



UNIVERSITEIT VAN PRETORIA  
UNIVERSITY OF PRETORIA  
YUNIBESITHI YA PRETORIA

**FORECASTING WITH DSGE MODELS: THE CASE OF  
SOUTH AFRICA**

Guangling “Dave” Liu

M.Com., Stellenbosch University, 2004

A Thesis

Submitted in Fulfillment of the Requirements for the Degree of

Doctor of Philosophy (Economics)

at the

University of Pretoria

2008



UNIVERSITEIT VAN PRETORIA  
UNIVERSITY OF PRETORIA  
YUNIBESITHI YA PRETORIA

Copyright by

Guangling “Dave” Liu

2008



UNIVERSITEIT VAN PRETORIA  
UNIVERSITY OF PRETORIA  
YUNIBESITHI YA PRETORIA

Doctor of Philosophy Thesis

**FORECASTING WITH DSGE MODELS: THE CASE OF  
SOUTH AFRICA**

Presented by

Guangling “Dave” Liu, M.Com.

Major Advisor

Prof. Rangan Gupta

Associate Advisor

Prof. Eric Schaling

University of Pretoria

2008



UNIVERSITEIT VAN PRETORIA  
UNIVERSITY OF PRETORIA  
YUNIBESITHI YA PRETORIA

# FORECASTING WITH DSGE MODELS: THE CASE OF SOUTH AFRICA

Guangling “Dave” Liu, Ph.D.

University of Pretoria, 2008

## ABSTRACT

The objective of this thesis is to develop alternative forms of Dynamic Stochastic General Equilibrium (DSGE) models for forecasting the South African economy and, in turn, compare them with the forecasts generated by the Classical and Bayesian variants of the Vector Autoregression Models (VARs). Such a comparative analysis is aimed at developing a small-scale micro-founded framework that will help in forecasting the key macroeconomic variables of the economy.

The thesis consists of three independent papers. The first paper develops a small-scale DSGE model based on Hansen’s (1985) indivisible labor Real Business Cycle (RBC) model. The results suggest that, compared to the VARs and the Bayesian VARs, the DSGE model produces large out-of-sample forecast errors.

In the basic RBC framework, business cycle fluctuations are purely driven by real technology shocks. This one-shock assumption makes the RBC models stochastically singular. In order to overcome the singularity problem in the RBC model developed in the first paper, the second paper develops a hybrid model (DSGE-VAR), in which the theoretical model is augmented with unobservable

errors having a VAR representation. The model is estimated via maximum likelihood technique. The results suggest DSGE-VAR model outperforms the Classical VAR, but not the Bayesian VARs. However, it does indicate that the forecast accuracy can be improved alarmingly by using the estimated version of the DSGE model.

The third paper develops a micro-founded New-Keynesian DSGE (NKDSGE) model. The model consists of three equations, an expectational IS curve, a forward-looking version of the Phillips curve, and a Taylor-type monetary policy rule. The results indicate that, besides the usual usage for policy analysis, a small-scale NKDSGE model has a future for forecasting. The NKDSGE model outperforms both the Classical and Bayesian variants of the VARs in forecasting inflation, but not for output growth and the nominal short-term interest rate. However, the differences of the forecast errors are minor. The indicated success of the NKDSGE model for predicting inflation is important, especially in the context of South Africa — an economy targeting inflation.

## ACKNOWLEDGEMENTS

I wish to extend my deep gratitude to my parents, brother and sister, for their support and patience throughout many years since I first entered university. I owe my family a debt I may never be able to repay.

I would like to thank my thesis advisor, Prof. Rangan Gupta, for his guidance and support. Without his encouragement, and intellectual enthusiasm, this research would not have been possible. I am indebted to him for the time and energy he has spent on me. His guidance throughout my research and career so far is highly appreciated. Working with Prof. Gupta was a pleasant journey. I am looking forward to our continued collaboration. I would also like to thank my thesis co-advisor, Prof. Eric Schaling, for his critical insights and challenging ideas.

I would like to thank my colleagues at the Department of Economics for their encouragement and discussions. Special thanks goes out to Prof. Jan van Heerden and Prof. Steven Koch for their consistent support during the last two years. I would also like to borrow this opportunity to thank Prof. Stan du Plessis, my former lecturer and advisor at Stellenbosch University. His influence has been instrumental on my choice to pursue macroeconomics as a research area.

I am grateful to two of my old friends, Prof. Jianxin Chi and Mr. Daping Wang for their consistent encouragement and help throughout the last few years.

Chapter two of the thesis has been published in the June issue of 2007 of the South African Journal of Economics. In this regard, I would like to thank an anonymous referee for valuable comments. Chapter three of the thesis has been published as a working paper of Economic Research Southern Africa (ERSA). It was also presented at the 27th Annual International Symposium of Forecasting (ISF), in New York, U.S.A, and in the 12th Annual African Econometric Society (AES) conference in Cape Town. I thank members of the audiences at these presentations for many helpful comments and suggestions. I also presented the preliminary results of this paper in the ERSA macroeconomic workshop, held at South African Reserve Bank in May 2007. For this, I am grateful to Prof. Nicola Viegi of the University of Cape Town for his invitation and valuable comments. Finally, I would also like to thank the Department of Economics at University of Cape Town for inviting me to present the fourth chapter of my thesis in their seminar series in September 2007.

# TABLE OF CONTENTS

<b>Chapter 1:</b>	<b>Introduction</b>	<b>1</b>
<b>Chapter 2:</b>	<b>A Small-Scale DSGE Model for Forecasting the South African Economy</b>	<b>7</b>
2.1	Introduction . . . . .	7
2.2	The Model Economy . . . . .	11
2.3	Calibration . . . . .	13
2.4	Empirical Performance of the Model . . . . .	15
2.4.1	Data moments and cross-correlation . . . . .	15
2.4.2	Impulse response analysis . . . . .	16
2.4.3	Forecast accuracy . . . . .	19
2.4.3.1	Classical and Bayesian VARs . . . . .	20
2.4.3.2	DSGE vs. VARs . . . . .	23
2.5	Conclusion . . . . .	27
	Appendix . . . . .	28
<b>Chapter 3:</b>	<b>Forecasting the South African Economy: A DSGE-VAR Approach</b>	<b>32</b>
3.1	Introduction . . . . .	32
3.2	The Model Economy . . . . .	34
3.3	The Hybrid Model: A DSGE-VAR Approach . . . . .	36





3.4	Results . . . . .	41
3.4.1	Classical and Bayesian VARs . . . . .	42
3.4.2	Forecast accuracy . . . . .	45
3.5	Conclusion . . . . .	50
	Appendix . . . . .	51
<b>Chapter 4: A New-Keynesian DSGE Model for Forecasting the South African Economy</b>		<b>56</b>
4.1	Introduction . . . . .	56
4.2	The Model . . . . .	59
4.2.1	The Representative Household . . . . .	59
4.2.2	Final-Goods Production . . . . .	61
4.2.3	Intermediate-Goods Production . . . . .	63
4.2.4	The Monetary Authority . . . . .	64
4.3	Solution of the Model . . . . .	67
4.4	Results . . . . .	70
4.4.1	Classical and Bayesian VARs . . . . .	71
4.4.2	Forecast accuracy . . . . .	74
4.5	Conclusion . . . . .	77
	Appendix . . . . .	79
<b>Chapter 5: Conclusions</b>		<b>86</b>
<b>Bibliography</b>		<b>87</b>



