# FORECASTING SOUTH AFRICAN MACROECONOMIC DATA WITH A NONLINEAR DSGE MODEL

Mehmet Balcilar*, Rangan Gupta† & Kevin Kotzé‡

March 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 (which includes the Minnesota prior as well). The structure of the DSGE models follows that of New-Keynesian varieties, which allow for several nominal and real rigidities. The nonlinear DSGE model makes use of the second-order solution method of **?** and a particle filter to generate values for the unobserved variables. Most of the parameters in the models are estimated using maximum likelihood techniques. The models are applied to South African macroeconomic data, with an initial in-sample period of 1960Q1 to 1999Q4. 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 significantly superior to that of it's 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

*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 amount 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 developing-world economies, such as South Africa, the forecasting performance of these models has been somewhat mixed. Initial studies by **?**, **?**, **?**, **?** indicate that Bayesian vector autoregressive (BVAR) models tend to provide more superior forecasts for critical macroeconomic variables relative to the DSGE models. However, studies by **?**, **?**, **?** and **?** point out that when one allows for open-economy features and various kinds of rigidities, DSGE models are able to compete favorably with BVAR models. Against this backdrop, the objective of this paper is to analyze 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.

The decision to use a closed-economy model for South Africa, when small open-economy features have tended to improve forecast performances of key variables, is due to the computational burden associated with estimating a small open-economy nonlinear DSGE model. Furthermore, of the seven types of rigidities that are included in the small open-economy New-Keynesian DSGE-VAR model of **?**, the rigidity that is responsible for the most significant improvement in forecasting ability, is domestic price stickiness, which is included in the closed-economy model that is described below.[3,4]

The majority of DSGE models that are used for forecasting purposes usually apply a first-order linear approximation to a theoretical model that incorporates several nonlinear features and a number of forward-looking expressions.[5] After applying such a log-linear approximation, one is then able 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).[6] 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.[7]

The use of DSGE models that are estimated with higher-order approximations and nonlinear filters is not that widespread.[8] Early advocates of this procedure include, **?**, **?**, **?**, **?**,

[1]See, **?** for an overview of the use of these models in central banks. **?** provides details of the DSGE model that has been used at the Federal Reserve Bank, whilst **?** 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 **?**.

[2]See, **?** for an early indication of the forecasting performance of these models.

[3]The rigidities that are include in the DSGE-VAR model of **?** are 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.

[4]Given the results of this paper, future research into the estimation and forecasting of a small open-economy nonlinear DSGE models for emerging economies, such as South Africa, may provide extremely interesting results.

[5]The proliferation of forecasting models that makes use of first-order approximations has also been facilitated by the excellent software platform, *Dynare*.

[6]Solution methods for linear rational expectations problems are provided by, **?**, **?**, **?**, **?**, amongst others.

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

[8]**?** suggests that this is largely due to computational complexities that are involved with these approximations. DSGE models that make use of higher-order approximations are often referred to as nonlinear DSGE models.

and ?. These studies usually employ the second-order solution method proposed by ? before utilising a particle filter to derive the likelihood function.[9] Early investigations into the forecasting performance of these nonlinear models is provided in ?; 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 macroeconomic data from an emerging 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 economy, the nonlinear DSGE model may provide a superior out-of-sample fit, when compared to it's linear counterpart

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.[10] 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 it's 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.[11] The nonlinear DSGE model also appears to outperform the VAR and BVAR models when forecasting consumer inflation.[12] In addition, when forecasting output, the nonlinear DSGE model would appear to provide significantly better forecasts over longer horizons; 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 represented by a single household, intermediate producer, final good producer, and central

[9]Nonlinear 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 ?, ?, ?, and others. Whilst most of the research that makes use of nonlinear DSGE models utilise a particle filter to derive the likelihood function, ? suggest that the use of the Efficient-Information-Sampling filter may lead to improved results; when applied to structural macroeconomic models.

[10]The specification of BVAR models with a Minnesota prior is discussed in ?, ?, ? and ?. The application of SVS techniques in a BVAR model is described in ? and ?.

[11]These forecasts are evaluated after calculating the relative-root-mean-squared errors and statistics of ?, which consider the significance of any improvement.

