



University of Pretoria
Department of Economics Working Paper Series

**A Dynamic Factor Model for Forecasting Macroeconomic
Variables in South Africa**

Rangan Gupta

University of Pretoria

Alain Kabundi

University of Johannesburg

Working Paper: 2008-15

June 2008

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

A Dynamic Factor Model for Forecasting Macroeconomic Variables in South Africa

Rangan Gupta *

and

Alain Kabundi[†]

March 11, 2008

Abstract

This paper uses the Dynamic Factor Model (DFM) framework, which accommodates a large cross-section of macroeconomic time series for forecasting the per capita growth rate, inflation, and the nominal short-term interest rate for the South African economy. The DFM used in this study contains 267 quarterly series observed over the period 1980Q1-2006Q4. The results, based on the RMSEs of one- to four-quarters-ahead out of sample forecasts over 2001Q1 to 2006Q4, indicate the DFM outperforms the NKDSGE in forecasting per capita growth, inflation and the nominal short-term interest rate. Moreover, the DFM performs no worse compared to the VARs, both classical and Bayesian, indicating the blessing of dimensionality.

Journal of Economic Literature Classification: C11, C13, C33, C53.

Keywords: Dynamic Factor Model, VAR, BVAR, NKDSGE Model, 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.

[†] 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

For a long time policy makers, the general public, and academics have been interested in producing accurate forecasts economic variables for various reasons. For example, recent economic credit crisis has shown the urgency of predicting such events ahead of time, instead of being behind the curve. Model builders have exploited the recent development in computation to write simple and complex models that simulate reality with high degree of accuracy. To mimic economic relationships there is an increasing need of large information. However, traditional economic models, such univariate time series, Kalman Filter, and multivariate VAR are limited, in that they cannot accommodate large number of time series. Although VAR is popular, compared to traditional macroeconomic models for its forecasting ability, its limitations-the most conspicuous being that it cannot accommodate a large panel of series without the risk of running short of degrees of freedom. In practice forecasters and policymakers often extract information from many series than the one included in a VAR. The main problem of small models is the decision regarding the choice of correct variables to include.

As Bernanke and Boivin (2003) argue eloquently, central banks monitor and analyze literally thousands of data from various sources. Since central banks pay the costs of analyzing a wide range data to improve their decisions, econometric models should considerably take into account the marginal benefits that increasing information brings in forecasting. Progress has been made in last decades to accommodate large panel of time aeries in forecasting through the use of factor models. The initial contribution in this era comes from the work of Burns and Mitchell (1946) in their analysis of business cycles. Subsequent works of Sargent and Sims (1977) and Geweke (1977) improved the original work of Burns and Michthell by introducing the static factor approach through the generalization to the dynamic case. They exploited the dynamic interrelationship of variables and then reduced the number of common factors even further. However, the approach followed by Sargent and Sims (1977) and Geweke (1977) is too restrictive, in that, it imposes orthogonality of idiosyncratic components. Chamberlain (1983) and Chamberlain and Rothschild (1983) admit the possibility of serial correlation and weakly cross-sectional correlation of idiosyncratic components.

These dynamic factor models have been improved recently through advances in estimation techniques proposed by Stock and Watson (2002), Forni et al. (2005), and Kapetanios and Marcellino (2004). This progress made in estimation generates an increasing interest in academia, international organizations, central banks, and the governments as they

accommodate large panel of time series in forecasting economic variables. However, there still is divergence as to whether factor models with large cross-section of time series outperform traditional econometric models with limited number of variables. There are several empirical researches that 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 Watson, 2002a, 2002b, 1999, 1991, and 1989). However, recent empirical studies find no or only minor improvements in forecasting ability (Schumacher, 2006; Schumacher and Dreger, 2004; Gosselin and Tkacz, 2001; Angelini et al., 2001). These conflicting results have been issue of fascinating debate as to whether the victory proclaimed by proponents of large models was not too early. Some attribute the success of large models to the different circumstances. Banerjee et al. (2005), for example, find that factor models are relatively good at forecasting real variables in the US compared to the euro area, while, euro-area nominal variables are easier to predict using factor models than US nominal variables. Furthermore, Boivin and Ng (2006) claim that the composition of the dataset and the size of the cross-section dimension matter to produce better forecasts with factor models.