## LIST OF TABLES

1	Parameters calibrated to the model economy . . . . .	15
2	Statistical moments: Baselinemodel and RSA data . . . . .	17
3	MAPE (2001:1-2005:4): Real GNP in logs . . . . .	25
4	MAPE (2001:1-2005:4): Final consumption expenditure by households in logs . . . . .	25
5	MAPE (2001:1-2005:4): Investment expenditure in logs .	26
6	MAPE (2001:1-2005:4): Employment in logs . . . . .	26
7	MAPE (2001:1-2005:4): Real treasury bill rate (91 days)	26
8	RMSE (2001Q1-2005Q4): Output . . . . .	46
9	RMSE (2001Q1-2005Q4): Consumption . . . . .	47
10	RMSE (2001Q1-2005Q4): Investment . . . . .	47
11	RMSE (2001Q1-2005Q4): Hours worked . . . . .	47
12	Across-Model Test Statistics . . . . .	49
13	RMSE (2001Q1-2006Q4): Output Growth . . . . .	75
14	RMSE (2001Q1-2006Q4): Inflation . . . . .	76
15	RMSE (2001Q1-2006Q4): TBILL . . . . .	76
16	Across-Model Test Statistics . . . . .	77



## LIST OF FIGURES

1	Impulse responses to technology shock(baseline model) . .	18
2	Impulse responses to technology shock(actual data) . . . .	19



# Chapter 1

## Introduction

Generally, economy-wide forecasting models, at business cycle frequencies, are in the form of simultaneous-equations structural models. However, two problems often encountered with such models are as follows: (i) the correct number of variables needs to be excluded, for proper identification of individual equations in the system which are, however, often based on little theoretical justification (Cooley and LeRoy, 1985); and (ii) given that projected future values are required for the exogenous variables in the system, structural models are poorly suited to forecasting.

The Vector Autoregression (VAR) model, though 'atheoretical' is particularly useful for forecasting purposes. Moreover, as shown by Zellner (1979) and Zellner and Palm (1974) any structural linear model can be expressed as a VAR moving average (VARMA) model, with the coefficients of the VARMA model being combinations of the structural coefficients. Under certain conditions, a VARMA

model can be expressed as a VAR and a VMA model. Thus, a VAR model can be visualized as an approximation of the reduced-form simultaneous equation structural model.

Though, both the large-scale econometric models and the VARs perform reasonably well as long as there are no structural changes whether in or out of the sample. Specifically, Lucas (1976) indicates that estimated functional forms obtained for macroeconomic models in the Keynesian tradition, as well as VARs, are not “deep” because these models do not correctly account for the dependence of private agents’ behavior on anticipated government policy rules, used for generating current and future values for government policy variables. Under such circumstances, while such models may be useful for forecasting future states of the economy conditional on a given government policy rule, they are fatally flawed when there are changes to government policy rules. Econometrically, this means that in a later time period,  $T + t$ , this problem would show up as an occurrence of a “structural break” in the estimate for the parameters of the model at  $T$ . In other words, if the sampling period were broken up into two subsamples, one spanning periods prior to  $T$ , and one spanning periods after  $T$ , it would be seen that the “best-fit” estimates for the parameters of the model, over these two subsamples, are statistically different from each other.

Furthermore, the standard econometric models, as well as the VARs, are linear and hence fail to take account of the nonlinearities in the economy. One and perhaps the best response to these objections has been the development of

micro-founded DSGE models that are capable of handling both the possibilities of structural changes and the issues of nonlinearities, since DSGE models are able to identify that the actions of rational agents are not only dependent on government policy variables, but also on government policy rules.

Since Kydland and Prescott (1982), a vast literature has evolved attempting to model the business cycle, as an equilibrium outcome of the representative agents' response to a productivity shock ( Hansen,1985; Hansen and Sargent, 1988; Christiano and Eichenbaum, 1992; King *et al*, 1988). Hansen and Prescott (1993) suggest the 1990-91 recession in the U.S. economy can be explained by a real business cycle model with technology shocks. However, the weakness of their analysis, with regard to forecasting, is that it cannot actually forecast the recession since the measurements of technology shocks are *ex post*. Ingram and White-man (1994) show that forecasting with BVAR models, in which priors are generated by real business cycle models, outperforms the one based on standard VAR models. Recently, based on the work done by Christiano, *et al.* (2003), Smets and Wouters (2003, 2004) develop micro-founded DSGE models with sticky prices and wages for the European economy. By employing the Bayesian techniques, the authors investigate the relative importance of the various frictions and shocks in explaining the European business cycle as well as its prediction performance. They find that the estimated DSGE model is able to outperform the unrestricted VAR and BVAR models in out-of-sample predictions. This result clearly suggests

that the micro-founded DSGE models can be used as forecasting tools by central banks.

The objectives of the thesis are twofold, with the primary objective being to develop alternative DSGE models for forecasting South African economy. It is worth noting that all the DSGE models used for forecasting discussed above suggest that productivity shock plays a leading role in all the models. This research starts off with a Real Business Cycle model but extends it to account for nominal shocks. This is extremely important in the case of the South African economy, given the structure and policy changes over time. Both calibrated and estimated versions of Real Business Cycle (RBC) and New Keynesian Macroeconomic (NKM) DSGE models have been employed to forecast the South African economy.

The second objective is to evaluate the forecasting performances of the alternative DSGE models by comparing them with both the Classical and Bayesian variants of VARs. This comparison study allows us to analyze the forecasting abilities of alternative models, and in turn help us to select a suitable model for predicting the economy.

The thesis consists of three independent papers. The first paper develops a small-scale DSGE model based on Hansen's (1985) indivisible labor RBC model. The calibrated model is used to forecast output and its main components, and a measure of the short-term interest rate (91 days Treasury Bill rate). The results suggest that, compared to the VARs and the BVARs, the DSGE model produces

large out-of-sample forecast errors. In the basic RBC framework, business cycle fluctuations are purely driven by real technology shocks (Kydland and Prescott, 1982). This one-shock assumption makes the RBC models stochastically singular. As indicated by Rotemberg and Woodford (1995), output is unforecastable with only one state variable.

In order to overcome the singularity problem in the RBC model developed in the first paper, the second paper develops a hybrid model (DSGE-VAR) model. In the hybrid model, the theoretical model is augmented with unobservable errors having a VAR representation. This allows one to combine the theoretical rigor of a micro-founded DSGE model with the flexibility of an atheoretical VAR model in the hybrid model. The model is estimated via maximum likelihood technique. The results suggest that the estimated hybrid DSGE (DSGE-VAR) model outperforms the Classical VAR, but not the Bayesian VARs. However, it does indicate that the forecast accuracy can be improved alarmingly by using the estimated version of the DSGE model.

The third paper develops a micro-founded New-Keynesian DSGE (NKDSGE) model. The model consists of three equations, an expectational IS curve, a forward-looking version of the Phillips curve, and a Taylor-type monetary policy rule. Furthermore, the model is characterized by four shocks: a preference shock; a technology shock; a cost-push shock; and a monetary policy shock. Essentially, by incorporating four shocks, that generally tends to affect a macroeconomy, the



paper attempts to model the empirical stochastics and dynamics in the data better, and hence, improve the predictions. The results indicate that, besides the usual usage for policy analysis, a small-scale NKDSGE model has a future for forecasting. The NKDSGE model outperforms both the Classical and Bayesian variants of the VARs in forecasting inflation, but not for output growth and the nominal short-term interest rate. However, the differences of the forecasts errors are minor. The indicated success of the NKDSGE model for predicting inflation is important, especially in the context of South Africa — an economy targeting inflation.