[12]The BVAR with Minnesota prior appears to provide the second best results in this instance

bank.

Household's maximize 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,

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

where, $\beta$ represents the subjective time discount factor, $\theta$ represents a demand shock, $c_t$ represents consumption, $\tau$ represents the household's preference for consumption, $M_t/P_t$ represents real monetary balances, $\chi_m$ represents the household's preference for monetary holdings, $h_t$ represents leisure, and $\eta_h$ represents the household's preference for leisure. It is presumed that the demand shock follows an autoregressive structure, such that,

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

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

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

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, $\delta$ represents the 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,

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

The representative firm that is 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 amount of finished goods that are produced is given as,

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

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

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

where the demand for each intermediate good is given by,

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

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

$$y_t(j) = k_t(j)\alpha\Big(z_t h_t(j)\Big)^{1-\alpha} \tag{8}$$

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

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

Sticky-prices are introduced through quadratic functions that describe the cost of adjusting prices. The specification follows **?**, where,

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

where $\bar{\pi}$ denotes the steady-state value for inflation and $\psi_p$ represents the size of adjustment costs.

At the end of period $t$ the firm distributes profits to the representative household through dividend payments, where,

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

The objective of each firm is then to maximize it's total market value, for which we construct the Lagrangian,

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

Finally, to close the model, we assume that the central bank conducts monetary policy by following a Taylor rule that may be described as,

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

## 2.2 Model solution, likelihood functions and parameter estimates

The nonlinear DSGE model is solved using second order perturbation methods, as described by **?**. For comparative purposes we also make use of the first order approximation method of **?**, which is used to approximate a linear DSGE model. The likelihood function may then be constructed with the following measurement equation in a state-space representation,

$$Y_t \;\; = \;\; GX_t + \upsilon_t, \quad \text{where} \;\; \upsilon_{i,t} \sim N(0, \sigma_\upsilon^2) \tag{14}$$

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

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

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 **?**. The specification of this filter may be expressed as,

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

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 $\prod_{t=1}^{T}p\left(X_t|Y_t^{T-1};\Xi\right)$. The specification of the likelihood function for the linear model, $\bar{\xi}$, makes use of a traditional Kalman filter that is provided by **?**, such that,

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

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.[13] Furthermore, when using the particle filter for the nonlinear model, the likelihood function is not continuous with respect to the parameter vector, $\Xi$. In cases such as this, **?** suggests that traditional gradient-based numerical optimization techniques may be of little use when seeking to maximize the likelihood function. As an alternative, we employed the simulated annealing global optimization approach, which provided much improved results.[14]

Despite seeking to estimate as many of the parameters as possible, it was necessary to calibrated certain parameters in the model. Specifically, we follow **?**, 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 the household spends 30% of it's time working in the steady-state (to match the steady-state ratio between real balances and quarterly output). Finally, as in **?**, the measurement error variances was calibrated to be 10% of the variance of the respective data series.[15]

After estimating the remaining parameters we are able to forecast generate the respective forecasts for the model, whereby we are seeking to generate a future value of $Y_{t+h}$ given the most recent values of all variables, as well as the estimated parameters values $\hat{\Xi}$ and model structure, $\xi$ which may be linear or nonlinear. When seeking to generate values for the linear model, $\bar{\xi}$, one is able to use the Kalman filter to derive, $E_t\left(Y_{t+h}|X_t;\bar{\xi},\hat{\bar{\Xi}}\right)$. One is able to derive similar values for the particle filter by making use of either Monte Carlo methods or numerical integration. In this we make use of numerical integration to derive $E_t\left(Y_{t+h}|X_t;\tilde{\xi},\hat{\bar{\Xi}}\right)$ as it was less time intensive.

[13] See, **?** for more on the complexity of the likelihood functions in DSGE models. Such complexities may result in potential difficulties with parameter identification.

[14] Our initial efforts, which made use of various gradient-based techniques to solve the nonlinear DSGE models, were largely unsuccessful. **?** suggests that when solving complex models, traditional optimization techniques that look for the perfect answer, fail to generate reasonable results in many instances.

