2006:Q4 are then examined. The model, DFM or any of the VARs, that produces the lowest average value for the RMSE is selected, as the 'optimal' model for a specific variable.

To evaluate the accuracy of forecasts generated by the DFM, we need alternative forecasts. To make the RMSEs comparable with the DFM, we report the same set of statistics for the out-of-sample forecasts generated from an unrestricted classical VAR, the three-variable BVARs and the symmetric and asymmetric 267-variables BVARs. In Tables 1 to 3, we compare the RMSEs of one- to four-quarters-ahead out-of-sample-forecasts for the period of 2001:Q1 to 2006:Q4, generated by the abovementioned models. At this stage, a few words need to be said regarding the choice of the evaluation criterion for the out-of-sample forecasts generated from Bayesian models. As Zellner (1986: 494) points out the "optimal" Bayesian forecasts will differ depending upon the loss function employed and the form of predictive probability density function". In other words, Bayesian forecasts are sensitive to the choice of the measure used to evaluate the out-of-sample forecast errors. However, Zellner (1986) points out that the use of the mean of the predictive probability density function for a series, is optimal relative to a squared error loss function and the Mean Squared Error (MSE), and, hence, the RMSE is an appropriate measure to evaluate performance of forecasts, when the mean of the predictive probability density function is used. This is exactly what we do below in Tables 1 through 3, when we use the average RMSEs over the one- to four-quarter-ahead forecasting horizon. The conclusions, regarding each of the three variables, based on the average one- to four-quarters-ahead RMSEs, from these tables can be summarized as follows:

[INSERT TABLES 1 THROUGH 3]

(i) Per Capita Growth Rate: The DFM is outperformed by all the VAR models, small or large-scale, classical or Bayesian, symmetric or asymmetric. The three variable classical VAR performs better than all the small-scale BVARs, as well as, the large-scale asymmetric BVARs. Amongst the VARs, though, it is the large-scale symmetric BVAR which performs the best. In fact, the best-suited model for forecasting per capita growth rate is the BVAR model with the most tight prior (w = 0.1 and d = 2.0).

(ii) CPI Inflation: For CPI inflation, the DFM outperforms all the other models. Amongst the VARs, the small-scale classical VAR outperforms all the large-scale symmetric BVAR, and the three-variable BVAR except for the case where priors are most loose (w = 0.3 and d = 0.5). Further, except for two cases of relatively loose specification of the prior, specifically when w = 0.3, d = 0.5 and w = 0.2 and d = 1.0, the VAR, in general, is also better suited than the large-scale asymmetric BVARs. But, overall amongst the VARs, it is the large-scale asymmetric BVAR model with w = 0.3, d = 0.5 that is best-suited for forecasting CPI inflation.

(iii) 91-days Treasury bill rate: As with the CPI inflation, the DFM stands out in forecasting the Treasury bill rate, when compared to other alternative models. Amongst the VARs, the small-scale classical VAR outperforms all the large-scale symmetric BVAR, but is outperformed in turn, by all the small-scale BVARs and the large-scale asymmetric BVARs. Further, as with the CPI inflation, it is the large-scale asymmetric BVAR model with w = 0.3, d = 0.5 that is best-suited for forecasting the short-term interest rate amongst the VARs.

---

[15] Note that if A_{t+n} denotes the actual value of a specific variable in period $t + n$ and ${}_tF_{t+n}$ is the forecast made in period t for $t + n$, the RMSE statistic can be defined as: $\sqrt{\frac{1}{N}\sum(A_{t+n} - {}_tF_{t+n})} \times 100$. For $n = 1$, the summation runs from 2001:Q1 to 2006:Q4, and for $n = 2$, the same covers the period of 2001:Q2 to 2006:Q4, and so on.

So, to summarize, we find that the DFM outperforms the VARs, small or large-scale, classical or Bayesian, symmetric or asymmetric, by quite a margin in terms of the average RMSEs for one- to four-quarters-ahead forecasts for the CPI inflation and the Treasury bill rate, but, in turn, is outperformed by the large-scale symmetric BVAR with most tight priors. Furthermore, setting the DFM aside for a moment, we observe that the next best performing model for forecasting the CPI inflation and the Treasury bill rate is the large-scale asymmetric BVAR with most loose priors. Recall, that the large-scale symmetric BVAR treats all the variables, other than the dependent variable identically. Hence, the looseness and asymmetry in the prior specification is most-likely indicative of, respectively, persistence and the importance of relevant variables required for the determination of CPI inflation and the interest rate, especially when one takes into account of the role of foreign factors in determining these two variables. The fact that growth rate per capita is forecasted best by a symmetric BVAR, irrespective of the degree of tightness or looseness of the prior, is, perhaps, due to the equal importance of variables, other than the per capita growth rate in explaining itself.