The main contribution of the thesis lies in its ability to show that econometrically estimated models which have strong theoretical foundations can be used for forecasting key macroeconomic variables. Moreover, a theoretically sound framework, well-suited for forecasting, has the simultaneous advantage of being used for policy analysis at business cycle frequencies. This thesis, using South Africa as a case study, hence, attempts to bridge the gap between Econometricians and the Business Cycle Theorists. The thesis shows that, when compared with the atheoretical econometric models, the theoretically well equipped models have worthwhile future in carrying out economy-wide predictions.



## Chapter 2

# A Small-Scale DSGE Model for Forecasting the South African Economy

### 2.1 Introduction

This paper develops a small-scale Real Business Cycle Dynamic Stochastic General Equilibrium (DSGE) model for the South African economy, and forecasts real Gross National Product (GNP), consumption, investment, employment, and a measure of short-term interest rate (91 days Treasury Bill rate), over the period of 1970Q1-2000Q4. The out-of-sample forecasts from the DSGE model is then compared with the forecasts based on an unrestricted Vector Autoregression (VAR) and Bayesian VAR (BVAR) models for the period 2001Q1-2005Q4.

Generally, economy-wide forecasting models, at business cycle frequencies, are in the form of simultaneous-equations structural models. However, two problems often encountered with such models are as follows: (i) the correct number of

variables needs to be excluded, for proper identification of individual equations in the system which are, however, often based on little theoretical justification (Cooley and LeRoy, 1985); and (ii) given that projected future values are required for the exogenous variables in the system, structural models are poorly suited to forecasting.

The Vector Autoregression (VAR) model, though 'atheoretical' is particularly useful for forecasting purposes. Moreover, as shown by Zellner (1979) and Zellner and Palm (1974) any structural linear model can be expressed as a VAR moving average (VARMA) model, with the coefficients of the VARMA model being combinations of the structural coefficients. Under certain conditions, a VARMA model can be expressed as a VAR and a VMA model. Thus, a VAR model can be visualized as an approximation of the reduced-form simultaneous equation structural model.

Though, both the large-scale econometric models and the VARs perform reasonably well as long as there are no structural changes whether in or out of the sample. Specifically, Lucas (1976) indicates that estimated functional forms obtained for macroeconomic models in the Keynesian tradition, as well as VARs, are not "deep" because these models do not correctly account for the dependence of private agents' behavior on anticipated government policy rules, used for generating current and future values for government policy variables. Under such circumstances, while such models may be useful for forecasting future states of the economy conditional on a given government policy rule, they are fatally

flawed when there are changes to government policy rules. Econometrically, this means that in a later time period,  $T + t$ , this problem would show up as an occurrence of a “structural break” in the estimate for the parameters of the model at  $T$ . In other words, if the sampling period were broken up into two subsamples, one spanning periods prior to  $T$ , and one spanning periods after  $T$ , it would be seen that the “best-fit” estimates for the parameters of the model, over these two subsamples, are statistically different from each other.<sup>1</sup>

Furthermore, the standard econometric models, as well as the VARs, are linear and hence fail to take account of the nonlinearities in the economy. One and perhaps the best response to these objections has been the development of micro-founded DSGE models that are capable of handling both the possibilities of structural changes and the issues of nonlinearities, since DSGE models are able to identify that the actions of rational agents are not only dependent on government policy variables, but also on government policy rules.

Since Kydland and Prescott (1982), a vast literature has evolved attempting to model the business cycle, as an equilibrium outcome of the representative agents’ response to a productivity shock ( Hansen, 1985; Hansen and Sargent, 1988; Christiano and Eichenbaum, 1992; King *et al*, 1988)<sup>2</sup> . Hansen and Prescott (1993)

---

<sup>1</sup>Even though we do not explicitly incorporate the role of government policy in the DSGE model, but given that the model is micro-founded, the set-up would have been immune to the “Lucas Critique”, if a government policy was in fact present. See section 2 for further details.

<sup>2</sup>For an exceptional source of research along this line, see *Journal of Monetary Economics*, 1988, vol. 21 (March/May).

suggest that the 1990-91 recession in the U.S. economy can be explained by a real business cycle model with technology shocks. However, the weakness of their analysis, with regard to forecasting, is that it cannot actually forecast the recession since the measurements of technology shocks are *ex post*. Ingram and Whiteman (1994) show that forecasting with BVAR models, in which priors are generated by real business cycle models, outperforms the one based on standard VAR models. Recently, based on the work done by Christiano, *et al.* (2003), Smets and Wouters (2003, 2004) develop micro-founded DSGE models with sticky prices and wages for the European economy. By employing the Bayesian techniques, the authors investigate the relative importance of the various frictions and shocks in explaining the European business cycle as well as its prediction performance. They find that the estimated DSGE model is able to outperform the unrestricted VAR and BVAR models in out-of-sample predictions. This result clearly suggests that the micro-founded DSGE models can be used as forecasting tools by central banks.

Besides the introduction and conclusion, the paper is organized as follows: section 2 lays out the theoretical model, while section 3 presents the calibration of the model economy; section 4 discusses the performance of the DSGE model in terms of explaining the business cycle properties of South African economy and evaluating the accuracy of forecasts relative to the VARs.

## 2.2 The Model Economy

The model economy, here, is based on the benchmark real business cycle model developed by Hansen (1985). Equilibrium models have been criticized for depending heavily on individuals' substitution of leisure and work responding to the change in interest rate or wage. Hansen (1985) argues that in the real economy labor is indivisible. Individuals either work full time or not at all. Other features of Hansen's indivisible labor are exactly the same as standard real business model, such as Kydland and Prescott (1982). The economic environment is described below.

The model economy is populated by infinitely-lived households. The preferences of households are assumed to be identical. Households maximize the expected utility over life time:

$$U(C_t, N_t) = E_t \sum_{t=0}^{\infty} \beta^t \left( \frac{C_t^{1-\eta} - 1}{1-\eta} - AN_t \right), \quad 0 < \beta < 1 \quad \eta > 0 \quad (1)$$

where  $C_t$  and  $N_t$  are consumption and labor respectively,  $\beta$  is the discount factor that households apply to future consumption, and  $\eta$  is the coefficient of relative risk aversion.

The technology is defined as a standard Cobb-Douglas production function:

$$Y_t = Z_t K_{t-1}^\rho N_t^{1-\rho} \quad (2)$$

where  $\rho$  is the fraction of aggregate output that goes to the capital input and  $1 - \rho$  is the fraction that goes to the labor input.  $Z_t$  is total factor productivity (TFP) which is exogenously evolving according to the law of motion:

$$\log Z_t = (1 - \psi)\log \bar{Z} + \psi \log Z_{t-1} + \epsilon_t, \quad \epsilon_t \sim i.i.d.(0, \sigma^2) \quad (3)$$

where  $\psi$  and  $\bar{Z}$  are parameters, and  $0 < \psi < 1$ .