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

# 3 Vector Autoregressive Models

The competing forecasting models make use of a vector autoregression functional form. These models include classical unrestricted VAR models and restricted BVAR models.

## 3.1 Classical unrestricted VAR models

Using a classical unrestricted VAR structure,

$$y_t = c + \varphi_1 y_{t-1} + \varphi_2 y_{t-1} + \cdots + \varphi_p y_{t-p} + \varepsilon_t \tag{18}$$

the lag length, $p$, for the recursive models 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 of the various 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 **?**, **?**, **?** and **?**. These models make use of a small selection of hyperparameters that impose various prior restrictions on the model for the estimated coefficient mean, $\hat{\bar{\varphi}}$, tightness, $\zeta$, decay, $\kappa$, and variation due to other-lagged variables, $\omega$.[16]

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) equal zero. Hence, in the above model the prior is set such that, $\hat{\bar{\varphi}} = 0$.

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

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

In this specification, the values for the $\zeta$ 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, ranging 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 $\kappa$ 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 of their respective lags. Therefore, for explanatory variables that are not lagged dependent variables, a smaller variance is assigned in relative terms by choosing a value for $\omega$ between 0 and 1 (where the ratio $\sigma_i^2/\sigma_j^2$ 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 follows **?** and **?**. Using the formulation of an unrestricted vector autoregressive model in (18), we allow for the respective coefficient

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

matrices, $\varphi$, to be multiplied through by the indicator matrix, $\gamma_{i,j}$, which has elements that follow 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 subjected to the likelihood function to derive the final posterior values.[17]

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.[18] 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.[19] When specifying this model, we condition the coefficient matrix by assuming that the off-diagonal elements of the co-variance matrix are zero.

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 $\vartheta_{i,t} \sim N(0, \sigma^2_{\vartheta})$. 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, $\pi_t$, and a measure of 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 2005 constant prices was obtained from the South African Reserve Bank.[20] The data for interest rate relates to the 3-month Treasury bill, which was obtained from the International Financial Statistics (IFS) 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 and then seasonally adjusted using the X-12 procedure developed by the Department of Commerce, U.S. Census Bureau.

As discussed above, we select the time period 1960:Q1 through 2011:Q4 for our analysis. This gives a total sample of 208 observations on each series with the first 160 (1960:Q1 through 1999:Q4) used for in-sample analysis, following the existing literature on forecasting for South Africa. The remaining 48 (2000:Q1 through 2011:Q4) is used for the out-of-sample forecasting, over which the models are recursively estimated by increasing the size of the in-sample by one observation to produce one- to eight-steps-ahead forecasts. Note, the choice

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

[18]This 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.

[19]See, **?** for further details on the comparative performance of LASSO and Normal-Jeffreys priors.

[20]For 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.

of the starting point of the out-of-sample 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 $1, 2, \ldots, 8$ step step-ahead forecasts over the entire out-of-sample period.[21]

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 is then used to calculate the relative root-mean-squared error for the other models.[22] The **?** statistic is then used to determine whether this difference in forecasting performance is statistically significant.

## 5.1 Consumer price inflation

The results for the $\pi_t$ forecasts are contained in table (5.2). 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 there is a positive relative root-mean-squared error, where the largest value is 27.97.

The model that is 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.[23] 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.[24]

To ensure that these results are not attributable to any outlier that may be 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.[25,26] Over longer horizons the results of the nonlinear

[21] Further details of this forecasting exercise have been contained in the online appendix, which considers various other ways of summarizing 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 $\pi_t$ 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.

[22] For 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.

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

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

[25] Further details of this analysis are contained in the online appendix.

[26] When 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 linear and nonlinear DSGE models are almost equivalent.

DSGE are more impressive, which would confirm that it is responsible for significantly smaller forecasting errors at longer horizons.[27]

## 5.2 Output

The relative root-mean square errors for the forecasts for $y_t$ are contained in table (5.2). 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.[28]

Furthermore, there are a number of occasions where the Diebold-Mariano statistic suggests that the nonlinear model provides significantly better results when the forecasting horizon is at least a year.[29] 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.[30]

Once again, when comparing the nonlinear DSGE model with it's 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.[31] 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.

[27] The 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 comparison between 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.