6. Conclusions

This paper exploits information contained in a large cross-section of time series to forecast three key macroeconomic variables, namely, per capita growth rate, the Consumer Price Index (CPI) inflation, and the 91 days Treasury Bill rate for the South African economy, using a DFM and BVARs, based on alternative values of the hyperparameters specifying the prior. The two-types of models are first estimated over the period of 1980:01 to 2000:04 using quarterly data, and are then used to generate one- to four-quarters-ahead out-of-sample forecasts over a 24 quarter forecasting horizon of 2001:01 to 2006:04. The performance of these two large-scale models are also compared with each other and with an unrestricted Vector Autoregressive (VAR) model and BVAR models, with identical hyperparameter values as the large-scale BVARs, but including only the three variables we are concerned of. As stressed in the introduction, given the modeling strategies, the current study can be easily generalized to any other country and should not be dubbed as specific to the South African economy.

In general, we find that the DFM outperforms all the alternative forms of the VARs by quite a distance in terms of the average RMSEs for one- to four-quarters-ahead forecasts for the CPI inflation and the Treasury bill rate, but, in turn, is outperformed by the large-scale symmetric BVAR with most tight priors. But, setting the DFM aside for a moment, we observe that the next best performing model for forecasting the CPI inflation and the Treasury bill rate is the large-scale asymmetric BVAR with most loose priors. Based on these results, what is of more importance, is the observation that, whether it is the DFM or the BVAR (symmetric or asymmetric), it is always a large-scale model that is best-suited for forecasting the three variables of our concern. And further, it seems that, perhaps, the DFM might have an edge over the large-scale BVARs, especially when it comes to forecasting variables such as the inflation and the interest rates, which are likely to be influenced more by a specific set of important variables, unlike growth rate, which, at least in the short-run, is less likely to be driven by proper theoretical foundations. The first part of the statement is, in some sense, vindicated by the better performance of the asymmetric BVARs within the VAR category of models, which, in fact, does allow for bit of theory in terms of the Small Open Economy assumptions. While, second half of the statement in concern follows from the result that the large-scale symmetric BVAR is best-suited for forecasting growth rate per capita.

Overall, our results indicate that data-rich models – DFM or large-scale BVARs, are better suited in forecasting key macroeconomic variables relative to small-scale models involving only the few variables of interest. However, it is important to point out that, given that there are at least two major limitations to using a Bayesian approach for forecasting, the DFM is, perhaps, a better model to base ones' forecasts on. The two shortcomings of the Bayesian models are as follows: Firstly, as it is clear from Tables 1 to 3, the forecast accuracy is sensitive to the choice of the priors. So if the prior is not well specified, an alternative model used for forecasting may perform better. Secondly, in case of the Bayesian models, one requires to specify an objective function, for

example the average RMSEs, to search for the 'optimal' priors, which, in turn, needs to be optimized over the period for which we compute the out-of-sample forecasts. However, there is no guarantee that the chosen parameter values specifying the prior will continue to be 'optimal' beyond the period for which it was selected. Nevertheless, the importance of large-scale BVARs cannot be ignored[16], especially when one realizes that they are the best possible alternative to the DFM, as far as accommodating large number of time series is concerned. Further, it is also important to check for the robustness of our conclusions, by redoing the exercise with BVARs based on alternative forms of priors, other than the Minnesota-type used in this paper. In this regard, a good starting point would be to use the double exponential priors as in De Mol *et al.* (2006). In addition to this, one might want to revisit the forecast performances of the BVARs by assuming a more general error structure, as in Gupta (2007a), to account for non-constant variance of the variables, and, also look at Bayesian Vector Error Correction Models (BVECMs). As pointed out by LeSage (1990), Gupta (2006, 2007b) and Zita and Gupta (2007), even though non-stationarity is not an issue with the Bayesian approach BVECMs, in general, tends to outperform BVARs, since Error Correction Models (ECMs) use long-run equilibrium relationships from economic theory to explain short-run dynamics of data.