As in a neoclassical growth model, capital stock depreciates at the rate  $\delta$ , and households invest a fraction of income in capital stock in each period. This amount of investment forms part of productive capital in current period. Therefore the law of motion for aggregate capital stock is

$$K_t = (1 - \delta)K_{t-1} + I_t, \quad 0 < \delta < 1 \quad (4)$$

Although in this indivisible model households do not choose hours worked in competitive equilibrium, the objective of the benevolent social planner is also to maximize the utility of the households (1), subject to the aggregate resource constraints

$$Y_t = C_t + I_t \quad (5)$$

$$Y_t = Z_t K_{t-1}^\rho N_t^{1-\rho} \quad (6)$$

$$K_t = (1 - \delta)K_{t-1} + I_t \quad (7)$$

$$\log Z_t = (1 - \psi)\log \bar{Z} + \psi \log Z_{t-1} + \epsilon_t, \quad \epsilon_t \sim i.i.d.(0, \sigma^2) \quad (8)$$

Uhlig (1995) illustrates the numerical solution methods for solving nonlinear stochastic dynamic models. The following section describes how to calibrate the model economy. Once all the parameters have been assigned, we can then log-linearize the DSGE model <sup>3</sup> and numerically solve the dynamic problems by employing the method of undetermined coefficients.

### 2.3 Calibration

This section follows the three-step process as outlined in Cooley and Prescott (1995). This involves moving from the general framework described in the previous section to quantitative measurements of the variables of interest — output, employment, investment, and so on. The first step is restricting the model to display balanced growth, that is, in steady state capital, consumption and investment all grow at a constant rate. The second step is defining the consistent measurements of the conceptual framework of the model economy and the real data. The parameter values of the model economy are then assigned according to the measured data during the sample period of 1970 to 2000.

The annual aggregate capital depreciation rate  $\delta$  is obtained from annual averaged values of  $\frac{I}{Y}$  and  $\frac{K}{Y}$ . This yields an annual depreciation rate of 0.076, or a quarterly rate of 0.019.

The standard real business cycle literature suggests that capital and labor shares of output have been approximately constant. The capital output share ( $\rho$ )

---

<sup>3</sup>The log-linearized equilibrium conditions are presented in Appendix A.



is equal to 0.26<sup>4</sup> , obtained from the steady state equation, whereas the labor output share  $(1 - \rho)$  is 0.74.

The measurement of technology shock, also known as Solow residual in growth accounting literature (Solow, 1957), is computed as follows:

$$\log Z_t - \log Z_{t-1} = (\log Y_t - \log Y_{t-1}) - (1 - \rho)(\log N_t - \log N_{t-1}) \quad (9)$$

Omitting the capital part of the expression <sup>5</sup> is not a serious problem given the fact that capital stock has very little contribution to the cyclical fluctuations of output (Kydland and Prescott, 1982; Backus, *at al*, 1995).

The parameter  $\bar{Z}$ , in the law of motion for TFP (3), is set equal to one. Therefore (3) becomes a first-order linear Markov process:

$$\log Z_t = \psi \log Z_{t-1} + \epsilon_t, \quad \epsilon_t \sim i.i.d.(0, \sigma^2) \quad (10)$$

The persistence parameter  $\psi$  is set equal to 0.95, which is consistent with the literature (Hansen, 1985). From (4) we can compute a set of innovations of technology  $\epsilon_t$ . These innovations have a standard deviation of 0.0083.

The discount factor  $\beta$  is set equal to 0.99, as in Hansen (1985), which implies an annual real interest rate of four percent in steady state. The coefficient of relative risk aversion  $\eta$ , is set equal to one. The parameter  $A$ , in the utility

---

<sup>4</sup>The capital output share for the South African economy is 0.39 in Zimmermann (2001), and 0.31 in Smit and Burrows (2002).

<sup>5</sup>There is no quarterly capital stock data available.

Table 1: **Parameters calibrated to the model economy**

$\rho$	$1 - \rho$	$A$	$\bar{Z}$	$\delta$	$\sigma_\epsilon$	$\beta$	$\psi$	$\eta$
0.26	0.74	2.6712	1.00	0.019	0.0083	0.99	0.95	1.00

function (1), is equal to 2.6712, obtained from (A.7). As shown in Table 1, all parameters of the model have now been assigned.

## 2.4 Empirical Performance of the Model

### 2.4.1 Data moments and cross-correlation

In this section, we compare the stylized facts of the actual data to those of obtained from the baseline model. Table 2 reports a number of statistics for both the baseline model and the actual data. All data are obtained from South African Reserve Bank Quarterly Bulletin except employment and population (aged 15 – 64) from the World Bank database.

The standard deviation of GNP is 2.18% in the baseline model, but 0.93% in the actual data. In other words, the baseline model exaggerates the variability of output. So does the investment (10.11% vs. 4.49%). Moreover, the baseline model underestimates the variability of the short term real interest rate<sup>6</sup> (0.06% vs. 2.77%). But, in general, the baseline model mimics most of the stylized facts of the business cycle. Employment is more or less as volatile as output

---

<sup>6</sup>The short term real interest rate,  $R$ , in actual data is 91 days Treasury Bill rate minus GNP deflator, a risk-free bank rate, which is comparable with the interest rate in the baseline model.

(0.76% vs. 0.74%), while investment is much more volatile than output (4.46% vs. 4.85%). Consumption is less volatile than output (0.29% vs. 0.86%). In order to be consistent with the model, in which the durability is disregarded, we use the measurement of non-durable goods consumption here. The measurement of consumption, elsewhere in this paper, is total consumption. Total consumption is more variable relative to output (1.07%) in actual data <sup>7</sup>. This scenario differs from the empirical regularity. For instance, Backus *et al.* (1995) show that output is more than 2-3 times variable relative to consumption in the economies of Canada, Japan, United Kingdom, and United States. It indicates that South African economy has a more volatile total consumption than other economies in general.

In the baseline model, consumption, investment, and employment are highly pro-cyclical, compared to those in actual data. Interest rate also has a high correlation with output, 95%, whereas there is little correlation between short term real interest rate and output in the actual data.

#### 2.4.2 Impulse response analysis

This section analyzes the responses of aggregate variables with respect to the productivity shock. As shown in Figure 1, the aggregates follow a hump-shaped pattern in response to the shock. In other words, the productivity shock

---

<sup>7</sup>The standard deviation of total consumption is 0.99%, slightly greater than that of output, 0.93%. It results the ratio of standard deviation to that of output, 1.07%.

Table 2: **Statistical moments: Baselinemodel and RSA data**

Variable	Baseline Model			RSA data		
	SD(%)	SD ratio to GNP	Corr.	SD(%)	SD ratio to GNP	Corr.
GNP	2.18	1.00	1.00	0.93	1.00	1.00
CON	0.63	0.29	0.87	0.80	0.86	0.44
INV	10.11	4.64	0.99	4.49	4.85	0.65
EMP	1.66	0.76	0.98	0.69	0.74	0.45
R	0.06	0.03	0.95	2.77	2.97	0.11

Notes: Statistics are based on Hodrick-Prescott-filtered data.

has a transitory output effect, which dies out over time. The response of short term interest rate is minimal, while investment responds the most among the five aggregates. In fact, investment increases more than 10% in the period that the positive shock occurs.

The scenarios in the actual data are more complicated. Figure 2 shows there is no significant hump-shaped pattern associated with the shock. The short term interest rate also responds little to the shock. The peak effect occurs with a longer lag than that in the baseline model. For instance, the peak effect occurs in the second period after the shock on consumption, third period on investment, and fourth period on labor time <sup>8</sup>. However, in the baseline model, the peak effect on all aggregates happens in the same period when the shock occurs. Investment does not exhibit the most response to the shock. Instead, the shock has a negative effect in the first period after the shock and a positive effect in the second period, then negative effect again from the third period onwards. The most serious

<sup>8</sup>In order to compare with the baseline model, we generate labor time by dividing employment with population aged 15-64 ( $N/L$  in Figure 2).