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

[29] Indeed, 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

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

[31] It 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 | 0.24 | 1.32 | 2.36 | 3.98 | 5.73 | 6.9 | 8.77 |
| | [*0.29*] | [*-0.21*] | [*-1.09*] | [*-1.41*] | [*-1.83*] | [*-2.3*]⋆⋆ | [*-2.53*]⋆⋆ | [*-2.8*]⋆⋆ |
| RW | 10.18 | 18.27 | 20.61 | 21.33 | 27.2 | 27.97 | 24.7 | 17.67 |
| | [*-1.73*] | [*-2.46*]⋆⋆ | [*-2.44*]⋆⋆ | [*-2.52*]⋆⋆ | [*-3.49*]⋆⋆ | [*-2.94*]⋆⋆ | [*-2.28*]⋆⋆ | [*-1.76*] |
| VAR | 3.88 | 3.39 | 4.09 | 5.94 | 3.18 | 2.44 | 3.38 | 6.3 |
| | [*-1.48*] | [*-0.89*] | [*-0.73*] | [*-0.92*] | [*-0.47*] | [*-0.33*] | [*-0.41*] | [*-0.65*] |
| BVAR | 3.87 | 3.43 | 4 | 5.86 | 3.13 | 2.34 | 3.21 | 6.18 |
| | [*-1.49*] | [*-0.91*] | [*-0.72*] | [*-0.91*] | [*-0.47*] | [*-0.32*] | [*-0.39*] | [*-0.64*] |
| Minnes1 | 3.11 | 1.92 | 1.72 | 3.01 | 0.94 | -0.19 | -0.23 | 1.85 |
| | [*-1.52*] | [*-0.61*] | [*-0.4*] | [*-0.56*] | [*-0.16*] | [*0.03*] | [*0.03*] | [*-0.2*] |
| Minnes2 | 3.05 | 1.51 | 1.47 | 2.89 | 0.64 | -0.34 | -0.08 | 2.24 |
| | [*-1.4*] | [*-0.47*] | [*-0.33*] | [*-0.53*] | [*-0.11*] | [*0.05*] | [*0.01*] | [*-0.23*] |
| Minnes3 | 2.15 | 0.77 | 1.22 | 2.83 | 0.24 | -0.34 | 0.67 | 3.32 |
| | [*-0.91*] | [*-0.21*] | [*-0.25*] | [*-0.48*] | [*-0.04*] | [*0.05*] | [*-0.08*] | [*-0.32*] |
| Minnes4 | 1.84 | 0 | 2.02 | 4.17 | 0.26 | 0.17 | 2.28 | 5.49 |
| | [*-0.46*] | [*0*] | [*-0.31*] | [*-0.58*] | [*-0.04*] | [*-0.02*] | [*-0.25*] | [*-0.48*] |
| Minnes5 | 0.82 | 0.31 | 1.39 | 3.06 | 0.11 | -0.06 | 1.63 | 4.44 |
| | [*-0.32*] | [*-0.08*] | [*-0.25*] | [*-0.48*] | [*-0.02*] | [*0.01*] | [*-0.19*] | [*-0.41*] |
| Minnes6 | 1.01 | 0.45 | 3.15 | 5.2 | 0.63 | 0.58 | 2.91 | 6.1 |
| | [*-0.23*] | [*-0.07*] | [*-0.43*] | [*-0.66*] | [*-0.08*] | [*-0.07*] | [*-0.31*] | [*-0.52*] |
| Minnes7 | 3.56 | 2.74 | 3.84 | 5.91 | 2.98 | 2.59 | 4.04 | 7.17 |
| | [*-1.24*] | [*-0.66*] | [*-0.66*] | [*-0.9*] | [*-0.44*] | [*-0.35*] | [*-0.48*] | [*-0.72*] |
| Minnes8 | 2.88 | 2.15 | 4 | 6.31 | 2.99 | 3.02 | 5.15 | 8.5 |
| | [*-0.85*] | [*-0.44*] | [*-0.63*] | [*-0.91*] | [*-0.43*] | [*-0.4*] | [*-0.59*] | [*-0.82*] |
| Minnes9 | 3.83 | 2.46 | 5.79 | 8.29 | 3.56 | 3.82 | 6.62 | 10.19 |
| | [*-0.68*] | [*-0.36*] | [*-0.74*] | [*-1.05*] | [*-0.47*] | [*-0.47*] | [*-0.72*] | [*-0.92*] |
| Minnes10 | 1.85 | 2.07 | 4.63 | 7 | 3.25 | 3.58 | 6.19 | 9.58 |
| | [*-0.5*] | [*-0.37*] | [*-0.67*] | [*-0.96*] | [*-0.45*] | [*-0.46*] | [*-0.69*] | [*-0.89*] |
| Minnes11 | 3.33 | 3.25 | 6.93 | 9.2 | 3.86 | 4.07 | 6.95 | 10.48 |
| | [*-0.55*] | [*-0.42*] | [*-0.82*] | [*-1.1*] | [*-0.49*] | [*-0.5*] | [*-0.74*] | [*-0.93*] |
| Minnes12 | 3.86 | 3.39 | 4.09 | 5.94 | 3.18 | 2.44 | 3.39 | 6.31 |
| | [*-1.48*] | [*-0.89*] | [*-0.73*] | [*-0.92*] | [*-0.47*] | [*-0.33*] | [*-0.41*] | [*-0.65*] |
