

FORECASTING MACROECONOMIC DATA FOR AN EMERGING MARKET WITH A NONLINEAR DSGE MODEL

Mehmet Balcilar*, Rangan Gupta[†] & Kevin Kotzé[‡]

November 2013

Abstract

This paper considers the forecasting performance of a nonlinear dynamic stochastic general equilibrium (DSGE) model. The results are compared to a wide selection of competing models, which include a linear DSGE model and a variety of vector autoregressive (VAR) models. The parameters in the VAR models are estimated with classical and Bayesian techniques; where some of the Bayesian models are augmented with stochastic-variable-selection, time-varying parameters, endogenous structural breaks and various forms of prior-shrinkage (where the Minnesota prior is included as a special case). The structure of the DSGE models follow that of New-Keynesian varieties, which allow for nominal and real rigidities. The nonlinear DSGE model makes use of the second-order solution method of Schmitt-Grohé and Uribe (2004) and a particle filter is used to generate values for the unobserved variables. Most of the parameters in these models are estimated using maximum likelihood techniques. The models are applied to the macroeconomic data of South Africa, which is classified as an emerging market economy. The initial in-sample period of 1960Q1 to 1999Q4 is used to generate an eight-step ahead forecast. The models are then estimated recursively, by extending the in-sample period by a quarter, to generate successive forecasts over the out-of-sample period, 2000Q1 to 2011Q4. We find that the forecasting performance of the nonlinear DSGE model is almost always superior to that of the linear counterpart; particularly over longer forecasting horizons. The nonlinear DSGE model also outperforms the selection of VAR models in most cases.

JEL Classifications: E0; C5; C11; C61; C63

Keywords: Macroeconomic Forecasting; Linear and Nonlinear New-Keynesian DSGE; Vector Autoregressions; Bayesian Methods, Emerging Markets

*Department of Economics, Eastern Mediterranean University, Famagusta, Turkish Republic of Northern Cyprus, via Mersin 10, Turkey.

[†]Corresponding author. Department of Economics, University of Pretoria, Pretoria, 0002, South Africa. Email: rangan.gupta@up.ac.za

[‡]The School of Economics, Faculty of Commerce, University of Cape Town, Middle Campus, Private Bag, Rondebosch, 7701, South Africa.

1 Introduction

Dynamic stochastic general equilibrium (DSGE) models are used by central banks and other policy-making institutions for policy investigations and forecasting purposes.¹ Modern variants of these models incorporate a prodigious amount of nominal and real rigidities, and a relatively large number of shocks; which has improved their forecasting potential to the extent where they have outperformed most other multivariate models, when applied developed-world macroeconomic data.²

When applied to an emerging market economy, such as South Africa, the forecasting performance of these models has been somewhat mixed.³ Initial studies by Liu and Gupta (2007), Liu, Gupta and Schaling (2009), Liu, Gupta and Schaling (2010), and Gupta and Kabundi (2011) suggest that the forecasting potential of Bayesian vector autoregressive (BVAR) models may be superior to that of small DSGE models, for key macroeconomic variables. However, studies by Steinbach, Mathuloe and Smit (2009), Gupta and Kabundi (2010), Alpanda, Kotzé and Woglom (2011) and Gupta and Steinbach (forthcoming) indicate that when one allows for open-economy features and a relatively large number of rigidities, DSGE models would appear to compete favourably with BVAR models.

Against this backdrop, the objective of this paper is to analyse whether allowing for nonlinearities in a New-Keynesian closed-economy DSGE model could be used to obtain comparable, if not better, forecasts relative to BVAR models. Our decision to make use of a closed-economy model for this analysis is based on the following reasons. Firstly, the results of previous studies that were applied to South African data suggest that the forecasting performance of small closed-economy DSGE models is usually inferior to that of other forecasting models. Hence, if we find that the forecasting performance of this nonlinear model is superior to that of other models, then it would suggest that there could be important nonlinear features in the underlying data generating process of an emerging market economy.

Furthermore, whilst the inclusion of small open-economy features have tended to improve forecasting performances for South African macroeconomic variables, the computational burden associated with estimating a small open-economy nonlinear DSGE model is not insignificant. In addition, whilst we would expect *a priori* that the small closed-economy model would provide inferior forecasts, given earlier findings, Gupta and Steinbach (forthcoming) provide evidence which suggests that models that include domestic price stickiness could outperform models that do not include this feature.^{4,5}

The majority of DSGE models that are used for forecasting purposes would usually make use of a first-order linear approximation of the theoretical model that incorporates several nonlinear features and a number of forward-looking expressions.⁶ After applying such a

¹See, Tovar (2009) for an overview of the use of these models in central banks. Edge, Kiley and Laforge (2010) provides details of the DSGE model that has been used at the Federal Reserve Bank, whilst Ratto, Roeger and in 't Veld (2008) describe the model that has been used at the European Central Bank. An early exposition of the multi-country DSGE model that has been used at the International Monetary Fund (IMF) is provided in Carabenciov, Ermolaev, Freedman, Juillard, Kamenik, Korshunov, Laxton and Laxton (2008).

²See, Smets and Wouters (2007) for an early indication of the forecasting performance of these models.

³South Africa is classified as an emerging market economy by the International Monetary Fund (2013), and at the time of writing it was also classified as such by the FTSE, S&P, Dow Jones, and MSCI.

⁴The small open-economy New-Keynesian DSGE-VAR model of Gupta and Steinbach (forthcoming), includes the following rigidities: wage indexation to past consumer inflation, domestic price inflation indexation to its past values, imperfect exchange rate pass-through, the degree of stickiness in domestic prices, wage and imported prices and the degree of habit persistence. After testing the effects of each of these rigidities they found that the most significant improvement in forecasting ability, is attributable to domestic price stickiness, which is included in the closed-economy DSGE model that has been estimated in this paper.

⁵Given the results of this paper, it may be of interest to consider the predictive ability of a small open-economy nonlinear DSGE model for an emerging market economy, such as South Africa.

⁶The proliferation of forecasting models that make use of first-order approximations has been facilitated by the development of the excellent software platform, *Dynare*. Further details for which can be found in

log-linear approximation, one is then able to derive the model solution; before making use of the Kalman filter to approximate the likelihood function of the model (which may include several unobserved variables).⁷ Whilst this procedure has been successfully applied to many problems, a first-order linear approximation may exclude important nonlinear features of the theoretical model. In addition, it might also curtail large deviations from the steady-state of the respective variables.⁸

The use of DSGE models that are estimated with higher-order approximations and nonlinear filters is not that widespread.⁹ Early advocates of this procedure include, Fernández-Villaverde and Rubio-Ramírez (2005), An and Schorfheide (2007), An (2008), Primiceri and Justiniano (2008), and DeJong and Dave (2011). These studies usually employ the second-order solution method proposed by Schmitt-Grohé and Uribe (2004) before utilising a particle filter to derive the likelihood function.¹⁰ Early investigations into the forecasting performance of these nonlinear models is provided in Pichler (2008); which suggests that whilst nonlinear models may outperform their linearised counterpart when applied to simulated data, they do not do so for actual macroeconomic data for the United States economy.

To the best of our knowledge, the current literature does not include an example of a nonlinear DSGE model that is applied to the macroeconomic data of an emerging market economy. Such an investigation would be of interest, as one would expect that this data would incorporate larger deviations from the steady-state (as well as potentially more complex nonlinear relationships). Hence, it may be the case that when applied to an emerging market economy, the nonlinear DSGE model may provide a superior out-of-sample fit, when compared to a linear counterpart. In addition, it also may have the potential to outperform other reduced-form forecasting models.

In this paper we estimate a linear and nonlinear DSGE model (as well as a large selection of competing forecasting models) for the South African economy. The competing forecasting models include classical vector-autoregressive (VAR) models and a number of BVAR varieties. These BVAR models have been estimated with various forms of the Minnesota prior and stochastic variable selection (SVS) techniques.¹¹ Some of the BVAR models that employ SVS have been extended to allow for time-varying parameters, endogenous structural breaks, and least absolute shrinkage and selection operators.