This paper exploits the sparse of information sets in a large-dimensional dynamic factor model framework, developed by Forni et al. (2005), that contains 267 quarterly series of South African economy observed over the period 1980Q1-2006Q4. 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) of America. The in-sample period contains data from 1980Q1 to 2000Q4. The rational behind the use of DFM is to extract few common factors that that drive all series included in the panel. The model estimated here shows in total five common factors explain the maximum variation of series in South Africa. The DFM is then used to forecast per capita growth (percentage change in real per capita GDP), inflation (percentage change in the implicit GDP deflator) and a measure of short-tem nominal interest rate (91-days Treasury bill rate) over the out-of-sample horizon spanning the period of 2001Q1 to 2006Q4. The forecasting performance of the DFM is evaluated in comparison to three other alternative models, namely, an unrestricted classical VAR, optimal Bayesian VARs¹ (BVARs) and a New-Keynesian Dynamic Stochastic General Equilibrium (NKDSGE), on the basis of the Root Mean Square Error (RMSE) of the out-of-sample forecasts. Although Kabundi (2007) used the DFM to

¹See Section 5 for further details regarding the issue of optimality of BVARs.

assess the synchronization of South Africa and the US, and the channels through which the US supply and demand shocks are transmitted, to our knowledge this the first attempt to use a DFM to forecast key macroeconomic variables in South Africa. Besides, the introduction and the conclusions, the remainder of the paper is organized as follows: Section 2 lays out the DFM, while Section 3 discusses the data use to estimate the DFM. Section 4 outlines the basics of the VAR, BVAR and the NKDSGE models, and Section 5 presents the results from the forecasting exercise.

2 The Model

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 output growth, inflation rate, and nominal interest rates. 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 $q \ll n$, 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 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 $Y_t = (y_{1t}, \dots, y_{nt})'$. 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:

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

or, in vector notation:

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

where f_t is a $q \times 1$ vector of dynamic factors, $B(L) = B_0 + B_1L + \dots + B_sL^s$ 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 $\chi_{it} = B(L)f_t$ 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 eigenvectors decomposition of spectral density matrix of X_t , which is a generalization of orthogonalization process in case of static principal components. The spectral density matrix of X_t , which is estimated using the frequency $-\pi < \theta < \pi$, can be decomposed into the spectral densities of the common and the idiosyncratic component:

$$\Sigma(\theta) = \Sigma_\chi(\theta) + \Sigma_\xi(\theta) \quad (3)$$

where $\Sigma_\chi(\theta) = B(e^{-i\theta})\Sigma_f(\theta)B(e^{-i\theta})'$ is the spectral density matrix of the common component χ_t and $\Sigma_\xi(\theta)$ is the spectral density matrix of the idiosyncratic component ξ_t . The rank $\Sigma_\chi(\theta)$ of is equal to the number of dynamic factors, q . Similarly, the covariance matrix of X_t can be decomposed as:

$$\Gamma_k = \Gamma_k^\chi + \Gamma_k^\xi \quad (4)$$

where $\Gamma_k^\chi = \Lambda \Gamma_k^F \Lambda'$, Γ_k^F is the covariance of F_t at the lag k and Γ_k^ξ is the covariance of ξ_t . The rank of Γ_k^χ is r , the number of static factors.

The forecast of the i th variable h -steps ahead forecast is not feasible in practice since the common factors are unobserved. However, if data follow an approximate dynamic factor model, the set of common factors F_t can be consistently estimated by appropriate cross-sectional averages, or aggregators in the terminology of Forni and Reichlin (1998) and Forni and Lippi (2001). The rationale is that using the law of large numbers, only the pervasive common

sources survive the aggregation, since the weakly correlated idiosyncratic errors are averaged out. Building on Chamberlain and Rothschild (1983), Forni, Hallin, Lippi, and Reichlin (2000) and Stock and Watson (2002) have shown that principal components of the observed variables X_t , are appropriate averages. That is, the common component can be approximated by projecting either on the first r principal components of the covariance matrix (see Stock and Watson (2002)) or on the first q dynamic principal components (see Forni, Hallin, Lippi, and Reichlin (2000)).

Empirically we estimate the autocovariance matrix of standardized data, $\hat{X}_t = (\hat{x}_{1t}, \dots, \hat{x}_{nt})'$ by:

$$\Gamma_k = \frac{1}{T - K - 1} \sum_{t=k}^T \hat{X}_t \hat{X}_t' \quad (5)$$

where T is the sample size. Following Forni et al. (2005) the spectral density matrix will be estimated by averaging a given number m of autocovariances:

$$\hat{\Sigma}(\theta) = \frac{1}{2\pi} \sum_{k=-m}^m w_k \hat{\Gamma}_k e^{-i\theta k} \quad (6)$$

where w_k are barlett-lag window estimator weights $w_k = 1 - \frac{|k|}{m+1}$. The consistent estimates are ensured, provided that $m \rightarrow \infty$ and $\frac{m}{T} \rightarrow 0$ as $T \rightarrow \infty$. In the empirical section we will use $m = \sqrt{T}$, which satisfies the above asymptotic requirements.