Figure 1: Impulse responses to technology shock(baseline model)

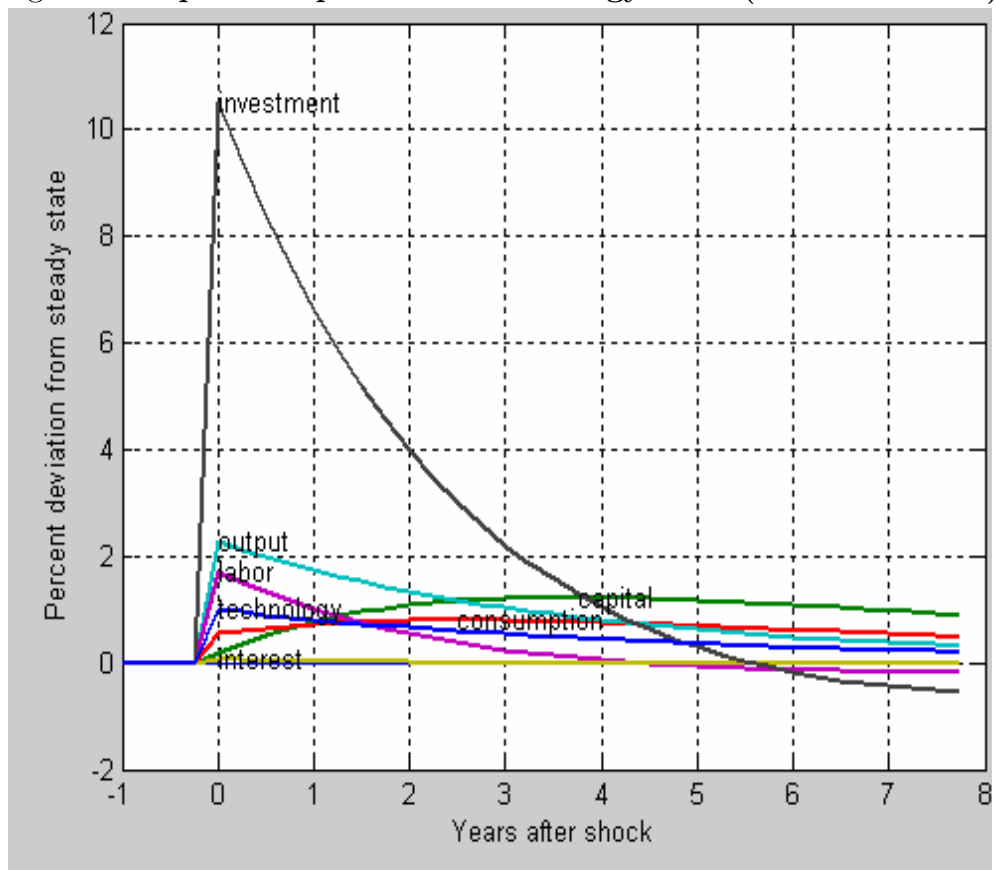
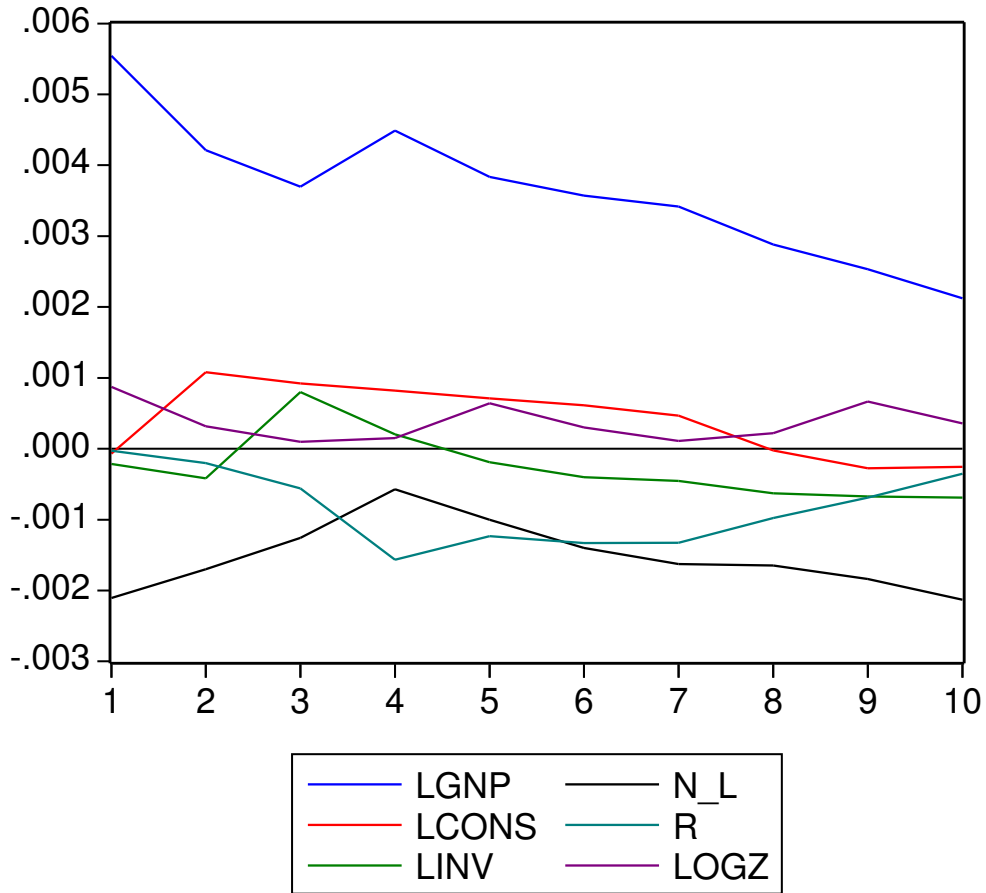


Figure 2: Impulse responses to technology shock(actual data)



problem is labor time, which exhibits a negative response to the shock. So is the short term real interest rate.

### 2.4.3 Forecast accuracy

In this section, we compare the out-of-sample forecasting perform of the DSGE model with the VARs in terms of the Mean Absolute Percentage Errors (MAPEs)<sup>9</sup>

. Before this, however, it is important to lay out the basic structural difference and, hence, the advantages of using BVARs over traditional VARs for forecasting.

### 2.4.3.1 Classical and Bayesian VARs

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

$$\chi_t = C + \lambda(L)\chi_t + \varepsilon_t \quad (11)$$

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

A crucial drawback of the VAR forecasts is “overfitting” due to the inclusion too many lags and too many variables, some of which may be insignificant. The problem of “overfitting” results in multicollinearity and loss of degrees of freedom, leads to inefficient estimates and large out-of-sample forecasting errors. Thus, it

---

<sup>9</sup>Whitley (1994: 187) argues that although the forecast accuracy can be evaluated by the comparison of MAPEs from different forecast models, there is no absolute measure of forecast performance against which to judge them.

$MAPE = \left[ \frac{1}{n} \sum_{t=1}^n \frac{|F_t - \hat{F}_t|}{F_t} \right] \times 100$ , where  $n$  is the number of observations,  $F_t$  is the actual value of the specific variable for period  $t$  and  $\hat{F}_t$  is the forecast value for period  $t$ . The summation is calculated as the following: for one period ahead forecast MAPE, the summation runs from 2001Q1 to 2005Q4; for two period ahead forecast MAPE, it runs from 2001Q2 to 2005Q4; and so on.

can be argued the performance of VAR forecasts will deteriorate rapidly as the forecasting horizon becomes longer.