The results of this investigation suggest that the nonlinear DSGE model appears to outperform the linear counterpart for all variables in most instances. In addition, the findings suggest that these improvements are statistically significant when forecasting consumer inflation and interest rates over the medium to long horizon, and output over short to medium horizons.¹² The nonlinear DSGE model also appears to outperform the VAR and BVAR

Adjemian, Bastani, Juillard, Karame, Mihoubi, Perendia, Pfeifer, Ratto and Villemot (2011).

⁷Solution methods for linear rational expectations problems are provided by, Blanchard and Kahn (1980), Klein (2000), Sims (2001), Uhlig (1999), amongst others.

⁸After taking the first order Taylor series expansion around the central point of a nonlinear exponential function you will find that the extreme points of the nonlinear function will usually be well beyond those of the linear approximation.

⁹An (2008) suggests that this is largely due to computational complexities that are involved with such approximations. DSGE models that make use of second-order (or higher) approximations are often referred to as nonlinear DSGE models.

¹⁰Nonlinear filters have been applied in many settings to model various features of time series data. The forms that some of these filters take is discussed in Kitagawa and Gersch (1996), Doucet, Godsill and Andrieu (2000), DeJong and Dave (2011), and others. Whilst most of the research that makes use of nonlinear DSGE models utilise a particle filter to derive the likelihood function, DeJong and Dave (2011) suggest that the use of the Efficient-Information-Sampling filter may lead to improved results; when applied to structural macroeconomic models.

¹¹The specification of BVAR models with a Minnesota prior is discussed in Litterman (1986a), Litterman (1986b), Doan, Litterman and Sims (1984) and Sims and Zha (1998). The application of SVS techniques in a BVAR model is described in Koop and Korobilis (2010) and Korobilis (2011).

¹²These forecasts are evaluated after calculating the relative-root-mean-squared errors and the Diebold and Mariano (1995) statistics, which consider the significance of any observed improvement.

models when forecasting consumer inflation.¹³ In addition, when forecasting output over longer horizons, the predictive ability of the nonlinear DSGE model would appear to be superior; whilst over a shorter horizon there are a few cases where a BVAR model generates better forecasts. The forecasts for interest rates are all fairly similar, however, one of the BVAR models with Minnesota prior is able to outperform the nonlinear DSGE over the medium to long horizon.

In what follows, section 2 describes the theoretical structure and empirical techniques that are employed to estimate the DSGE models. Section 3 considers the specification of the wide selection of VAR and BVAR models. In section 4, we describe the data that is used in this study, before we discuss the results in section 5. The final section contains the conclusion.

2 Dynamic Stochastic General Equilibrium Models

2.1 Theoretical Structure

The structure of the DSGE models is consistent with the New Keynesian framework in that it incorporates features that describe monopolistic competition, capital accumulation, capital adjustment costs and various other nominal and real rigidities. The economic environment is described by the actions of households, intermediate producers, final good producers, and the central bank.¹⁴

Households maximise utility for different measures of consumption, real money balances, and leisure activities; such that after incorporating separable preferences their utility function can be expressed as,

$$E_0 \sum_{t=0}^{\infty} \beta^t U \left[\Theta_t \left(\frac{c_t^{1-\tau} - 1}{1-\tau} \right) + \chi_m \log \left\{ \frac{M_t}{P_t} \right\} + \eta_h (1 - h_t) \right] \quad (1)$$

where, β represents the subjective time discount factor, θ represents a demand shock, c_t represents consumption, τ represents the households preference for consumption, M_t/P_t represents real monetary balances, χ_m represents the households preference for monetary holdings, h_t represents leisure, and η_h represents the households preference for leisure. It is presumed that the demand shock follows an autoregressive structure, such that,

$$\log \Theta_{t+1} = \rho_{\Theta} \log \Theta_t + \epsilon_{\Theta,t+1}, \quad \text{where } \epsilon_{\Theta,t+1} \sim N(0, \sigma_{\Theta}^2) \quad (2)$$

The households utility function is then subject to a budget constraint that incorporates capital adjustment costs,

$$\frac{M_{t-1} + B_{t-1} + W_t h_t + Q_t k_t + D_t + L_t}{P_t} \geq c_t + x_t + \frac{\psi_k}{2} \left(\frac{x_t}{k_t} - \delta \right)^2 k_t + \frac{B_t/i_t + M_t}{P_t} \quad (3)$$

where, B_t represents bond holdings, W_t represents wage rate, Q_t represents the rate of return on capital, k_t represents productive capital, D_t represents dividend payments, L_t represents lump-sum transfers from government, x_t represents investment, δ represents the

¹³The BVAR with Minnesota prior appears to provide the second best results in this instance

¹⁴The structure of this model follows Pichler (2008). This allows for a comparison between the results of the models that have been applied to the macroeconomic data of a developed and emerging market economy. Hence, if the predictive ability of this nonlinear model is superior to that of the linear counterpart, then this result could be attributed to the underlying characteristics of the data generating process in the emerging market economy, given the results that are reported in Pichler (2008).

depreciation rate of capital, and i_t represents the gross nominal interest rate. The capital stock is then assumed to evolve according to the expression,

$$k_{t+1} = (1 - \delta)k_t + x_t \quad (4)$$

The firms that are involved in the production of finished goods makes use of constant returns-to-scale production technology where the j intermediate goods, $y_t(j)$, serve as the only inputs. Hence, the quantity of finished goods that are produced is determined by the expression,

$$y_t = \left[\int_0^1 y_t(j) \frac{(\theta - 1)}{\theta} dj \right]^{\frac{\theta}{(\theta - 1)}} \quad (5)$$

where, θ represents the elasticity of substitution between intermediate inputs. The price of these goods is then given by,

$$P_t = \left[\int_0^1 P_t(j) (\theta - 1) dj \right]^{\frac{1}{(1 - \theta)}} \quad (6)$$

where the demand for each intermediate good is given by,

$$y_t(j) = \left(\frac{P_t(j)}{P_t} \right)^{-\theta} y_t \quad (7)$$

Firms that are involved in intermediate production face a Cobb-Douglas production function with labour augmenting technology change,

$$y_t(j) = k_t(j) \alpha \left(z_t h_t(j) \right)^{1 - \alpha} \quad (8)$$

where, α represents capital's share of output and z_t is the technology shock that follows an autoregressive process,

$$\log z_{t+1} = (1 - \rho_z) \log \bar{z} + \rho_z \log z_t + \epsilon_{z,t}, \quad \text{where } \epsilon_{z,t} \sim N(0, \sigma_z^2). \quad (9)$$

Sticky-prices are introduced through quadratic functions that describe the cost of adjusting prices. The specification follows Rotemberg (1982), where,

$$PAC_t(j) = \frac{\psi_p}{2} \left[\frac{P_t(j)/P_{t-1}(j)}{\bar{\pi}} - 1 \right]^2 y_t P_t \quad (10)$$

where $\bar{\pi}$ denotes the steady-state value for inflation and ψ_p represents the size of adjustment costs. At the end of period t the firms distribute profits to the respective households through dividend payments, where,

$$D_t(j) = P_t(j)y_t(j) - W_t h_t(j) - Q_t K_t(j) - \frac{\psi_p}{2} \left[\frac{P_t(j)/P_{t-1}(j)}{\bar{\pi}} - 1 \right]^2 y_t P_t \quad (11)$$

The objective of each firm is then to maximise their total market value, for which we construct the optimisation problem,

$$\max_{h_t(j), k_t(j), P_t(j)} E_0 \sum_{t=0}^{\infty} \beta^t \lambda_t \frac{D_t(j)}{P_t} \quad (12)$$

where λ is the Lagrangian multiplier. Finally, to close the model, we assume that the central bank conducts monetary policy by following a Taylor rule that may be described as,

$$\log \frac{i_t}{\bar{i}} = \phi_i \log \frac{i_{t-1}}{\bar{i}} + \phi_y \log \frac{y_t}{\bar{y}} + \phi_\pi \log \frac{\pi_t}{\bar{\pi}} + \epsilon_{i,t}, \quad \text{where } \epsilon_{i,t} \sim N(0, \sigma_i^2). \quad (13)$$