References

Bai, Jushan, and Serena Ng (2002), "Determining the Number of Factors in Approximate Factor Models". *Econometrica* 70: 191–221.

Bai, Jushan, and Serena Ng (2007), "Determining the Number of Primitive Shocks in Factor Models". *Journal of Business and Economic Statistics* 25: 52–60.

Bernanke, Ben, and Jean Boivin (2003), "Monetary Policy in a Data-Rich Environment". *Journal of Monetary Economics* 50: 525–546.

Burger, Philippe and Lizelle, J Van Rensburg (2008), "Metropolitan House Prices in South Africa: Do they Converge? *Forthcoming South African Journal of Economics.*

Chamberlain, Gary (1983), "Funds, Factors, and Diversification in Arbitrage Pricing Models". *Econometrica* 51: 1281–1304.

Chamberlain, Gary, and M. Rothschild (1983), "Arbitrage, Factor Structure and Mean-Variance Analysis in Large Markets". *Econometrica* 51: 1305–1324.

Cristadoro, Ricardo., Mario, Forni., Lucrezia, Reichlin, and Veronese Giovanni, (2005), "A Core Inflation Indicator for the Euro Area". *Journal of Money, Credit and Banking* 37: 539–560.

Das, Sonali., Rangan Gupta, and Alain Kabundi (2008), "Is a DFM Well-Suited for Forecasting Regional House Price Inflation?" Forthcoming *Economic Research Southern Africa Working Paper.*

De Mol, Christine., Domenico Giannone, and Lucrezia Reichlin (2006), "Forecasting Using a Large Number of Predictors: Is Bayesian Regression a Valid Alternative to Principal Components?" *CEPR Discussion Papers*, No. 5829.

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

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

Forni, Mario., Marc, Hallin, Marco, Lippi, and Lucrezia Reichlin (2000), "The Generalized Dynamic Factor Model: identification and estimation", *Review of Economics and Statistics* 82: 540–554.

Forni, Mario., Marc Hallin, Marco Lippi, and Lucrezia Reichlin (2001), "Coincident and Leading Indicators for the Euro Area". *The Economic Journal* 111: 62-85.

Forni, Mario., Marc Hallin, Marco Lippi, and Lucrezia Reichlin (2005), "The Generalized Dynamic Factor Model, One Sided Estimation and Forecasting". *Journal of the American Statistical Association* 100: 830–840.

Geweke, John (1977), "The dynamic factor analysis of economic time series". In *Latent variables in socio-economic models* (Aigner, and A. Goldberger, eds.) Amsterdam: North Holland, 365–383.

Giannone, Domenico., and Troy D. Matheson (2007), "A New Core Inflation Indicator for New

[16] Please refer to De Mol *et al.* (2007) for further details.