The procedure of by Forni, Hallin, Lippi, and Reichlin (2005) consists of two steps. The first step is the problem of the spectral density matrix, defined, at a given frequency θ , as:

$$\hat{\Sigma}(\theta)V_q(\theta) = V_q(\theta)D_q(\theta) \quad (7)$$

where $D_q(\theta)$ is a diagonal matrix having the diagonal on the first q largest eigenvalues of $\hat{\Sigma}(\theta)$ and $V_q(\theta)$ is the $n \times q$ matrix whose columns are the corresponding eigenvectors. Let X_t be driven by q dynamic factors, the spectral density matrix of the common component is given by:

$$\hat{\Sigma}_\chi(\theta) = V_q(\theta)D_q(\theta)V_q(\theta)' \quad (8)$$

Hence, the spectral density matrix of the idiosyncratic part is estimated as a residual:

$$\hat{\Sigma}_\xi(\theta) = \hat{\Sigma}(\theta) - \hat{\Sigma}_\chi(\theta) \quad (9)$$

The covariance matrices of common and idiosyncratic parts are computed through an inverse Fourier transform of spectral density matrices:

$$\hat{\Sigma}_k^\chi = \frac{2\pi}{2m+1} \sum_{k=-m}^m \hat{\Sigma}_\chi(\theta_j) e^{ik\theta_j} \quad (10)$$

$$\hat{\Sigma}_k^\xi = \frac{2\pi}{2m+1} \sum_{k=-m}^m \hat{\Sigma}_\xi(\theta_j) e^{ik\theta_j} \quad (11)$$

where $\theta_j = \frac{2\pi}{2m+1}j$ and $j = -m, \dots, m$

In a second step, the estimated covariance matrix of the common components is used to construct the factor space by r contemporaneous averages. These r contemporaneous are solutions from the generalized principal components (GPC) problem:

$$\hat{\Gamma}_0^\chi V_{rg} = \hat{\Gamma}_0^\xi V_{rg} D_{rg} \quad (12)$$

$$\text{s.t. } V_{rg}' \hat{\Gamma}_0^\xi V_{rg} = I_r$$

where D_{rg} is a diagonal matrix having on the diagonal the first r largest generalized eigenvalues of the pair $(\hat{\Gamma}_0^\chi, \hat{\Gamma}_0^\xi)$ and V_{rg} is the $n \times r$ matrix whose columns are the corresponding eigenvectors.

The first r GPCs are defined as:

$$\hat{F}_t^g = V_{rg}' \hat{X}_t \quad (13)$$

The out off-diagonal elements of $\hat{\Gamma}_0^\xi$ are set to zero to overcome the problem of instability that is common in the generalized principal component methodology. With such restrictions, the generalized principal components can be seen as static principal components computed on weighed data, in that these weights are inversely proportional to the variance of the idiosyncratic components. Such a weighting scheme should provide more efficient estimates of the common factors.

$$\hat{\chi}_{iT+r|T} = \hat{\Gamma}_{i,r}^\chi V_{rg} (V_{rg}' \hat{\Gamma}_0^\xi V_{rg})^{-1} V_{rg}' \hat{X}_T \quad (14)$$

and

$$\hat{\xi}_{iT+r|T} = \left[\hat{\Gamma}_{ii,r}^{\xi}, \dots, \hat{\Gamma}_{ii,r+p}^{\xi} \right] W_{i,k}^{-1} [\hat{x}_{iT}, \dots, \hat{x}_{iT-p}]' \quad (15)$$

where

$$W_{i,k} = \begin{bmatrix} \hat{\Gamma}_{ii,0} & \cdots & \hat{\Gamma}_{ii,-(k-1)} \\ \cdots & \cdots & \cdots \\ \hat{\Gamma}_{ii,k-1} & \cdots & \hat{\Gamma}_{ii,0} \end{bmatrix}$$

The forecast of $y_{i, T+h|T}$ is computed as follows:

$$\hat{y}_{i, T+h|T} = \hat{\sigma}_i (\hat{\chi}_{i, T+k|T} + \hat{\xi}_{i, T+k|T}) + \hat{\mu}_i \quad (16)$$

3 Data

It is imperative in factor analysis framework to extract common components from a data rich environment. After extracting common components of output growth, inflation rate, and nominal interest rates, 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) 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. The Appendix

A to this paper contains details about the statistical treatment of all data. The in-sample period contains data from 1980Q1 to 2000Q4, while the out-of-sample set is 2001Q1-2006Q4.²

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 (2006) 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 $0.005 < 0.05$.

4 Alternative Forecasting Models

In this study, the DFM is our benchmark model. However, to evaluate the forecasting performance of the DFM, we require alternative models. In our case these are namely, the unrestricted classical VAR, BVARs, and a NKDSGE model, developed recently by Liu *et al.* (2007) for forecasting the South African economy. This section outlines the basics of the above-mentioned competing models.

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

$$\chi_t = C + \lambda(L)\chi_t + \varepsilon_t \quad (17)$$

where χ is a $(n \times 1)$ vector of variables being forecasted; $\lambda(L)$ is a $(n \times n)$ polynomial matrix in the backshift operator L with lag length p , i.e., $\lambda(L) = \lambda_1 L + \lambda_2 L^2 + \dots + \lambda_p L^p$; C is a $(n \times 1)$ vector of constant terms; and ε is a $(n \times 1)$ vector of white-noise error terms. The VAR model, thus, posits a set of relationships between the past lagged values of all variables and the current value of each variable in the model.