2.2 Model solution, likelihood functions and parameter estimates

The nonlinear DSGE model is solved using second order perturbation methods, as described by Schmitt-Grohé and Uribe (2004). For comparative purposes we make use of the first order approximation method of Klein (2000), which is used to approximate a linear DSGE model.¹⁵ The likelihood function may then be constructed with the following measurement equation in a state-space representation,

$$Y_t = GX_t + v_t, \quad \text{where } v_{i,t} \sim N(0, \sigma_v^2) \quad (14)$$

In this case, the off-diagonal elements in $\sigma_v^2 = 0$ and the unobserved variables are related to the observed variables through the coefficient matrix G . The respective state equations are then contained in the X_t matrix, where,

$$X_{t+1} = \xi(X_t, \epsilon_{t+1}; \Xi) \quad (15)$$

To evaluate the likelihood function of the nonlinear model, $\tilde{\xi}$, for parameters $\Xi = [\mu, \Sigma_v]$, we make use of Monte Carlo methods and the particle filter that was used in Fernández-Villaverde and Rubio-Ramirez (2005). The specification of this filter may be expressed as,

$$\ell(Y_t^T | \tilde{\xi}, \Xi) = \prod_{t=1}^T \frac{1}{N} \sum_{i=1}^N p(Y_t | \tilde{x}_{t|t-1}^i; \tilde{\xi}, \Xi) \quad (16)$$

where p denotes the probability density and $\sum_{i=1}^N \tilde{x}_{t|t-1}^i$ represent draws from each density in the sequence $\prod_{t=1}^T p(X_t | Y_t^{T-1}; \Xi)$. The specification of the likelihood function for the linear model, $\bar{\xi}$, makes use of a traditional Kalman filter that is provided by Hamilton (1994), such that,

$$\begin{aligned} \ell(Y_t^T | \bar{\xi}, \Xi) &= \prod_{t=1}^T p(Y_t | Y_t^{T-1}; \bar{\xi}, \Xi) \\ &= \prod_{t=1}^T \int p(Y_t | X_t, Y_t^{T-1}; \bar{\xi}, \Xi) p(X_t | Y_t^{T-1}; \bar{\xi}, \Xi) dX_t \end{aligned} \quad (17)$$

It is worth noting that these likelihood functions are extremely complex; where portions of them are flat and the possibility of several local minima and maxima would often arise.¹⁶

¹⁵Once again, we make use of the approach for the derivation of the model solution and parameter estimation that was followed in Pichler (2008), to allow for a comparison of the results for an emerging market economy to those that were generated for a developed world economy.

¹⁶See, Canova and Sala (2009) for more on the complexity of the likelihood functions in DSGE models. Such complexities may result in potential difficulties with parameter identification.

Furthermore, when using the particle filter for the nonlinear model, the likelihood function is not continuous with respect to the parameter vector, Ξ . In cases such as this, Judd (1998) suggests that traditional gradient-based numerical optimisation techniques may be of little use when seeking to maximise the likelihood function. As an alternative, we employed the simulated annealing global optimisation approach, as in Fernández-Villaverde and Rubio-Ramirez (2005), which provided much improved results.¹⁷

Despite seeking to estimate as many of the parameters as possible, it was necessary to calibrate certain parameters in the model. Specifically, we follow Liu et al. (2009), and set the elasticity of output with respect to capital to 0.26, the depreciation rate was set to 0.019, the capital adjustment costs parameter was set to 10, and the mark-up parameter was set to 6. Furthermore, we choose the parameters corresponding to leisure and real money balances in the utility function, such that households spend 30% of their time working in the steady-state (to match the steady-state ratio between real balances and quarterly output). Finally, as in Pichler (2008), the measurement error variances were calibrated to be 10% of the variance of the respective data series.¹⁸

After estimating the remaining parameters, we are then able to generate the respective h -step ahead forecasts, $E_t[Y_{t+h}]$, given the most recent values of all variables, Y_t , and the estimated parameters values, $\hat{\Xi}$. When seeking to generate values for the linear model, one is able to use the Kalman filter to derive the expected values, $E_t[Y_{t+h}|X_t; \tilde{\xi}, \hat{\Xi}]$. Similarly, for the nonlinear model, one is able to make use of either Monte Carlo methods or numerical integration to generate future expected values, $E_t[X_{t+h}]$, with the aid of the particle filter. In this case we follow Pichler (2008) and make use of numerical integration to derive $E_t(Y_{t+h}|X_t; \tilde{\xi}, \hat{\Xi})$.

3 Vector Autoregressive Models

The competing forecasting models make use of various VAR representations for all of the reduced-form models. These models include classical unrestricted VAR models and restricted BVAR models.

3.1 Classical unrestricted VAR models

The classical unrestricted VAR models make use of the structure,

$$y_t = c + \varphi_1 y_{t-1} + \varphi_2 y_{t-1} + \dots + \varphi_p y_{t-p} + \varepsilon_t \quad (18)$$

where the lag length, p , was determined by the Schwartz-Bayes information criterion. After estimating successive models for the respective end-of-sample periods, 2000Q1 to 2010Q4, we found that each model suggested an optimal lag length of 2 periods. This maximum lag length was also applied to all the BVAR models.

3.2 Bayesian VAR models

3.2.1 Minnesota shrinkage priors

The first group of BVAR models make use of shrinkage priors that follow the work of Litterman (1986a), Litterman (1986b), Doan et al. (1984) and Sims and Zha (1998). These models

¹⁷Our initial efforts, which made use of various gradient-based techniques to solve the nonlinear DSGE models, were largely unsuccessful. Judd (1998) suggests that when solving complex models, traditional optimisation techniques that look for the perfect answer, fail to generate reasonable results in many instances.

¹⁸Real money balances is measured by M2 deflated by the consumer price inflation and covers 1965Q1-2011Q4; since M2 is only available from 1965Q1.

make use of a small selection of hyperparameters that impose various prior restrictions on the model for the estimated coefficient mean, $\hat{\varphi}$, tightness, ζ , decay, κ , and variation due to other-lagged variables, ω .¹⁹

Since the data in all the BVAR models are assumed to be stationary, the prior means follow the specification for white-noise, where all these hyperparameters (including the first own-lag) are set to zero. Hence, in the above model the prior is set such that, $\hat{\varphi} = 0$.

When seeking to specify the variance-covariance elements, $v_{ij,l}$, we impose the following specification for the priors of variable j in equation i and lag l ;

$$v_{ij,l} = \begin{cases} \zeta/\kappa l & \text{if } i = j \\ \frac{\zeta\omega\sigma_i^2}{\kappa l\sigma_j^2} & \text{if } i \neq j \end{cases}$$

In this specification, the values for the ζ hyperparameter control the degree to which the coefficient of the first lag of the dependent variable is believed to be concentrated around zero. Various values for this tightness parameter have been used, from between 0.1 and 2.0, where small values will force the own-lags of the dependent variable to be close to the prior mean.

Since it is assumed that the coefficients for more immediate lags are possibly going to be more influential, we assume that the variance for these coefficients will decrease with an increasing lag length, l . Once again, we make use of various values for the decay, where κ ranges between 0 and 2.

In addition, it is also assumed that most of the variation in each of the respective variables may be explained by the variation in their respective lags. Therefore, for the explanatory variables that are not lagged dependent variables, a smaller variance is assigned (in relative terms) by choosing a value for ω that is between 0 and 1 (where the ratio σ_i^2/σ_j^2 accounts for the differences in the variability of the respective variables).