Zealand. *International Journal of Central Banking* 3: 145-180.
Gupta, Rangan (2006), "Forecasting the South African Economy with VARs and VECMs". *South African Journal of Economics* 74, 611-628.
Gupta, R.angan (2007a), "Forecasting the South African Economy with Gibbs Sampled BVECMs". *South African Journal of Economics* 75: 631-643.
Gupta, Rangan. (2007b), "Bayesian Methods of Forecasting Inventory Investment in South Africa", Working Paper 200704, University of Pretoria, Department of Economics.
Gupta, Rangan, and Sonali Das (2008), "Spatial Bayesian Methods for Forecasting House Prices in Six Metropolitan Areas of South Africa". *Forthcoming South African Journal of Economics.*
Gupta, R.angan and Alain Kabundi (2008), "A Dynamic Factor Model for Forecasting Macroeconomic Variables in South Africa", Mimeo.
Gupta, Rangan, and Moses Sichei (2006), "A BVAR Model for the South African Economy". *South African Journal of Economics* 74: 391-409.
Kabundi, Alain (2004), "Estimation of Economic Growth Using Business Survey Data. International Monetary Fund", Working Paper 04/69.
Kinal, Tennesse, and Jonathan Ratner (1986), "A VAR Forecasting Model of a Regional Economy: Its Construction and Comparison". *International Regional Science Review*, vol. 10, 113-126.
Kwiatowski, Denis., Peter C.B. Phillips., Peter Schmidt, and Yongcheol Shin (1992), "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root?" *Journal of Econometrics* 54: 159–178.
LeSAGE, James P. (1990). "A Comparison of the Forecasting Ability of ECM and VAR Models". *The Review of Economics and Statistics* 72 (4):664-671.
LeSAGE, James. P. (1999). *Applied Econometrics Using MATLAB*, www.spatial-econometrics.com.
Litterman, Robert B. (1981), "A Bayesian Procedure for Forecasting with Vector Autoregressions". Working Paper, Federal Reserve Bank of Minneapolis.
Litterman, Robert B. (1986), "Forecasting with Bayesian Vector Autoregressions – Five Years of Experience". *Journal of Business and Economic Statistics* 4 (1):25-38.
Sargent, Thomas J., and Christopher A. Sims (1977), "Business cycle modelling without pretending to have too much a priori economic theory". In *New methods in business research* (C. Sims, eds.) Federal Reserve Bank of Minneapolis.
Shoesmith, Gary L. (1992), "Co-integration, Error Correction and Medium-Term Regional VAR Forecasting". *Journal of Forecasting* 11: 91-109.
Sims, Christopher A., James. H Stock, and Mark W. Watson, (1990), "Inference in Linear Time Series Models with Some Unit Roots". *Econometrica* 58: 113-144.
Sims, Christopher A. (1980), "Macroeconomics and Reality". *Econometrica* 48:1-48.
Spencer, David E. (1993), "Developing a Bayesian Vector Autoregression Model". *International Journal of Forecasting"*, 9: 407-421.
Stock, James H., and Mark W. Watson (1989), "New indexes of coincident and leading economic indicators". *NBER Macroeconomics Annual* 351-393.
Stock, James H., and Mark W. Watson (1991), "A probability model of the coincident indicators". In *Leading economic indicators: New approaches and forecasting record* (K. Lahiri, and G. Moore, eds.) Cambridge: Cambridge University Press, pp. 63–95.
Stock, James H., and Mark W. Watson (1999), "Forecasting Inflation". *Journal of Monetary Economics* 44: 293-335.
Stock, James H., and Mark W. Watson (2002a), "Forecasting Using Principal Components from a Large Number of Predictors". *Journal of the American Statistical Association* 97: 147–162.
Stock, James H., and Mark W. Watson (2002b), "Macroeconomic Forecasting Using Diffusion Indexes". *Journal of Business and Economics Statistics* 20: 147–162.
Stock, James H., and Mark W. Watson (2005), "Implications of Dynamic Factor Models for VAR Analysis", NBER Working Paper 11467.
Theil, Herni (1971), *Principles of Econometrics*. New York, John Wiley.
Todd, Richard M. (1984), "Improving Economic Forecasting with Bayesian Vector Autoregression". *Quarterly Review*, Federal Reserve Bank of Minneapolis, Fall, 18-29.
Van Nieuwenhuyze, Christophe (2007), "A Generalized Dynamic Factor Model for the Belgian Economy Identification of the Business Cycle and GDP Growth Forecasts". *Journal of Business Cycle Measurement and Analysis* 2: 213–248.

Zellner, Arnold (1986), "A Tale of Forecasting 1001 Series: The Bayesian Knight Strikes Again". *International Journal of Forecasting* 2:494-494.
Zita, Samuel E, and Rangan Gupta (2007), "Modelling and Forecasting the Metical-Rand Exchange Rate". Working Paper 200702, University of Pretoria, Department of Economics.

Table1. RMSEs for Per Capita Growth Rate (2001:01-2006:04)

		1	2	3	4	Average
	DFM	0.4556	0.4856	0.5219	0.7869	0.5625
	VAR	0.2733	0.2167	0.1148	0.0599	0.1662
w=0.3,d=0.5	**BVAR1**	0.4473	0.2111	0.1985	0.0029	0.2149
	BVAR2	0.2700	0.2100	0.1100	0.0600	0.1625
	BVAR3	0.5013	0.1428	0.3264	0.3314	0.3255
w=0.2,d=1	**BVAR1**	0.4573	0.1857	0.1554	0.0154	0.2034
	BVAR2	0.2500	0.1900	0.1000	0.0500	0.1475
	BVAR3	0.4365	0.0004	0.2979	0.2987	0.2584
w=0.1,d=1	**BVAR1**	0.4498	0.1945	0.1481	0.0263	0.2047
	BVAR2	0.2100	0.1600	0.0700	0.0300	0.1175
	BVAR3	0.3776	0.0109	0.1693	0.2050	0.1907
w=0.2,d=2	**BVAR1**	0.4846	0.1562	0.2732	0.1277	0.2604
	BVAR2	0.1000	0.0900	0.0400	0.0100	0.0600
	BVAR3	0.3520	0.2204	0.2390	0.3031	0.2786
w=0.1,d=2	**BVAR1**	0.4579	0.1941	0.2198	0.1345	0.2516
	BVAR2	0.0000	0.0100	0.0000	0.0600	**0.0175**
	BVAR3	0.5013	0.1428	0.3264	0.3314	0.3255