A forecaster can overcome this “overfitting” problem by using Bayesian techniques. The motivation for the Bayesian analysis is based on the knowledge that more recent values of a variable are more likely to contain useful information about its future movements than older values. From a Bayesian perspective, the exclusion restriction in the VAR is, on the other hand, an inclusion of a coefficient without a prior probability distribution (Litterman, 1986a).

The Bayesian model proposed by Litterman (1981), Doan, *et al.* (1984), and Litterman (1986b), imposes restrictions on those coefficients by assuming they are more likely to be near zero. The restrictions are imposed by specifying normal prior<sup>10</sup> distributions with zero means and small standard deviations for all the coefficients with standard deviation decreasing as lag increases. One exception is that the mean of the first own lag of a variable is set equal to unity to reflect the assumption that own lags account for most of the variation of the given variable. To illustrate the Bayesian technique, suppose the “Minnesota prior” means and variances take the following form:

$$\begin{aligned}\beta_i &\sim N(1, \sigma_{\beta_i}^2) \\ \beta_j &\sim N(0, \sigma_{\beta_j}^2)\end{aligned}\tag{12}$$

---

<sup>10</sup>Note Litterman (1981) uses a diffuse prior for the constant, which is popularly referred to as the “Minnesota prior” due to its development at the University of Minnesota and the Federal Reserve bank at Minneapolis.



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

Doan *et al.* (1984) propose a formula to generate standard deviations as a function of small number of *hyperparameters*<sup>11</sup> :  $w$ ,  $d$ , and a weighting matrix  $f(i, j)$ . This approach allows the forecaster to specify individual prior variances for a large number of coefficients based on only a few *hyperparameters*. The specification of standard deviation of the distribution of the prior imposed on variable  $j$  in equation  $i$  at lag  $m$ , for all  $i, j$  and  $m$ , defined as  $S(i, j, m)$ :

$$S(i, j, m) = [w \times g(m) \times f(i, j)] \frac{\hat{\sigma}_i}{\hat{\sigma}_j} \quad (13)$$

where:

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

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

The term  $w$  is the measurement of standard deviation on the first own lag, which indicates the overall tightness. A decrease in the value of  $w$  results a tighter prior. The parameter  $g(m)$  measures the tightness on lag  $m$  relative to lag 1,

---

<sup>11</sup>The name of *hyperparameter* is to distinguish it from the estimated coefficients, the parameters of the model itself.

and is assumed to have a harmonic shape with a decay of  $d$ . An increasing in  $d$ , tightens the prior as lag increases. The parameter  $f(i, j)$  represents the tightness of variable  $j$  in equation  $i$  relative to variable  $i$ . Reducing the interaction parameter  $k_{ij}$  tightens the prior.  $\hat{\sigma}_i$  and  $\hat{\sigma}_j$  are the estimated standard errors of the univariate autoregression for variable  $i$  and  $j$  respectively. In the case of  $i \neq j$ , the standard deviations of the coefficients on lags are not scale invariant (Litterman, 1986b: 30). The ratio,  $\frac{\hat{\sigma}_i}{\hat{\sigma}_j}$  in (13), scales the variables so as to account for differences in the units of magnitudes of the variables.

The BVAR model is estimated using Theil's (1971) mixed estimation technique, which involves supplementing the data with prior information on the distribution of the coefficients. For each restriction imposed on the parameter estimated, the number of observations and degrees of freedom are increased by one in an artificial way. Therefore, the loss of degrees of freedom associated with the unrestricted VAR is not a concern in the BVAR.

#### 2.4.3.2 DSGE vs. VARs

The BVAR model is estimated in levels<sup>12</sup> with four lags for the period of 1970Q1 to 2000Q4. Consumption, investment and GNP are seasonally adjusted in order to address the fact that as pointed out by Hamilton (1994: 362), the Minnesota prior is not well suited for seasonal data. All variables except for the

---

<sup>12</sup>Sims *et al.* (1990) indicate that with the Bayesian approach entirely based on the likelihood function, the associated inference does not need to take special account of non-stationarity, since the likelihood function has the same Gaussian shape regardless of the presence of non-stationarity.

interest rate are measured in logarithms. We then perform the one- to eight-period-ahead forecasts for the period of 2001Q1 to 2005Q4. Following Dua *et al.* (1999), the overall tightness parameter ( $w$ ) is set equal to 0.1 and 0.2, 1 and 2 for the harmonic lag decay parameter ( $d$ ). Moreover, as in Dua and Ray (1995), we also report the results for a combination of  $w = 0.3$  and  $d = 0.5$ .

Table 3 to 7 summarizes the MAPEs of DSGE model and the VARs. In general, for all the five variables the DSGE model performs the worst. This is not a surprising result since the DSGE model is based on only two state variables, the previous capital stock and the productivity shock. The model is, thus, not rich enough to capture most of the movements of the real data. In addition, theoretically speaking, the methodology applied in this paper, involving calibration and forecasting based on simulated data, is not a preferable option in terms of forecasting. Ideally, these models need to be estimated using the real data.

Regarding forecasting performances of the VARs, the BVARs outperform the unrestricted VAR for predicting output, employment, and the short term real interest rate. In the cases of consumption and investment, the unrestricted VAR does a better job than the BVARs. As far as the BVAR itself is concerned, it is unclear whether a BVAR with a relatively loose or tight prior produces lower out-of-sample forecast errors. Our results indicate that for consumption and investment, a BVAR with the most loose prior ( $w = 0.3$ ,  $d = 0.5$ ) performs the best, whereas for employment and the short term real interest rate, a BVAR with the most tight prior ( $w = 0.1$ ,  $d = 2$ ) produces the best predictions. Whereas

Table 3: MAPE (2001:1-2005:4): Real GNP in logs

QA	VAR	DSGE	BVARs				
			(w=0.3,d=0.5)	(w=0.2,d=1)	(w=0.2,d=2)	(w=0.1,d=1)	(w=0.1,d=2)
1	0.0003	7.2790	0.0003	0.0003	0.0003	0.0001	0.0005
2	0.0035	7.2678	0.0035	0.0036	0.0045	0.0039	0.0050
3	0.0057	6.8978	0.0058	0.0060	0.0074	0.0065	0.0077
4	0.0009	7.2569	0.0009	0.0012	0.0029	0.0018	0.0033
5	0.0041	7.2468	0.0040	0.0037	0.0016	0.0029	0.0011
6	0.0023	7.2267	0.0023	0.0019	0.0003	0.0011	0.0009
7	0.0079	7.2110	0.0078	0.0074	0.0051	0.0066	0.0046
8	0.0067	7.1908	0.0066	0.0062	0.0038	0.0054	0.0032
AVE	0.0039	7.1971	0.0039	0.0038	<b>0.0032</b>	0.0035	0.0033

MAPE: mean absolute percentage error; QA: quarter ahead.