3.2.2 BVAR models with stochastic variable selection

The specification of the BVAR models with SVS follow Koop and Korobilis (2010) and Korobilis (2011). Using the formulation of an unrestricted vector autoregressive model in (18), we allow for the respective coefficient matrices, φ , to be multiplied through by the indicator matrix, $\gamma_{i,j}$, which has elements that take on a Bernoulli distribution. Essentially, these indicator parameters determine whether or not the variable should be included in the final representation that will be used to generate the forecasts. Using a Bayesian framework, these parameters are treated as random variables, for which we assign a prior that is taken to the likelihood function to derive the final posterior values.²⁰

In addition to the basic model that employs these variable selection techniques, we also include a model where $\gamma_{i,j} = 1$ to investigate whether or not these techniques make a significant difference to the forecasting performance.²¹ Thereafter, we include a model that makes use of priors that impose hierarchical Bayesian shrinkage using least absolute shrinkage and selection operators (LASSO). Several studies have suggested that these priors have outperformed other types of hierarchical Bayesian shrinkage estimates (such as the Normal-Jeffreys priors), whilst providing comparable forecasting results to the models that employ a Minnesota prior.²² When specifying this model, we condition the coefficient matrix by assuming that the off-diagonal elements of the co-variance matrix are zero.

¹⁹Of course, the posterior estimates may override these prior restrictions if the data provides strong evidence that the prior is inappropriate.

²⁰These posterior values are generated from simulation techniques that make use of a Gibbs sampler.

²¹This model would take the form of a traditional BVAR model, which is estimated with the aid of a Gibbs sampler and a flat prior. The results from this model could be compared to the VAR, which is estimated with frequentist techniques.

²²See, Korobilis (2011) for further details on the comparative performance of LASSO and Normal-Jeffreys priors.

We have also included the results of a BVAR model with SVS and time-varying parameters. In this case the coefficients in equation (18) may be expressed as, $\varphi_{i,t} = \varphi_{i,t-1} + \vartheta_{i,t}$, where $\vartheta_{i,t} \sim N(0, \sigma_{\vartheta}^2)$. Hence, this model would allow for a degree of stochastic variation in each of the coefficients. The final BVAR model with SVS allows for an endogenous structural break. This is achieved by incorporating a restricted Markov chain, where the respective processes are able to move to a second regime (during a structural break). However, in contrast with a traditional regime-switching model that makes use of a Markov chain, the process is not able to move back into the initial regime, and as such, it starts afresh from the breakpoint in the time series.

4 Data

The respective models make use of three observed variables, namely detrended output (in logarithms), y_t , quarter-on-quarter consumer price inflation, π_t , and a measure of the nominal interest rate, i_t . The data is measured at a quarterly frequency from 1960Q1 to 2011Q4, with the start and end date of the sample being governed by data availability. To derive a measure of detrended output we took the logarithm of real gross domestic product from which we removed the linear trend. The data on the seasonally adjusted real gross domestic product at constant prices (for the year 2005) was obtained from the South African Reserve Bank.²³ The data for interest rate relates to the 3-month Treasury bill, which was obtained from the International Financial Statistics (IFS) database, which is maintained by the International Monetary Fund (IMF). Consumer price inflation was derived from the first difference of the logarithm of the consumer price index. The data on the seasonally-unadjusted Consumer Price Index was obtained from the Global Financial Database. The seasonal was removed with the aid of the X-12 procedure, which has been developed by the Department of Commerce, U.S. Census Bureau.

As discussed above, we select the time period 1960Q1 through 2011Q4 for our analysis. This gives a total sample of 208 observations on each series, where the first 160 observations (1960Q1 through 1999Q4) were used for the initial in-sample analysis, following the existing literature on forecasting for South Africa. The remaining 48 observations (2000Q1 through 2011Q4) were used for out-of-sample forecasting evaluation, over which the models are recursively estimated by increasing the size of the in-sample by one observation to produce one to eight step-ahead forecasts. Note, the choice of the starting point of the out-of-sample period coincides with South Africa's decision to move formally to an inflation-targeting regime in the February of 2000.

5 Results

In total we make use of twenty-two models that have been estimated forty-one times to generated forecasts for the period 2000Q1 to 2011Q4. When comparing these models, we consider the results for each of the 1, 2, ..., 8 step-ahead forecasts over the entire out-of-sample period.²⁴

In the following sub-sections we investigate the forecasting performance for each of the respective variables that are contained in tables (5.2) through (5.2). The first line of these tables contains the root-mean-squared error for the nonlinear DSGE model. This statistic

²³For the BVAR models, we had to use the growth rates of detrended output due to issues of convergence. So, after generating all the forecasts from the BVARs, we transform this variable back into detrended output, before calculating the evaluation statistics.

²⁴Further details of this forecasting exercise have been contained in the online appendix, which considers various other ways of summarising this data, such as an evaluation of the forecasts at each point in time, as well as the average over the first 2, 4, and 8 step-ahead forecasting horizons. From the results in the appendix one is also able to ascertain that the root-mean-squared-errors are generally larger for π_t than y_t , whilst the errors for i_t are extremely small. It is also worth noting that the findings from these additional investigations support the results that are discussed here.

is then used to calculate the relative root-mean-squared error for the other models.²⁵ The Diebold and Mariano (1995) statistic is then used to determine whether this difference in forecasting performance is statistically significant.

5.1 Consumer price inflation

The results for the π_t forecasts are contained in table (1). Note firstly, that there are only 7 negative relative root-mean-squared errors, and the largest negative value is -0.34 . In contrast, there are 161 instances where a positive relative root-mean-squared error is observed. The largest positive value is 27.97.

The models that are able to provide a comparatively lower root-mean-squared error (when compared to the nonlinear DSGE model) are the first three BVAR models with Minnesota prior. These instances occur at the 6 and 7 step ahead forecasts.²⁶ The Diebold-Mariano statistic for these 7 negative relative root-mean-squared errors are all particularly small, and as such it is not surprising to note that there are no occasions where any of the competing models are able to generate statistically significant improvements.

Furthermore, what is also worth noting is that when we compare the nonlinear DSGE model to the linear variant, the nonlinear model provides a lower root-mean-squared error at all but the one-step ahead forecast horizon. However, at the longer forecasting horizon the nonlinear DSGE is clearly superior, as the 6, 7 and 8 step-ahead Diebold-Mariano statistics are all well below negative two.²⁷

To ensure that these results have not been overly influenced by an outlying forecast that may have been generated for a particular point in time, we also count the number of significant Diebold-Mariano statistics that were generated for each forecast from 2000Q1 to 2011Q4. In this case, the number of significant Diebold-Mariano statistics at the 2 step-ahead horizon continues to favour the nonlinear DSGE model.^{28,29} Over longer horizons the results of the nonlinear DSGE are more impressive, which would confirm that it is responsible for significantly smaller forecasting errors at longer horizons.³⁰

5.2 Output

The relative root-mean square errors for the forecasts of y_t are contained in table (2). Once again, the nonlinear DSGE model provides lower root-mean-squared errors at the longer horizon, when compared to the various vector autoregressive and random walk models.³¹

Furthermore, there are a number of occasions where the Diebold-Mariano statistic suggests that the nonlinear model provides significantly better results when the forecasting

²⁵For example, the relative root-mean-squared error for the classical-VAR at a 1-step ahead forecasting horizon is calculated as $[(\text{RMSE}_{\text{classical-VAR}}/\text{RMSE}_{\text{nonlinear-DSGE}}) - 1] \times 100$, where RMSE is the 1-step ahead root-mean-squared error over 2000Q1 to 2011Q4.

²⁶The only exception arises at a one-step ahead forecasting horizon, where the root-mean-squared error of the linear DSGE model is slightly lower.

²⁷The nonlinear DSGE model is also clearly superior to that of a Random-Walk.

²⁸Further details of this analysis are contained in the online appendix.

²⁹When using this method of evaluation, the most impressive of the competing vector autoregressive models is the BVAR with Minnesota prior, which has 12 significant statistics; whereas the nonlinear DSGE has 14. Over this horizon, the results of the linear and nonlinear DSGE models are almost equivalent.

³⁰The nonlinear DSGE model is responsible for four additional significant Diebold-Mariano statistics, when compared to the vector autoregressive models at the 4 step-ahead horizon. At the 8 step-ahead horizon, the nonlinear DSGE generates five additional significant Diebold-Mariano statistics. The results for the linear and nonlinear model are similar at the 4 step-ahead horizon, but the difference in the number of significant Diebold-Mariano statistics at the 8 step-ahead horizon is eight; in favour of the nonlinear model.