One drawback of VAR models is that many parameters are needed to be estimated, some of which may be in-

²A detailed list of variables and statistical treatment of all data is available upon request.

significant. 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. A popular approach to overcoming this overparameterization, as described in Litterman (1981,1986a, 1986b), Doan *et al.* (1984), Todd (1984), and Spencer (1993), is to use a Bayesian VAR (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, however, the coefficient on the first own lag of a variable, which 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 Minnesota prior means take the following form:

$$\begin{aligned}\beta_i &\sim N(1, \sigma_{\beta_i}^2) \\ \beta_j &\sim N(0, \sigma_{\beta_j}^2)\end{aligned}\tag{18}$$

where β_i represents the coefficients associated with the lagged dependent variables in each equation of the VAR, while β_j represents coefficients other than β_i . The prior variances $\sigma_{\beta_i}^2$ and $\sigma_{\beta_j}^2$, specify the uncertainty of the prior means, $\beta_i = 1$ and $\beta_j = 0$, respectively.

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 , $S(i, j, m)$, is given as follows:

$$S(i, j, m) = [w \times g(m) \times f(i, j)] \frac{\hat{\sigma}_i}{\hat{\sigma}_j}\tag{19}$$

where:

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

$$g(m) = m^{-d}, \quad d > 0$$

The term w is the measurement of standard deviation on the first own lag, and also indicates the overall tightness. A decrease in the value of w results in a tighter prior. The function $g(m)$ measures the tightness on lag m relative to lag 1, and is assumed to have a harmonic shape with a decay of d . An increase in d , tightens the prior as the number of lag increases.³ The parameter $f(i, j)$ represents the tightness of variable j in equation i relative to variable i , thus, reducing the interaction parameter k_{ij} tightens the prior. $\hat{\sigma}_i$ and $\hat{\sigma}_j$ are the estimated standard errors of the univariate autoregression for variable i and j respectively. In the case of $i \neq j$, the standard deviations of the coefficients on lags are not scale invariant (Litterman, 1986b: 30). The ratio, $\frac{\hat{\sigma}_i}{\hat{\sigma}_j}$ in (19), scales the variables so as to account for differences in the units of magnitudes of the variables.

The motivation to also use a NKDSGE model, besides the VAR and the BVARs, as a competing forecasting model to the DFM, emanates from a recent study by Liu *et al.* (2007b). In this paper, the authors used a NKDSGE model⁴, along the lines of Ireland (2004), and forecasted the growth rate, inflation, and the 91-days Treasury bill rate for the South African economy over the period of 2001Q1 to 2006Q4. The results indicated that, in terms of out-of-sample forecasting, the NKDSGE model outperformed both the Classical and the Bayesian VARs for inflation, but not for per capita growth and the nominal short-term interest rate. However, the differences in the RMSEs were not significantly different across the models. Given that South Africa has moved to an inflation targeting framework, the ability of the NKDSGE model to outperform the VAR and the BVARs in terms of forecasting inflation, gave the model tremendous economic importance. Given this, we decided to incorporate the same NKDSGE model used by Liu *et al.* (2007b), as one of the alternative models to the DFM. Note, unlike the author, we, however, use a different period of estimation. While Liu *et al.* (2007b) used 1970Q1 to 2000Q4, we started from the period of 1980Q1, which is understandably the same as that for the DFM.

³In this paper, we set the overall tightness parameter (w) equal to 0.3, 0.2, and 0.1, and the harmonic lag decay parameter (d) equal to 0.5, 1, and 2. These parameter values are chosen so that they are consistent with the ones used by Liu *et al.* (2007).

⁴See also Liu and Gupta (2007) and Liu *et al.* (2007a) for usage of calibrated and estimated versions of the Hansen (1985) Real Business Cycle model in forecasting the South African economy.

Formally, the NKDSGE model is described by the following eight equations⁵:

$$\hat{x}_t = E_t \hat{x}_{t+1} - (\hat{r}_t - E_t \hat{\pi}_{t+1}) + \left(1 - \frac{1}{\eta}\right)(1 - \rho_a) \hat{a}_t \quad (20)$$

$$\hat{\pi} = \beta E_t \hat{\pi}_{t+1} + \psi \hat{x}_t - \hat{\theta}_t / \phi, \quad \psi = \eta \left(\frac{\theta - 1}{\phi} \right) \quad (21)$$

$$\hat{r}_t = \rho_r \hat{r}_{t-1} + \rho_\pi \hat{\pi}_t + \rho_g \hat{g}_t + \rho_x \hat{x}_t + \varepsilon_{rt} \quad \varepsilon_{rt} \sim i.i.d.(0, \sigma_r^2) \quad (22)$$