Table 4: MAPE (2001:1-2005:4): Final consumption expenditure by households in logs

QA	VAR	DSGE	BVARs				
			(w=0.3,d=0.5)	(w=0.2,d=1)	(w=0.2,d=2)	(w=0.1,d=1)	(w=0.1,d=2)
1	0.0030	5.0167	0.0030	0.0029	0.0029	0.0028	0.0026
2	0.0047	5.1038	0.0047	0.0048	0.0053	0.0050	0.0051
3	0.0064	5.1879	0.0064	0.0066	0.0076	0.0070	0.0074
4	0.0074	5.2687	0.0074	0.0077	0.0090	0.0081	0.0089
5	0.0096	5.3435	0.0097	0.0100	0.0116	0.0105	0.0116
6	0.0117	5.4093	0.0117	0.0121	0.0139	0.0127	0.0138
7	0.0141	5.4715	0.0141	0.0145	0.0165	0.0152	0.0163
8	0.0170	5.5248	0.0171	0.0175	0.0197	0.0183	0.0195
AVE	<b>0.0092</b>	5.2908	0.0093	0.0095	0.0108	0.0099	0.0107

MAPE: mean absolute percentage error; QA: quarter ahead.

for output, a BVAR with an average prior ( $w = 0.2$ ,  $d = 2$ ) generates the best forecasts.

Table 5: MAPE (2001:1-2005:4): Investment expenditure in logs

QA	VAR	DSGE	BVARs				
			(w=0.3,d=0.5)	(w=0.2,d=1)	(w=0.2,d=2)	(w=0.1,d=1)	(w=0.1,d=2)
1	0.0338	31.6635	0.0338	0.0338	0.0355	0.0339	0.0353
2	0.0438	31.0647	0.0439	0.0447	0.0498	0.0466	0.0519
3	0.0376	30.5257	0.0378	0.0396	0.0489	0.0432	0.0513
4	0.0385	30.0584	0.0388	0.0410	0.0536	0.0459	0.0582
5	0.0377	29.5514	0.0380	0.0406	0.0540	0.0457	0.0576
6	0.0652	28.9947	0.0656	0.0683	0.0828	0.0739	0.0871
7	0.0442	28.4545	0.0446	0.0479	0.0639	0.0541	0.0678
8	0.0377	27.9088	0.0382	0.0415	0.0587	0.0482	0.0629
AVE	<b>0.0423</b>	29.7777	0.0426	0.0447	0.0559	0.0489	0.0590

MAPE: mean absolute percentage error; QA: quarter ahead.

Table 6: MAPE (2001:1-2005:4): Employment in logs

QA	VAR	DSGE	BVARs				
			(w=0.3,d=0.5)	(w=0.2,d=1)	(w=0.2,d=2)	(w=0.1,d=1)	(w=0.1,d=2)
1	0.0130	38.9136	0.0131	0.0133	0.0149	0.0140	0.0162
2	0.0081	37.8742	0.0082	0.0086	0.0120	0.0096	0.0147
3	0.0084	36.9046	0.0082	0.0069	0.0022	0.0039	0.0088
4	0.0216	35.9728	0.0215	0.0205	0.0107	0.0176	0.0030
5	0.0270	34.9687	0.0268	0.0252	0.0130	0.0212	0.0040
6	0.0521	33.8304	0.0519	0.0502	0.0371	0.0461	0.0273
7	0.0944	32.7531	0.0941	0.0918	0.0760	0.0867	0.0646
8	0.1301	31.6181	0.1297	0.1274	0.1099	0.1219	0.0971
AVE	0.0443	35.3545	0.0442	0.0430	0.0345	0.0401	<b>0.0295</b>

MAPE: mean absolute percentage error; QA: quarter ahead.

Table 7: MAPE (2001:1-2005:4): Real treasury bill rate (91 days)

QA	VAR	DSGE	BVARs				
			(w=0.3,d=0.5)	(w=0.2,d=1)	(w=0.2,d=2)	(w=0.1,d=1)	(w=0.1,d=2)
1	0.1187	44.3911	0.1183	0.1200	0.1434	0.1218	0.1492
2	0.2318	46.1530	0.2326	0.2401	0.2621	0.2430	0.2398
3	0.0962	46.9272	0.0975	0.1012	0.1158	0.1187	0.1537
4	0.6453	46.0619	0.6486	0.6623	0.7351	0.7070	0.8118
5	0.4334	47.3252	0.4381	0.4609	0.5757	0.5215	0.6753
6	0.2345	47.3953	0.2291	0.2002	0.0395	0.1235	0.0803
7	0.4965	48.6340	0.4906	0.4569	0.2691	0.3729	0.1379
8	0.8325	50.4094	0.8265	0.7906	0.5865	0.7029	0.4500
AVE	0.3861	47.1621	0.3852	0.3790	0.3409	0.3639	<b>0.3372</b>

MAPE: mean absolute percentage error; QA: quarter ahead.

## 2.5 Conclusion

This paper is the first attempt in using a DSGE model for forecasting the South African economy. However, compared to the VARs and the BVARs, the DSGE model produces large out-of-sample forecast errors.

But one must realize that there are some inherent problems with the BVAR models, which the forecaster should keep in mind: firstly, the forecast accuracy depends critically on the specification of the prior, and secondly, the selection of the prior based on some objective function for the out-of-sample forecasts may not be “optimal” for the time period beyond the period chosen to produce the out-of-sample forecasts. Moreover, the choice of the variables, to be forecasted, using the BVAR models can also affect the tightness, and hence, the optimal prior. In a recent study, Gupta and Sichei (2006) while trying to forecast consumption, investment, GDP, CPI and short- and long-term interest rates for the South African economy, over the same period as in this study, finds the most tightest prior to be optimal.

As indicated by Rotemberg and Woodford (1995), output is unforecastable with only one state variable. The small-scale DSGE model, developed in this paper, should, thus, be extended to a more elaborate model that includes a wider set of state variables. In addition, others have found the estimated DSGE models to empirically outperform other econometric models in terms of forecasting, *inter alia*, Christiano, *et al.* (2005), Smets and Wouters (2004) hence, an estimated

version of the current DSGE model should be developed for forecasting the South African economy.

### A. The Log-linearized DSGE Model

This section presents the log-linearized DSGE model. The principle of log-linearization is to replace all equations by Taylor approximation around the steady state, which are linear functions in the log-deviations of the variables (Uhlig, 1995:4). Suppose  $X_t$  be the vector of variables,  $\bar{X}$  their steady state, and  $x_t$  the vector of log-deviations:

$$x_t = \log X_t - \log \bar{X} \tag{A.1}$$

in other words,  $x_t$  denote the percentage deviations from their steady state levels.

(A.1) can be written alternatively:

$$X_t = \bar{X} e^{x_t} \approx \bar{X}(1 + x_t) \tag{A.2}$$

In order to derive the log-linearized DSGE model, we need to use (A.2) to rewrite all the equations of the model and then take logarithms<sup>13</sup> .