³¹At the 8 step-ahead horizon the vector autoregressive and random walk models all have positive relative root-mean-squared errors.

horizon is at least a year.³² However, at the shorter horizon, some of the BVAR models with Minnesota prior (as well as the BVAR model without stochastic variable selection) appear to provide slightly better forecasts.³³

Once again, when comparing the nonlinear DSGE model with the linear counterpart, we note that the nonlinear DSGE model would appear to generate a lower loss function on most occasions. However, in this case the root mean squared errors for the nonlinear model are much lower at the short and medium-term horizon, whilst over the longer term, the results are fairly similar.³⁴ At the 1, 2 and 3 step-ahead horizons, the Diebold-Mariano statistics indicate that the forecasts of the nonlinear DSGE model are significantly better.

³²Indeed, there are 18 occasions where the Diebold-Mariano statistic indicates that the nonlinear DSGE model is able to significantly improve upon the forecasts of the various vector autoregressive models.

³³In this case, some of the BVAR models provide statistically significant improvements on six occasions.

³⁴It is also worth noting that when we count the number of significant Diebold-Mariano statistics for the forecasts that were generated at each point in time, we note that 13 statistics favour the nonlinear model, whilst 6 favour the linear model (when we consider the 8 step-ahead horizon). At the shorter horizons these statistics favour the nonlinear DSGE model by a larger degree.

	1 step	2 step	3 step	4 step	5 step	6 step	7 step	8 step
Nonlinear	0.0091	0.011	0.0113	0.0116	0.0123	0.0124	0.0121	0.0115
Linear	-0.28 [0.29]	0.24 [-0.21]	1.32 [-1.09]	2.36 [-1.41]	3.98 [-1.83]	5.73 [-2.3]**	6.9 [-2.53]**	8.77 [-2.8]**
RW	10.18 [-1.73]	18.27 [-2.46]**	20.61 [-2.44]**	21.33 [-2.52]**	27.2 [-3.49]**	27.97 [-2.94]**	24.7 [-2.28]**	17.67 [-1.76]
VAR	3.88 [-1.48]	3.39 [-0.89]	4.09 [-0.73]	5.94 [-0.92]	3.18 [-0.47]	2.44 [-0.33]	3.38 [-0.41]	6.3 [-0.65]
BVAR	3.87 [-1.49]	3.43 [-0.91]	4 [-0.72]	5.86 [-0.91]	3.13 [-0.47]	2.34 [-0.32]	3.21 [-0.39]	6.18 [-0.64]
Minnes1	3.11 [-1.52]	1.92 [-0.61]	1.72 [-0.4]	3.01 [-0.56]	0.94 [-0.16]	0.19 [0.03]	-0.23 [0.03]	1.85 [-0.2]
Minnes2	3.05 [-1.4]	1.51 [-0.47]	1.47 [-0.33]	2.89 [-0.53]	0.64 [-0.11]	-0.34 [0.05]	-0.08 [0.01]	2.24 [-0.23]
Minnes3	2.15 [-0.91]	0.77 [-0.21]	1.22 [-0.25]	2.83 [-0.48]	0.24 [-0.04]	-0.34 [0.05]	0.67 [-0.08]	3.32 [-0.32]
Minnes4	1.84 [-0.46]	0 [0]	2.02 [-0.31]	4.17 [-0.58]	0.26 [-0.04]	0.17 [-0.02]	2.28 [-0.25]	5.49 [-0.48]
Minnes5	0.82 [-0.32]	0.31 [-0.08]	1.39 [-0.25]	3.06 [-0.48]	0.11 [-0.02]	-0.06 [0.01]	1.63 [-0.19]	4.44 [-0.41]
Minnes6	1.01 [-0.23]	0.45 [-0.07]	3.15 [-0.43]	5.2 [-0.66]	0.63 [-0.08]	0.58 [-0.07]	2.91 [-0.31]	6.1 [-0.52]
Minnes7	3.56 [-1.24]	2.74 [-0.66]	3.84 [-0.66]	5.91 [-0.9]	2.98 [-0.44]	2.59 [-0.35]	4.04 [-0.48]	7.17 [-0.72]
Minnes8	2.88 [-0.85]	2.15 [-0.44]	4 [-0.63]	6.31 [-0.91]	2.99 [-0.43]	3.02 [-0.4]	5.15 [-0.59]	8.5 [-0.82]
Minnes9	3.83 [-0.68]	2.46 [-0.36]	5.79 [-0.74]	8.29 [-1.05]	3.56 [-0.47]	3.82 [-0.47]	6.62 [-0.72]	10.19 [-0.92]
Minnes10	1.85 [-0.5]	2.07 [-0.37]	4.63 [-0.67]	7 [-0.96]	3.25 [-0.45]	3.58 [-0.46]	6.19 [-0.69]	9.58 [-0.89]
Minnes11	3.33 [-0.55]	3.25 [-0.42]	6.93 [-0.82]	9.2 [-1.1]	3.86 [-0.49]	4.07 [-0.5]	6.95 [-0.74]	10.48 [-0.93]
Minnes12	3.86 [-1.48]	3.39 [-0.89]	4.09 [-0.73]	5.94 [-0.92]	3.18 [-0.47]	2.44 [-0.33]	3.39 [-0.41]	6.31 [-0.65]
Minnes13	3.62 [-1.04]	1.88 [-0.42]	3.26 [-0.55]	5.19 [-0.81]	2.06 [-0.32]	1.69 [-0.24]	2.96 [-0.36]	5.6 [-0.56]
SVS	1.35 [-0.53]	1.27 [-0.31]	2.5 [-0.45]	4.53 [-0.71]	1.65 [-0.25]	1.24 [-0.17]	2.5 [-0.29]	5.46 [-0.52]
SVS-TVP	1.69 [-0.67]	1.87 [-0.46]	3.19 [-0.57]	4.36 [-0.69]	0.72 [-0.11]	0.81 [-0.11]	2 [-0.24]	4.51 [-0.43]
SVS-SB	1.95 [-0.8]	2 [-0.51]	2.57 [-0.48]	4.11 [-0.68]	1.5 [-0.24]	0.55 [-0.08]	1.91 [-0.24]	5.05 [-0.52]
SVS-LAS	1.77 [-0.68]	1.86 [-0.45]	3.1 [-0.55]	5.08 [-0.78]	1.91 [-0.29]	1.49 [-0.2]	3.05 [-0.36]	6.13 [-0.59]

Table 1: Consumer Inflation - Root Mean Square Errors and Diebold & Mariano Statistics (2000Q1 - 2011Q4). Nonlinear DSGE model absolute RMSE and competing models relative RMSE with corresponding DM statistics in parenthesis. ** significant in favour of Nonlinear DSGE. * significant in favour of competing model.³⁵

³⁵Model acronyms: Nonlinear - Nonlinear DSGE model; Linear - Linear DSGE model; RW - Random Walk model; VAR - classical VAR; BVAR - Bayesian VAR; Minnes1 - BVAR with Minnesota prior ($\zeta = 2, \kappa = 2, \omega = 0.001$); Minnes2 - ($\zeta = 0.3, \kappa = 0.5, \omega = 0.001$); Minnes3 - ($\zeta = 0.2, \kappa = 1, \omega = 0.001$); Minnes4 - ($\zeta = 0.1, \kappa = 1, \omega = 0.001$); Minnes5 - ($\zeta = 0.2, \kappa = 2, \omega = 0.001$); Minnes6 - ($\zeta = 0.1, \kappa = 2, \omega = 0.001$); Minnes7 - ($\zeta = 0.3, \kappa = 0.5, \omega = 0.05$); Minnes8 - ($\zeta = 0.2, \kappa = 1, \omega = 0.05$); Minnes9 - ($\zeta = 0.1, \kappa = 0.1, \omega = 0.05$); Minnes10 - ($\zeta = 0.2, \kappa = 2, \omega = 0.05$); Minnes11 - ($\zeta = 0.1, \kappa = 2, \omega = 0.05$); Minnes12 - ($\zeta = 2, \kappa = 0, \omega = 1$); Minnes13 - ($\zeta = 0.3, \kappa = 0.5, \omega = 0.5$); SVS - BVAR with SVS; SVS-TVP - BVAR with SVS and time varying parameters; SVS-SB - BVAR with SVS and endogenous structural break; SVS-LAS - BVAR with SVS and LASSO prior.