$$\hat{x}_t = \hat{y}_t - \frac{1}{\eta} \hat{a}_t \quad (23)$$

$$\hat{g}_t = \hat{y}_t - \hat{y}_{t-1} + \hat{z}_t \quad (24)$$

$$\hat{a}_t = \rho_a \hat{a}_{t-1} + \varepsilon_{at} \quad 0 \leq \rho_a < 1, \varepsilon_{at} \sim i.i.d.(0, \sigma_a^2) \quad (25)$$

$$\hat{\theta}_t = \rho_\theta \hat{\theta}_{t-1} + \varepsilon_{\theta t} \quad 0 \leq \rho_\theta < 1, \varepsilon_{\theta t} \sim i.i.d.(0, \sigma_\theta^2) \quad (26)$$

$$\hat{z}_t = \varepsilon_{zt} \quad \varepsilon_{zt} \sim i.i.d.(0, \sigma_z^2) \quad (27)$$

Equations (20) and (21) models the the expectational IS curve and the New Keynesian Phillips curve, respectively, while (22) presents the interest rate rule pursued by the monetary authority. Following Ireland (2004) and Liu *et al.* (2007b), the terms of the NKDSGE model is defined as follows⁶: y_t measures the output; x_t is the output gap; r_t is the nominal short-term interest rate; π_t is the inflation rate; g_t output growth; η (≥ 1) captures the degree of marginal disutility from labor; a_t is the preference shock; $0 < \beta < 1$ is the discount factor; θ_t is the cost-push shock; ϕ governs the magnitude of the cost of price adjustment; $\psi = \eta \left(\frac{\theta-1}{\phi} \right)$; ε_{rt} captures the monetary policy shock; ρ_i , $i = a, r, x, \pi, g$, and θ , captures the persistence parameters, and; z_t is the technology shock.

As far as estimation is concerned, the BVAR model is estimated using Theil's (1971) mixed estimation technique, which involves supplementing the data with prior information on the distribution of the coefficients. For each restriction imposed on the parameter estimated, the number of observations and degrees of freedom are increased by one in an artificial way. Therefore, the loss of degrees of freedom associated with the unrestricted VAR is not a concern in the BVAR. On the other hand, the NKDSGE model is in a state-space form and can be estimated via the maximum likelihood approach.⁷

⁵See Ireland (2004) and Liu *et al.* (2007b) for details regarding the microfoundations of the model.

⁶A letter with a hat above indicates its deviation from its steady-state.

⁷For further details, please refer to Ireland (2004) and Liu *et al.* (2007b).

5 Results

In this section, we compare the one- to four-quarters-ahead RMSEs of the alternative models in relation to the DFM for the out-of-sample forecast horizon of 2001Q1 to 2006Q4. Before we proceed, three points must be emphasized: First, unlike the DFM, the VAR, BVAR and the NKDSGE are estimated using data⁸ on only the three variables of interest, with all the data obtained from the Quarterly Bulletins of the South African Reserve Bank (SARB), except for the population size, which is obtained from the World Development Indicators of the World Bank; Second, even though the DFM incorporates global variables, given that the NKDSGE model is based on closed-economy, we, as in Liu *et al.* (2007b), use the percentage change in the GDP deflator as an appropriate measure of inflation rather than the CPI, simply to ensure consistency of comparison between the alternative models; Third, the stableStability was ensured, as no roots were found to lie outside the unit circle. VAR and the BVARs were estimated with four lags, as determined by the unanimity of the sequential modified LR test statistic, Final Prediction Error (FPE), Akaike Information Criterion (AIC), Schwarz Information Criterion (SIC), and the Hannan-Quinn Information (HQIC) criterion, and; Fourth, the optimality of the BVARs is based on the minimum average RMSEs⁹ for the one- to four-quarters-ahead forecasts, produced by the combination of the values of the hyperparameters defining the overall weight (w) and tightness (d). The main results, as reported in Tables 1 to 3, can be summarized as follows:

- Except for the one-quarter-ahead forecasts of the inflation rate and the interest rate, the DFM outperforms the NKDSGE model, in terms of lower RMSEs for all the variables;
- Except for the inflation rate over all the steps and the interest rate for the one-quarter-ahead forecast, the DFM

⁸Note to be consistent with the NKDSGE model, we estimate the VAR and the BVARs using demeaned data for the three variables of interest, and, hence, stationarity is not an issue. The same was also vindicated by the Augmented Dickey-Fuller (ADF), Phillips-Perron (PP), Dickey-Fuller with GLS detrending (DF-GLS) and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests of stationarity. Besides, given that the Bayesian approach is entirely based on the likelihood function, the associated inference does not need to take special account of nonstationarity (Sims *et al.* (1990)).