| Minnes13 | 3.62 | 1.88 | 3.26 | 5.19 | 2.06 | 1.69 | 2.96 | 5.6 |
| | [*-1.04*] | [*-0.42*] | [*-0.55*] | [*-0.81*] | [*-0.32*] | [*-0.24*] | [*-0.36*] | [*-0.56*] |
| SVS | 1.35 | 1.27 | 2.5 | 4.53 | 1.65 | 1.24 | 2.5 | 5.46 |
| | [*-0.53*] | [*-0.31*] | [*-0.45*] | [*-0.71*] | [*-0.25*] | [*-0.17*] | [*-0.29*] | [*-0.52*] |
| SVS-TVP | 1.69 | 1.87 | 3.19 | 4.36 | 0.72 | 0.81 | 2 | 4.51 |
| | [*-0.67*] | [*-0.46*] | [*-0.57*] | [*-0.69*] | [*-0.11*] | [*-0.11*] | [*-0.24*] | [*-0.43*] |
| SVS-SB | 1.95 | 2 | 2.57 | 4.11 | 1.5 | 0.55 | 1.91 | 5.05 |
| | [*-0.8*] | [*-0.51*] | [*-0.48*] | [*-0.68*] | [*-0.24*] | [*-0.08*] | [*-0.24*] | [*-0.52*] |
| SVS-LAS | 1.77 | 1.86 | 3.1 | 5.08 | 1.91 | 1.49 | 3.05 | 6.13 |
| | [*-0.68*] | [*-0.45*] | [*-0.55*] | [*-0.78*] | [*-0.29*] | [*-0.2*] | [*-0.36*] | [*-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.[32]

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

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.[33]

[33]Model acronyms: See Footnote 31.

| | 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*] |

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.[34]

[34] Model acronyms: See Footnote 31.

### 5.3 Interest rates

An evaluation of the selected models ability to forecast $i_t$ is contained in table (5.2). In this case, most of the relative-root-mean-squared errors would indicate that competing models would often generate a lower forecasting error. However, these improvement are often insignificant, with the exception of one of the BVAR models with Minnesota prior, which is able to generate significantly better forecasts when the horizon is at least a year.[35]

When comparing the results for the nonlinear DSGE with its 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.[36]

## 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 it's linear counterpart), these results suggest that when seeking to make use of a structural macroeconomic model to inform policy (be it through an optimal policy investigation, etc.) one should seek to incorporate potentially important nonlinearities in the model structure.

In addition, the results also suggest that nonlinear DSGE model is also able to improve upon the forecasting performance of most vector autoregressive and random walk models. However, there are a number of instances where certain BVAR models provide significant improvements when forecasting output and interest rates. This occurs only 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.

[35]One 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 $\phi_i$) is usually above 0.9. See, **?** for more on the values of the coefficients in the structural macroeconometric model that incorporates an estimated Taylor rule.

[36]On 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.