	1 step	2 step	3 step	4 step	5 step	6 step	7 step	8 step
Nonlinear	0.0065	0.0111	0.0144	0.017	0.0194	0.0219	0.0245	0.0272
Linear	77.76	54.04	33.31	20.02	12.2	6.24	2.03	-1.58
	<i>[-4.03]**</i>	<i>[-3.87]**</i>	<i>[-3.07]**</i>	<i>[-1.96]</i>	<i>[-1.25]</i>	<i>[-0.7]</i>	<i>[-0.23]</i>	<i>[0.19]</i>
RW	4.18	13.02	21.33	28.62	32.51	32.99	32.15	30.23
	<i>[-0.47]</i>	<i>[-1.47]</i>	<i>[-2.22]**</i>	<i>[-2.69]**</i>	<i>[-2.92]**</i>	<i>[-2.81]**</i>	<i>[-2.55]**</i>	<i>[-2.27]**</i>
VAR	-5.88	-9.99	-8.47	-7.06	-5.69	-4.67	-2.31	0.04
	<i>[1.64]</i>	<i>[2.55]</i>	<i>[2.45]</i>	<i>[1.56]</i>	<i>[0.93]</i>	<i>[0.62]</i>	<i>[0.28]</i>	<i>[0]</i>
BVAR	-6	-10.21	-8.71	-7.26	-5.84	-4.82	-2.45	-0.14
	<i>[1.69]</i>	<i>[2.59]*</i>	<i>[2.5]*</i>	<i>[1.59]</i>	<i>[0.95]</i>	<i>[0.64]</i>	<i>[0.29]</i>	<i>[0.02]</i>
Minnes1	-8.32	-2.68	8.06	16.08	20.74	21.77	21.43	19.48
	<i>[0.82]</i>	<i>[0.27]</i>	<i>[-0.82]</i>	<i>[-1.46]</i>	<i>[-1.82]</i>	<i>[-1.96]</i>	<i>[-1.99]</i>	<i>[-1.84]</i>
Minnes2	-7.98	-1.93	8.77	16.83	21.46	22.45	22.04	20.03
	<i>[0.8]</i>	<i>[0.2]</i>	<i>[-0.9]</i>	<i>[-1.53]</i>	<i>[-1.89]</i>	<i>[-2.03]**</i>	<i>[-2.05]**</i>	<i>[-1.89]</i>
Minnes3	-7.27	-0.12	10.12	18.05	22.39	23.18	22.57	20.51
	<i>[0.77]</i>	<i>[0.01]</i>	<i>[-1.08]</i>	<i>[-1.69]</i>	<i>[-2.03]**</i>	<i>[-2.16]**</i>	<i>[-2.14]**</i>	<i>[-1.94]</i>
Minnes4	-4.8	3.58	12.9	20.36	24.16	24.57	23.63	21.42
	<i>[0.57]</i>	<i>[-0.44]</i>	<i>[-1.43]</i>	<i>[-1.97]</i>	<i>[-2.27]**</i>	<i>[-2.36]**</i>	<i>[-2.28]**</i>	<i>[-2.04]**</i>
Minnes5	-5.79	3.07	12.52	20.08	23.95	24.41	23.51	21.31
	<i>[0.66]</i>	<i>[-0.37]</i>	<i>[-1.38]</i>	<i>[-1.93]</i>	<i>[-2.24]**</i>	<i>[-2.34]**</i>	<i>[-2.26]**</i>	<i>[-2.03]**</i>
Minnes6	-3.31	6.13	14.75	21.8	25.26	25.42	24.29	21.98
	<i>[0.42]</i>	<i>[-0.8]</i>	<i>[-1.68]</i>	<i>[-2.14]**</i>	<i>[-2.41]**</i>	<i>[-2.48]**</i>	<i>[-2.36]**</i>	<i>[-2.1]**</i>
Minnes7	-6.76	-9.97	-8.15	-6.27	-4.73	-3.71	-1.49	0.62
	<i>[1.84]</i>	<i>[2.44]*</i>	<i>[2.2]*</i>	<i>[1.3]</i>	<i>[0.75]</i>	<i>[0.49]</i>	<i>[0.18]</i>	<i>[-0.07]</i>
Minnes8	-6.36	-8.71	-6.61	-4.2	-2.45	-1.43	0.56	2.34
	<i>[1.72]</i>	<i>[2.15]*</i>	<i>[1.68]</i>	<i>[0.81]</i>	<i>[0.37]</i>	<i>[0.18]</i>	<i>[-0.07]</i>	<i>[-0.26]</i>
Minnes9	-7.41	-6.9	-3.55	0.27	2.66	3.83	5.43	6.55
	<i>[1.67]</i>	<i>[1.49]</i>	<i>[0.72]</i>	<i>[-0.04]</i>	<i>[-0.36]</i>	<i>[-0.46]</i>	<i>[-0.61]</i>	<i>[-0.7]</i>
Minnes10	-3.95	-6.16	-4.52	-2.05	-0.34	0.56	2.32	3.9
	<i>[1.18]</i>	<i>[1.76]</i>	<i>[1.2]</i>	<i>[0.4]</i>	<i>[0.05]</i>	<i>[-0.07]</i>	<i>[-0.27]</i>	<i>[-0.43]</i>
Minnes11	-5.7	-4.97	-1.88	1.95	4.28	5.3	6.67	7.6
	<i>[1.42]</i>	<i>[1.17]</i>	<i>[0.39]</i>	<i>[-0.31]</i>	<i>[-0.58]</i>	<i>[-0.65]</i>	<i>[-0.76]</i>	<i>[-0.82]</i>
Minnes12	-6.02	-10.03	-8.51	-7.08	-5.71	-4.69	-2.32	0.03
	<i>[1.67]</i>	<i>[2.55]*</i>	<i>[2.45]*</i>	<i>[1.56]</i>	<i>[0.93]</i>	<i>[0.62]</i>	<i>[0.28]</i>	<i>[0]</i>
Minnes13	-8.4	-8.92	-4.17	-0.6	1.38	2	2.95	3.24
	<i>[1.47]</i>	<i>[1.41]</i>	<i>[0.72]</i>	<i>[0.09]</i>	<i>[-0.19]</i>	<i>[-0.26]</i>	<i>[-0.37]</i>	<i>[-0.4]</i>
SVS	-10.1	-5.94	-1.73	2.43	4.75	5.57	6.91	7.97
	<i>[1.43]</i>	<i>[0.99]</i>	<i>[0.29]</i>	<i>[-0.35]</i>	<i>[-0.6]</i>	<i>[-0.64]</i>	<i>[-0.75]</i>	<i>[-0.83]</i>
SVS-TVP	-10.59	-5.83	-1.75	2.36	4.71	5.47	6.75	7.91
	<i>[1.68]</i>	<i>[1.1]</i>	<i>[0.31]</i>	<i>[-0.33]</i>	<i>[-0.57]</i>	<i>[-0.61]</i>	<i>[-0.72]</i>	<i>[-0.81]</i>
SVS-SB	-9.44	-8.28	-5.33	-2.35	-0.39	0.5	2.2	3.89
	<i>[1.86]</i>	<i>[1.74]</i>	<i>[1.1]</i>	<i>[0.38]</i>	<i>[0.05]</i>	<i>[-0.06]</i>	<i>[-0.25]</i>	<i>[-0.41]</i>
SVS-LAS	-10.46	-7.06	-3.29	0.48	2.7	3.52	4.98	6.15
	<i>[1.63]</i>	<i>[1.29]</i>	<i>[0.61]</i>	<i>[-0.07]</i>	<i>[-0.36]</i>	<i>[-0.42]</i>	<i>[-0.57]</i>	<i>[-0.67]</i>

Table 2: Output - Root Mean Square Errors and Diebold & Mariano Statistics (2000Q1 - 2011Q4). Nonlinear DSGE model absolute RMSE and competing models relative RMSE with corresponding DM statistics in parenthesis. ** significant in favour of Nonlinear DSGE. * significant in favour of competing model.³⁶

³⁶Model acronyms: See footnote 35.