⁹Zellner (1986: 494) pointed out that “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) also indicated 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. Thus, the paper uses RMSEs to evaluate out-of-sample forecasting performances of alternative models.

outperforms the VAR in forecasting the growth rate and the Treasury bill rate;

- When compared to the optimal BVARs, the DFM can only do better in terms of predicting the short-term interest rate;
- Within the VAR models, both classical and Bayesian, the BVARs with $w = 0.1$, $d = 1.0$, and $w = 0.2$ and $d = 1.0$, respectively, outperforms the VAR in terms of forecasting the growth rate and the interest rate, however, the optimal BVAR with $w = 0.1$ and $d = 2.0$, is outperformed by the classical unrestricted VAR;
- Overall, a relatively tight BVAR with $w = 0.1$ and $d = 1.0$ is best suited for forecasting the growth rate of the economy. While, the classical unrestricted VAR and the DFM are the preferred models for forecasting the inflation rate and interest rate, respectively. The NKDSGE model, unlike, in Liu *et al.* (2007b) is found to be outperformed by all the models for all three variables.¹⁰

Table 1. RMSE (2001Q1-2006Q4): Per capita Growth

QA	1	2	3	4	Average
DFM	0.4556	0.4856	0.5219	0.7869	0.5625
VAR (4)	0.4791	0.54076	0.6190	0.7348	0.5935
BVAR ($w = 0.1, d = 2.0$)	0.3851	0.4027	0.4951	0.6408	0.4809
NKDSGE	0.9894	1.1259	1.7950	1.2081	1.2796

QA: Quarter Ahead; RMSE: Root Mean Squared Error in %.

Table 2. RMSE (2001Q1-2006Q4): Inflation

QA	1	2	3	4	Average
DFM	0.4158	0.3345	0.4551	0.5830	0.4471
VAR (4)	0.2617	0.3134	0.3703	0.4342	0.3449
BVAR ($w = 0.1, d = 1.0$)	0.2608	0.3105	0.3747	0.4649	0.3527
NKDSGE	0.2973	0.4212	0.5770	0.6954	0.4977

QA: Quarter Ahead; RMSE: Root Mean Squared Error in %.

¹⁰This result should not be surprising, given the sensitiveness of forecasts to the choice of sample period.

Table 3. RMSE (2001Q1-2006Q4): Treasury bill

QA	1	2	3	4	Average
DFM	1.6913	1.2054	1.1597	1.0351	1.2729
VAR (4)	0.9213	1.7685	2.5054	3.1125	2.0769
BVAR (w = 0.2, d = 1.0)	0.9183	1.6988	2.3432	2.8648	1.9563
NKDSGE	1.1301	1.9785	2.6223	3.9910	2.4305

QA: Quarter Ahead; RMSE: Root Mean Squared Error in %.

In order to evaluate the models' forecast accuracy, we perform the across-model test between the benchmark DFM with that of the VAR, optimal BVARs and the NKDSGE model. The across-model test is based on the statistic proposed by Diebold and Mariano (1995). The test statistic is defined as the following. For instance, let $\{e_t^i\}_{t=1}^T$, with $i = \text{VAR, BVARs, and NKDSGE}$, denote the associated forecast errors from the alternative models and $\{e_t^d\}_{t=1}^T$ denote the forecast errors from the DFM. The test statistic is then defined as $s = \frac{l}{\sigma_l}$, where l is the sample mean of the "loss differentials" with $\{l_t\}_{t=1}^T$ obtained by using $l_t = (e_t^i)^2 - (e_t^d)^2$ for all $t = 1, 2, 3, \dots, T$, and where σ_l is the standard error of l . The s statistic is asymptotically distributed as a standard normal random variable and can be estimated under the null hypothesis of equal forecast accuracy, i.e. $l = 0$. Therefore, in this case, a positive value of s would suggest that the DFM outperforms the specific alternative model the comparison is made against in terms of out-of-sample forecasting. Results are reported in Table 4. In general, the DFM performs significantly better than the NKDSGE model in predicting the three variables of our concern. Moreover, the DFM performs no worse than the VARs, both classical and Bayesian, since most of the test statistics, except for the fourth quarter of the inflation rate, are insignificant.

Table 4. Across-Model Test Statistics

Quarters Ahead	1	2	3	4
(A) Per capita growth				
DFM vs. VAR	0.0302	0.5017	1.1355	-1.0868
DFM vs. BVAR	-0.3360	-0.3070	-0.0954	-0.7406
DFM vs. NKDSGE	3.2364***	2.4632***	1.8891**	1.7622**
(B) Inflation				
DFM vs. VAR	-0.8479	-0.2026	-0.7181	-1.5884*
DFM vs. BVAR	-1.2773	-0.1654	-0.5788	-1.5020*
DFM vs. NKDSGE	-2.0645***	1.2602	0.7863	0.8469
(C) Treasury bill				
DFM vs. VAR	-0.3373	0.4463	1.005	1.7291**
DFM vs. BVAR	-0.4083	0.1809	0.3960	0.5662
DFM vs. NKDSGE	-0.8068	1.3774*	1.5184*	1.4960*

Note: ***(**)[*] indicates significance at 1%(5%)[10%] level of significance.