Note: BVAR1: 3-Variable BVAR; BVAR2: Large Symmetric BVAR; BVAR3: Large Asymmetric BVAR

Table2. RMSEs for CPI Inflation (2001:01-2006:04)

		1	2	3	4	Average
	DFM	0.0111	0.0112	0.0112	0.0112	**0.0112**
	VAR	0.1695	0.0826	1.1857	0.7337	0.5429
w=0.3,d=0.5	**BVAR1**	0.0437	0.0206	1.2979	0.7093	0.5179
	BVAR2	0.1699	0.0947	1.1892	0.7358	0.5474
	BVAR3	0.0680	0.2534	0.6183	0.5343	0.3685
w=0.2,d=1	**BVAR1**	0.5768	0.1676	1.6281	1.0162	0.8472
	BVAR2	0.1410	0.1703	1.2171	0.7565	0.5712
	BVAR3	0.2777	0.1419	1.1231	0.5669	0.5274
w=0.1,d=1	**BVAR1**	0.5339	0.2034	1.5672	1.0458	0.8376
	BVAR2	0.0644	0.2939	1.2773	0.8011	0.6092
	BVAR3	0.2395	0.2457	1.0539	0.7546	0.5734
w=0.2,d=2	**BVAR1**	1.0603	0.0546	1.9936	0.9676	1.0190
	BVAR2	0.0921	0.4841	1.3781	0.9006	0.7137
	BVAR3	0.7323	0.4874	1.2126	0.8873	0.8299
w=0.1,d=2	**BVAR1**	0.9343	0.1142	1.7622	1.0941	0.9762
	BVAR2	0.1570	0.5919	1.4823	0.9777	0.8022
	BVAR3	0.6059	0.5223	1.2071	0.7793	0.7786

Note: BVAR1: 3-Variable BVAR; BVAR2: Large Symmetric BVAR; BVAR3: Large Asymmetric BVAR

Table3. RMSEs for 91 Days Treasury Bill Rate (2001:01-2006:04)

		1	2	3	4	Average
	DFM	0.0169	0.0121	0.0116	0.0104	**0.0127**
	VAR	0.0395	0.1756	0.6709	0.8999	0.4465
w=0.3,d=0.5	**BVAR1**	0.1262	0.2213	0.1173	0.3531	0.2045
	BVAR2	0.0446	0.1823	0.6749	0.9045	0.4516
	BVAR3	0.1728	0.1113	0.1610	0.1775	0.1557
w=0.2,d=1	**BVAR1**	0.1456	0.1901	0.2735	0.4678	0.2693
	BVAR2	0.0771	0.2227	0.7010	0.9322	0.4833
	BVAR3	0.1455	0.2013	0.1506	0.2412	0.1846
w=0.1,d=1	**BVAR1**	0.1957	0.1505	0.4869	0.4935	0.3317
	BVAR2	0.1657	0.3330	0.7552	0.9974	0.5628
	BVAR3	0.1252	0.1858	0.0823	0.1111	0.1261
w=0.2,d=2	**BVAR1**	0.1425	0.1981	0.2585	0.4579	0.2642
	BVAR2	0.1246	0.2814	0.7330	0.9681	0.5268
	BVAR3	0.1252	0.1858	0.0823	0.1111	0.1261
w=0.1,d=2	**BVAR1**	0.1837	0.1789	0.4377	0.4835	0.3209
	BVAR2	0.1717	0.3291	0.7054	0.9454	0.5379
	BVAR3	0.1518	0.2750	0.1266	0.1495	0.1757

Note: BVAR1: 3-Variable BVAR; BVAR2: Large Symmetric BVAR; BVAR3: Large Asymmetric BVAR