	1 step	2 step	3 step	4 step	5 step	6 step	7 step	8 step
Nonlinear	0.0019	0.0032	0.0045	0.0055	0.0064	0.0071	0.0074	0.0075
Linear	-3.36	-0.97	1.03	2.7	4.57	6.3	8.5	10.91
	[2.4]*	[0.49]	[-0.5]	[-1.16]	[-1.66]	[-2.00]**	[-2.34]**	[-2.56]**
RW	-0.21	5.01	0.6	-3.26	-6.32	-7.74	-6.7	-3.36
	[0.01]	[-0.24]	[-0.04]	[0.27]	[0.59]	[0.77]	[0.68]	[0.35]
VAR	-25.49	-10.34	-9.26	-10.95	-12.55	-13.95	-14.64	-13.62
	[1.61]	[0.66]	[0.76]	[1.18]	[1.62]	[1.92]	[1.97]	[1.79]
BVAR	-25.5	-10.4	-9.34	-11.04	-12.58	-13.96	-14.58	-13.55
	[1.61]	[0.67]	[0.77]	[1.18]	[1.62]	[1.93]	[1.97]	[1.79]
Minnes1	-20.5	-6.92	-8.05	-11.01	-13.27	-14.74	-14.96	-13.07
	[1.22]	[0.4]	[0.59]	[1.03]	[1.44]	[1.72]	[1.75]	[1.52]
Minnes2	-17.96	-5.65	-7.38	-10.48	-12.96	-14.43	-14.28	-12.03
	[1.02]	[0.31]	[0.52]	[0.94]	[1.35]	[1.6]	[1.6]	[1.35]
Minnes3	-12.57	-2.5	-5.39	-8.93	-11.83	-13.4	-12.95	-10.38
	[0.67]	[0.13]	[0.36]	[0.77]	[1.17]	[1.41]	[1.38]	[1.12]
Minnes4	-0.92	2.77	-2.83	-7.53	-11.25	-13.18	-12.65	-9.94
	[0.05]	[-0.14]	[0.19]	[0.63]	[1.07]	[1.32]	[1.28]	[1.03]
Minnes5	-7.27	0.51	-3.4	-7.31	-10.47	-12.09	-11.45	-8.65
	[0.38]	[-0.03]	[0.22]	[0.61]	[1.01]	[1.23]	[1.18]	[0.9]
Minnes6	-0.88	2.8	-2.81	-7.51	-11.23	-13.16	-12.63	-9.92
	[0.04]	[-0.14]	[0.19]	[0.63]	[1.07]	[1.32]	[1.28]	[1.03]
Minnes7	-24.09	-10.31	-9.79	-11.5	-13.17	-14.35	-14.39	-12.7
	[1.42]	[0.61]	[0.76]	[1.17]	[1.59]	[1.85]	[1.84]	[1.61]
Minnes8	-20.27	-8.16	-8.64	-10.68	-12.67	-13.82	-13.39	-11.17
	[1.12]	[0.45]	[0.63]	[1.03]	[1.43]	[1.65]	[1.61]	[1.35]
Minnes9	-7.74	-1.78	-5.46	-9.02	-12.07	-13.66	-13.08	-10.53
	[0.4]	[0.09]	[0.38]	[0.81]	[1.24]	[1.47]	[1.42]	[1.16]
Minnes10	-16.97	-5.96	-7.15	-9.46	-11.66	-12.8	-12.1	-9.57
	[0.91]	[0.32]	[0.51]	[0.88]	[1.26]	[1.47]	[1.4]	[1.11]
Minnes11	-7.68	-1.55	-5.31	-8.95	-12.05	-13.66	-13.05	-10.47
	[0.4]	[0.08]	[0.37]	[0.8]	[1.22]	[1.46]	[1.41]	[1.15]
Minnes12	-25.49	-10.35	-9.27	-10.96	-12.56	-13.96	-14.64	-13.61
	[1.61]	[0.66]	[0.76]	[1.18]	[1.62]	[1.92]	[1.97]	[1.79]
Minnes13	-10.95	-12.11	-13.9	-15.83	-17.96	-19.41	-19.88	-18.97
	[0.73]	[0.92]	[1.45]	[2.03]*	[2.46]*	[2.68]*	[2.6]*	[2.37]*
SVS	-20.87	-7.4	-8.68	-11.75	-14.06	-15.61	-15.9	-14.11
	[1.25]	[0.43]	[0.64]	[1.11]	[1.54]	[1.83]	[1.86]	[1.65]
SVS-TVP	-18.49	-5.02	-6.51	-10.89	-14.3	-15.54	-15.32	-13.14
	[1.08]	[0.28]	[0.46]	[0.95]	[1.39]	[1.61]	[1.6]	[1.35]
SVS-SB	-21.04	-7.28	-8.3	-11.21	-13.43	-14.69	-14.58	-12.47
	[1.24]	[0.41]	[0.6]	[1.02]	[1.41]	[1.64]	[1.63]	[1.38]
SVS-LAS	-21.25	-7.83	-8.86	-11.75	-13.99	-15.47	-15.71	-13.93
	[1.26]	[0.45]	[0.66]	[1.12]	[1.55]	[1.84]	[1.87]	[1.65]

Table 3: Interest Rates - Root Mean Square Errors and Diebold & Mariano Statistics (2000Q1 - 2011Q4). Nonlinear DSGE model absolute RMSE and competing models relative RMSE with corresponding DM statistics in parenthesis. ** significant in favour of Nonlinear DSGE. * significant in favour of competing model.³⁷

³⁷Model acronyms: See footnote 35.

5.3 Interest rates

An evaluation of the selected models ability to forecast i_t is contained in table (3). In this case, most of the relative-root-mean-squared errors would indicate that competing models would often generate a lower forecasting error. However, with the exception of one of the BVAR models, these improvements are usually insignificant.^{38,39}

When comparing the results for the nonlinear DSGE model with the linear counterpart, we note that the linear model appears to provide forecasts that are significantly superior at the 1 step-ahead horizon. However, as the horizon increases the nonlinear model starts to perform significantly better; particularly for 6, 7 and 8 step-ahead forecasts. In addition, when we compare the individual forecasts that were generated for each point in time, we note that the results for the 2 and 4 step-ahead horizon are extremely similar; whilst the results at the 8 step-ahead horizon would indicate that the nonlinear model is clearly superior.⁴⁰

6 Conclusion

The results suggest that the forecasting performance of the nonlinear DSGE model is at least comparable, and in many cases superior, to that of an equivalent linear model for South African macroeconomic data. The improvements are statistically significant at the medium to longer horizon, for inflation and interest rates, and at shorter horizons for output growth. Hence, given the improved out-of-sample fit of the nonlinear DSGE model (over the linear counterpart), these results suggest that there are important nonlinear features (or relatively large departures from the steady state) in underlying data generating process of this emerging market economy.

This finding is in contrast to the result of Pichler (2008), which suggests that the incorporation of nonlinear features in the model solution does not improve upon the out-of-sample fit of such models, when applied to data from the United States economy. Indeed, our results suggest that when seeking to make use of a structural macroeconomic model to inform policy in an emerging market economy, one should possibly seek to incorporate potentially important nonlinearities in the model solution.

In addition, the results of this investigation also suggest that the nonlinear DSGE model is also able to improve upon the forecasting performance of most reduced-form vector autoregressive and random walk models. However, there are a number of instances where certain BVAR models may provide significant improvements (when forecasting output and interest rates). This only occurs at the shorter horizon, when forecasting output. When forecasting inflation at most horizons and output at longer horizons, the nonlinear DSGE model is often superior.

³⁸One of the BVAR models that makes use of a Minnesota prior is able to generate significantly better forecasts when the horizon is at least a year.