6 Conclusions

This paper assesses the forecasting performance of Dynamic Factor Model in a large panel time series in comparison to the classical VAR and the Bayesian VAR. The DFM accommodates 267 quarterly series of South Africa, a set of global variables, and some series of South Africa series of major trading partners. The model extracts five static factors and two dynamic factors that explain the most of variation in the entire panel. These factors are used to forecast output growth, inflation rate and the nominal interest rate. The three models are evaluated based on the minimum average RMSEs for the one- to four-quarters-ahead forecasts. Overall, the results show that the DFM performs significantly better than the NKDSGE model in predicting the three variables of our concern. The DFM outperforms VAR in

forecasting output and the nominal interest rate, while it only beats the BVAR in predicting the nominal interest rate. But, the statistical tests show that DFM performs no worse than the VARs, both classical and Bayesian, since most of the results are insignificant, indicating the blessing of dimensionality.

Note, even though the VAR and the “optimal” BVAR outperforms the DFM in terms of forecasting inflation and per capita growth respectively, based on the RMSEs, it is important to stress the following facts: (i) Practically speaking, a central bank, or for that matter any forecaster, would ideally want to include a large number of variables, into the forecasting model, to obtain forecasts for the variables of key interest. In this regard, the VAR estimation is disadvantaged due to the curse of dimensionality. The BVAR though, can be considered a valid alternative to DFM as it can equally accommodate large number of variables, given that its estimation is based on the Theil’s (1971) mixed estimation technique, which amounts to supplementing the data with prior information on the distribution of the coefficients, therefore, the loss of degrees of freedom associated with the unrestricted VAR is no longer a concern. (ii) In addition, due to fact that the problem associated with the degrees of freedom is no longer valid for the DFM and the BVAR, these models are also capable of forecasting simultaneously a large number of time series, beyond the possible key variables of interest; (iii) There are, however, limitations to using the Bayesian approach. Firstly, the forecast accuracy depends critically on the specification of the prior, and secondly, the selection of the prior based on some objective function for the out-of-sample forecasts may not be “optimal” for the time period beyond the period chosen to produce the out-of-sample forecasts, and; (iv) Finally, general to any traditional statistically estimated models, for example the DFM, VAR, and BVAR used for forecasting at the business cycle frequencies, there are couple of other concerns. Such procedures perform reasonably well as long there are no structural changes experienced in the economy, but changes of this nature, whether in or out of the sample, would then render the models inappropriate. Alternatively, these models are not immune to the ‘Lucas Critique’. Furthermore, the estimation procedures used here are linear in nature, and, hence, they fail to take into account of nonlinearities in the data. In this regard, the role of microfounded DSGE models cannot be disregarded. The fact that the NKDSGE, based on the sample period used, is outperformed by all the other models, mainly calls for better modeling of the South African economy, by extending the current model to incorporate facts like inflation persistence, role of capital in the production process, habit persistence, and, perhaps, more importantly the role of external shocks, given South Africa’s small open economy structure.

But, whatever the limitations of the DFM, as we show in this paper, one cannot gainsay the importance of this kind

of modeling strategy in forecasting three key variables, namely, per capita growth rate, inflation and the short-term interest rate for South Africa over the period of 2001Q1 to 2006Q4. Clearly, the DFM stands in great stead, relative to alternative popular forecasting methods, in predicting the South African economy.

References

- Angelini, E., J. Henry, and R. Mestre (2001) Diffusion index-based inflation forecasts for the euro area. *Working Paper No. 61* ECB.
- Bai, J., and S. Ng (2007) Determining the Number of Primitive Shocks in Factor Models. *Journal of Business and Economic Statistics* 25: 52–60.
- Bai, J., and S. Ng (2002) Determining the Number of Factors in Approximate Factor Models. *Econometrica* 70: 191–221.
- Banerjee, A., M. Marcellino, and I. Masten (2005) Leading indicators for euro-area inflation and GDP growth. *Oxford Bulletin of Economics and Statistics* 67: 785–814.
- Bernanke, B., and J. Boivin (2003) Monetary Policy in a Data-Rich Environment. *Journal of Monetary Economics* 50: 525–546.
- Boivin, J., and S. Ng (2006) Are more data always better for factor analysis. *Journal of Econometrics* 132: 169–194.
- Burns, A.M., and W.C. Mitchell (1946) Measuring Business Cycles. *National Bureau of Economic Research*
- Chamberlain, G., and M. Rothschild (1983) Arbitrage, Factor Structure and Mean-Variance Analysis in Large Markets. *Econometrica* 51: 1305–1324.
- Chamberlain, G. (1983) Funds, Factors, and Diversification in Arbitrage Pricing Models *Econometrica* 51: 1281–1304.
- Cristadoro, R., M. Forni, L. Reichlin, and G. Veronese (2005) A Core Inflation Indicator for the Euro Area *Journal of Money, Credit and Banking* 37: 539–560.
- Diebold, F. X. and R. S. Mariano (1995) Comparing predictive accuracy. *Journal of Business and Economic Statistics* 13: 253-263.
- Doan, T., Litterman, R. B. and C. A. Sims (1984) Forecasting and conditional projection using realistic prior distributions. *Econometric Reviews* 3: 1-144.
- Elliott, G., T.J. Rothenberg, and J. Stock (1996) Efficient Tests for an Autoregressive Unit Root. *Econometrica* 64: 813–36.
- Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2005) The Generalized Dynamic Factor Model: identification and