³⁹One reason as to why so many of these improvements are insignificant, is that the forecasting errors for i_t are mostly extremely small for all models. As such an improvement of 0.01 over 0.11 generates quite a large relative root-mean-squared error, which is in most cases insignificant. The small forecasting errors can largely be attributed to the high degree of persistence in this variable. For example, when the parameters in a Taylor rule are estimated for the South African economy, the smoothing coefficient (which in this case is ϕ_i) is usually above 0.9. See, Alpanda, Kotzé and Woglom (2010) for further details on the values of the estimated Taylor rule coefficients in structural macroeconomic models.

⁴⁰On 16 occasions the nonlinear DSGE model generates significantly better forecasts (as measured by the Diebold-Mariano statistic), when considering the 8 step-ahead horizon. This contrasts with the 9 occasions where the linear model is responsible for improvements that are statistically significant.

References

- Adjemian, S., H. Bastani, M. Juillard, F. Karame, F. Mihoubi, G. Perendia, J. Pfeifer, M. Ratto, and S. Villemot, “Dynare: Reference Manual, Version 4.,” Dynare Working Papers, 1 CEPREMAP 2011.
- Alpanda, Sami, Kevin Kotzé, and Geoffrey Woglom, “Should Central Banks of Small Open Economies Respond to Exchange Rate Fluctuations? The Case of South Africa,” *ERSA Working Paper*, 2010, No. 174.
- , Kevin Kotzé, and Geoffrey Woglom, “Forecasting Performance Of An Estimated Dsge Model For The South African Economy,” *South African Journal of Economics*, 2011, 79 (1), 50–67.
- An, Sungbae, “Bayesian Estimation of DSGE Models: Lessons from Second-Order Approximations,” Technical Report, Singapore Management University 2008.
- and Frank Schorfheide, “Bayesian Analysis of DSGE Models,” *Econometric Reviews*, 2007, 26, 211–219.
- Blanchard, Olivier J. and Charles M. Kahn, “The Solution of Linear Difference Models under Rational Expectations,” *Econometrica*, 1980, pp. 48(5):1305–1311.
- Canova, Fabio and Luca Sala, “Back to Square one: Identification issues in DSGE models,” *Journal of Monetary Economics*, 2009, 56 (4), 431–449.
- Carabenciov, Ioan, Igor Ermolaev, Charles Freedman, Michel Juillard, Ondra Kamenik, Dmitry Korshunov, Douglas Laxton, and Jared Laxton, “A Small Multi-Country Projection Model,” Technical Report 08/279, IMF working paper 2008.
- DeJong, David N. and Chetan Dave, *Structural Macroeconometrics*, 2nd ed., Princeton: Princeton University Press, 2011.
- Diebold, Francis X. and Roberto S. Mariano, “Predictive Accuracy,” *Journal of Business and Economic Statistics*, 1995, 13 (3), 253–263.
- Doan, Thomas A., Robert Litterman, and Christopher A. Sims, “Forecasting and conditional projection using realistic prior distributions,” *Econometric Reviews*, 1984, 3 (1), 1–100.
- Doucet, Arnaud, Simon Godsill, and Christophe Andrieu, “On sequential Monte Carlo sampling methods for Bayesian filtering,” *Statistics and computing*, 2000, 10 (3), 197–208.
- Edge, Rochelle M., Michael T. Kiley, and Jean-Philippe Laforte, “A comparison of forecast performance between Federal Reserve staff forecasts, simple reduced-form models, and a DSGE model,” *Journal of Applied Econometrics*, 2010, 25 (4), 720–754.
- Fernández-Villaverde, Jesús and Juan F. Rubio-Ramirez, “Estimating Dynamic Equilibrium Economies: Linear versus Nonlinear Likelihood,” *Journal of Applied Econometrics*, 2005.
- Gupta, Rangan and Alain Kabundi, “Forecasting Macroeconomic Variables in a Small Open Economy: A Comparison between Small- and Large-Scale Models,” *Journal of Forecasting*, 2010, 29 (1-2), 168–185.
- and —, “A Large Factor Model for Forecasting Macroeconomic Variables in South Africa,” *International Journal of Forecasting*, 2011, 27 (4), 1076–1088.
- and Rudi Steinbach, “A DSGE-VAR Model for Forecasting Key South African Macroeconomic Variables,” *Economic Modelling*, forthcoming.
- Hamilton, James D., *Time Series Analysis*, Princeton: Princeton University Press, 1994.

- International Monetary Fund**, *World Economic Outlook: Transitions and Tensions*, Washington D.C.: International Monetary Fund, October 2013.
- Judd, Kenneth L.**, *Numerical Methods in Economics* MIT Press Books, The MIT Press, June 1998.
- Kitagawa, Genshiro and Will Gersch**, *Smoothness priors analysis of time series*, Vol. 116, Springer Verlag, 1996.
- Klein, Paul**, “Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model,” *Journal of Economic Dynamics and Control*, 2000, 24 (1), 405–423.
- Koop, Gary and Dimitris Korobilis**, “Bayesian Multivariate Time Series Methods for Empirical Macroeconomics,” *Foundations and Trends in Econometrics*, 2010, 3, 267–358.
- Korobilis, Dimitris**, “VAR Forecasting Using Bayesian Variable Selection,” *Journal of Applied Econometrics*, 2011.
- Litterman, Robert**, “Forecasting with Bayesian Vector Autoregressions: Five Years of Experience,” *Journal of Business & Economic Statistics*, 1986, 4 (1), 25–38.
- , “A Statistical Approach to Economic Forecasting,” *Journal of Business & Economic Statistics*, 1986, 4 (1), 1–4.
- Liu, Guangling and Rangan Gupta**, “A Small-Scale DSGE Model for Forecasting the South African Economy,” *South African Journal of Economics*, 2007, 75 (2).
- , —, and **Eric Schaling**, “A New-Keynesian DSGE Model for Forecasting the South African Economy,” *Journal of Forecasting*, 2009, 28, 387–404.
- , —, and —, “Forecasting the South African Economy: A hybrid-DSGE Approach,” *Journal of Economic Studies*, 2010, 37 (2), 181–195.
- Pichler, Paul**, “Forecasting with DSGE Models: The Role of Nonlinearities,” *The B. E. Journal of Macroeconomics*, 2008, 8 (1).
- Primiceri, Giorgio and Alejandro Justiniano**, “The Time Varying Volatility of Macroeconomic Fluctuations,” *American Economic Review*, 2008, 98 (3), 604–641.
- Ratto, Marco, Werner Roeger, and Jan in ’t Veld**, “QUEST III: an estimated DSGE model of the euro area with fiscal and monetary policy,” European Economy - Economic Papers 335, Directorate General Economic and Monetary Affairs, European Commission July 2008.
- Rotemberg, Julio J.**, “Monopolistic Price Adjustment and Aggregate Output,” *Review of Economic Studies*, 1982, 49, 517–31.
- Schmitt-Grohé, Stephanie and Martin Uribe**, “Solving Dynamic General Equilibrium Models Using a Second Order Approximation of the Policy Function,” *Journal of Economic Dynamics and Control*, 2004, 28, 755–75.
- Sims, Christopher A.**, “Solving Linear Rational Expectations Models,” *Computational Economics*, 2001, 20, 1–20.
- and **Tao Zha**, “Bayesian methods for dynamic multivariate models,” *International Economic Review*, 1998, 39 (4), 949–968.
- Smets, Frank and Rafael Wouters**, “Shocks and Frictions in US Business Cycles: A Bayesian Approach,” *European Central Bank Working Papers Series*, 2007, 722.

- Steinbach, Rudi, Patience Mathuloe, and Ben W. Smit**, “An Open Economy New Keynesian DSGE Model of the South African Economy,” *South African Journal of Economics*, 2009, 77 (2), 207–227.
- Tovar, Camilo E.**, “DSGE Models and Central Banks,” *Economics - The Open-Access, Open-Assessment E-Journal*, 2009, 3 (16), 1–31.
- Uhlig, Harald**, “A Toolkit for Analyzing Nonlinear Dynamic Stochastic Models Easily,” in Ramon Marimon and Andrew Scott, eds., *Computational Methods for the Study of Dynamic Economies*, Oxford University Press 1999, pp. 30–61.