estimation. *Review of Economics and Statistics* 82: 540–554.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2003) Do financial variables help forecasting inflation and real activity in the euro area? *Journal of Monetary Economics* 50: 1243-1255.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2000) The Generalized Dynamic Factor Model: identification and estimation. *Review of Economics and Statistics* 82: 540–554.

Forni, M., and M. Lippi (2001) The Generalized Dynamic Factor Model: representation theory. *Econometric Theory* 17: 1113–1141.

Forni, M., and L. Reichlin (1998) Let's get real: a factor analytical approach to disaggregated business cycle dynamics. *Review of Economic Studies* 65: 453–473.

Geweke, J. (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, D., and T.D., Matheson (2007) A New Core Inflation Indicator for New Zealand. *International Journal of Central Banking* 3: 145-180.

Gosselin, M.-A., and G. Tkacz (2001) Evaluating factor models: an application to forecasting inflation in Canada. *Working Paper No. 18* Bank of Canada.

Hansen, G. D. (1985) Indivisible labor and the business cycle. *Journal of Monetary Economics* 16: 309-327.

Ireland, P. N. (2004) Technology shocks in the new Keynesian model. *Review of Economics and Statistics* 86: 923-936.

Kabundi, A. (2007) Synchronization between South Africa and the U.S.: A Structural Dynamic Factor Analysis. Economic Research Southern Africa, University of Cape Town, *Working Paper, No. 50*.

Kabundi, A., (2004) Estimation of Economic Growth Using Business Survey Data. International Monetary Fund, *Working Paper No. 04/69*.

Kapetanios and Marcellino(2004) A parametric estimation method for dynamic factor models of large dimensions. *mimeo*

Kwiatowski, D., P.C.B. Phillips, P. Schmidt, and Y. 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.

- Litterman, R. B. (1981) A Bayesian procedure for forecasting with vector autoregressions. Federal Reserve Bank of Minneapolis, *Working Paper*.
- Litterman, R. B. (1986a) A statistical approach to economic forecasting. *Journal of Business and Statistics* 4: 1-4.
- Litterman, R. B. (1986b) Forecasting with Bayesian vector autoregressions—five years of experience. *Journal of Business and Statistics*, 4: 25-38.
- Liu G. L. and R. Gupta (2007) A small-scale DSGE model for forecasting the South Africa economy. *South Africa Journal of Economics*, 75: 179-193.
- Liu G. L., R. Gupta, and E. Schaling (2007a) Forecasting the South African economy: A DSGE-VAR approach, Economic Research Southern Africa, University of Cape Town, *Working Paper, No.51*.
- Liu G. L., R. Gupta, and E. Schaling (2007b) A NKDSGE model for forecasting the South African Economy. *Working Paper*, University of Pretoria.
- Sargent, T. J., and C. 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.
- Schumacher, C. (2007) Forecasting German GDP using alternative factor models based on large datasets. *Journal of Forecasting* 26: 271–302.
- Schumacher, C., and C. Dreger (2004) Estimating large-scale factor models for economic activity in Germany: Do they outperform simpler models? *Jahrbcher fr Nationalökonomie und Statistik* 224: 731–750.
- Schneider, M., and M. Spitzer (2004) Forecasting Austrian GDP using the generalized dynamic factor model. Oesterreichische National Bank, *Working Paper No. 89*.
- Sims, C. A. (1980) Macroeconomics and reality. *Econometrica*, 48: 1-48.
- Sims, C. A., J. H. Stock, and M. W. Watson (1990) Inference in linear time series models with some unit roots. *Econometrica* 58: 113-144.
- Stock, J. H., and M. W. Watson (2002a) Forecasting Using Principal Components from a Large Number of Predictors. *Journal of the American Statistical Association* 97: 147–162.
- Stock, J. H., and M. W. Watson (2002b) Macroeconomic Forecasting Using Diffusion Indexes. *Journal of Business and Economic Statistics* 20: 147–162.
- Stock, J. H., and M. W. Watson (1999) Forecasting Inflation. *Journal of Monetary Economics* 44: 293-335.

Stock, J. H., and M. 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, J. H., and M. W. Watson (1989) New indexes of coincident and leading economic indicators. *NBER Macroeconomics Annual*, 351-393.

Theil, H. (1971) *Principles of econometrics*. New York: John Wiley.

Van Nieuwenhuyze, C. (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, A. (1986) A tale of forecasting 1001 series: The Bayesian knight strikes again. *International Journal of Forecasting*, 2: 494-494.