The complete model economy:

---

<sup>13</sup>For details of log-linearization, see Uhlig (1995).

$$Y_t = C_t + I_t \quad (\text{A.3})$$

$$Y_t = Z_t K_{t-1}^\rho N_t^{1-\rho} \quad (\text{A.4})$$

$$K_t = (1 - \delta)K_{t-1} + I_t \quad (\text{A.5})$$

$$1 = \beta E_t \left[ \left( \frac{C_t}{C_{t+1}} \right)^\eta R_{t+1} \right] \quad (\text{A.6})$$

$$A = C_t^{-\eta} (1 - \rho) \frac{Y_t}{N_t} \quad (\text{A.7})$$

$$R_t = \rho \frac{Y_t}{K_{t-1}} + (1 - \delta) \quad (\text{A.8})$$

$$\log Z_t = (1 - \psi) \log \bar{Z} + \psi \log Z_{t-1} + \epsilon_t, \quad \epsilon_t \sim i.i.d.(0, \sigma^2) \quad (\text{A.9})$$

In steady state, we have:

$$\bar{Y} = \bar{C} + \bar{I} \quad (\text{A.10})$$

$$\bar{Y} = \bar{Z} \bar{K}^\rho \bar{N}^{1-\rho} \quad (\text{A.11})$$

$$\bar{K} = \left( \frac{\rho \bar{Z}}{\bar{R} - 1 + \delta} \right)^{\frac{1}{1-\rho}} \bar{N} \quad (\text{A.12})$$

$$\bar{I} = \delta \bar{K} \quad (\text{A.13})$$

$$A = \frac{1}{\bar{N}} (1 - \rho) \frac{\bar{Y}}{\bar{C}^\eta} \quad (\text{A.14})$$

$$\bar{R} = \frac{1}{\beta} \quad (\text{A.15})$$

The log-linearized equations:



$$\bar{Y}y_t = \bar{C}c_t + \bar{I}i_t \quad (\text{A.16})$$

$$y_t = z_t + \rho k_{t-1} + (1 - \rho)n_t \quad (\text{A.17})$$

$$\bar{K}k_t = \bar{I}i_t + (1 - \delta)\bar{K}k_{t-1} \quad (\text{A.18})$$

$$0 = E_t[\eta(c_t - c_{t+1}) + r_{t+1}] \quad (\text{A.19})$$

$$0 = -\eta c_t + y_t - n_t \quad (\text{A.20})$$

$$\bar{R}r_t = \rho \frac{\bar{Y}}{\bar{K}}(y_t - k_{t-1}) \quad (\text{A.21})$$

$$z_t = \psi z_{t-1} + \epsilon_t, \quad \epsilon_t \sim i.i.d.(0, \sigma^2) \quad (\text{A.22})$$

## B. The Recursive Law of Motion

The principle of undetermined coefficients method is to write all variables as linear functions of a vector of endogenous variables  $x_{t-1}$  and exogenous variables  $z_t$ . These variables are also called predetermined variables in the sense that they cannot be changed at date  $t$  (Uhlig, 1995). In our simple real business cycle model, the endogenous variable is capital,  $k_{t-1}$ , and exogenous variable is the productivity shock,  $z_t$ . We further define a list of other endogenous variables  $y_t$ , which includes output  $Y$ , consumption  $C$ , investment  $I$ , employment  $N$ , and the short term interest rate  $R$ . The equilibrium relationships between vectors  $x_{t-1}$ ,  $y_t$ , and  $z_t$  are:

$$0 = Ax_t + Bx_{t-1} + Cy_t + Dz_t \quad (\text{B.1})$$

$$0 = E_t[Fx_{t+1} + Gx_t + Hx_{t-1} + Jy_{t+1} + Ky_t + Lz_{t+1} + Mz_t] \quad (\text{B.2})$$

$$z_t = Nz_{t-1} + \epsilon_t, \quad \epsilon_t \sim i.i.d.(0, \sigma^2) \quad (\text{B.3})$$

The recursive law of motion is derived using Uhlig's MATLAB program:<sup>14</sup>

$$y_t = Px_{t-1} + Qz_t \quad (\text{B.4})$$

where  $y_t$  here is a vector of all endogenous variables in log-deviations:

$$\begin{pmatrix} k_t \\ y_t \\ c_t \\ i_t \\ n_t \\ r_t \end{pmatrix} = \begin{pmatrix} 0.9256 & 0.1993 \\ -0.1602 & 2.2625 \\ 0.4142 & 0.5493 \\ -2.9183 & 10.4897 \\ -0.5743 & 1.7132 \\ -0.033 & 0.0650 \end{pmatrix} \times \begin{pmatrix} k_{t-1} \\ z_t \end{pmatrix}$$

---

<sup>14</sup>See Uhlig (1995) for details of solving recursive stochastic linear systems with the method of undetermined coefficients.



## Chapter 3

### Forecasting the South African Economy:

### A DSGE-VAR Approach

#### 3.1 Introduction

The controversy about methods for evaluating the empirical relevance of economic models is not new. However, two distinct approaches has emerged since the early 1980s. First, the standard econometric approach in which an economic model should be embedded within a complete probability model and analyzed using statistical methods (Watson, 1993). For instance, Vector Autoregression (VAR) models introduced by Sims (1980), which can be taken directly to the data to perform statistical hypothesis. VAR models also became popular in the forecasting literature pioneered by Litterman (1986b). Although VAR models

have been proved to be reliable tools in terms of data description and forecasting, they are subject to Lucas critique (Lucas, 1976) and also fail to take account of nonlinearities in the economy.

The second approach, pioneered by Kydland and Prescott (1982) and Long and Plosser (1983), has become increasingly popular for evaluating dynamic macroeconomic models. Dynamic stochastic general equilibrium (DSGE) models are explicitly derived from the first principles. DSGE models describe the general equilibrium of a model economy in which agents like consumers and firms maximize their objectives subject to budget and resource constraints (Del Negro and Schorfheide, 2003). Therefore, the DSGE structural (or 'deep') parameters, in principle, do not vary with the policy regime. However, the calibrated DSGE models are typically too stylized to be taken directly to the data and often yield fragile results (Stock and Watson, 2001; Ireland, 2004).

In this paper, we develop an estimated DSGE model for forecasting the Gross National Product (GNP), consumption, investment and hours worked for South African economy. Our proposed hybrid DSGE-VAR model combines a micro-founded DSGE model with the flexibility of a VAR framework. The model is estimated using maximum likelihood technique based on quarterly data obtained from the South African Reserve Bank over the period of 1970:1-2000:4. Based on a recursive estimation using the Kalman filter algorithm, the out-of-sample forecasts from the hybrid model are then compared with the forecasts generated

from the Classical and Bayesian variants of the VAR for the period 2001:1-2005:4.

The remainder of the paper is organized as follows. Section 2 lays out the theoretical model, while Section 3 describes the hybrid model. Results are presented in Section 4 and Section 5 concludes.

### 3.2 The Model Economy

The model economy, here, is based on the benchmark real business cycle model developed by Hansen (1985). Equilibrium models have been criticized for depending heavily on individuals' substitution of leisure and work responding to the change in interest rate or wage. Hansen (1985) and Rogerson (1988) argue that in the real economy labor is indivisible. Individuals either work full time or not at all. Other features of Hansen's indivisible labor are exactly the same as standard real business cycle models, such as Kydland and Prescott (1982). The economic environment is described below.

The model economy is populated by infinitely-lived households. The preferences of households are assumed to be identical. Households maximize expected life-time utility:

$$U(C_t, H_t) = E_t \sum_{t=0}^{\infty} \beta^t (\ln C_t - \gamma H_t), \quad 0 < \beta < 1 \quad \gamma > 0 \quad (1)$$

where  $C_t$  and  $H_t$  are consumption and hours worked respectively,  $\beta$  is the discount factor that households apply to future utility.

The technology is defined as a standard Cobb-Douglas production function with constant-returns-to-scale:

$$Y_t = Z_t K_{t-1}^\rho (\eta^t H_t)^{1-\rho}, \quad 0 < \rho < 1 \quad \eta > 1 \quad (2)$$

where  $\rho$  is the fraction of household's income that goes to the capital input and  $1 - \rho$  is the fraction that goes to the labor input.  $\eta$  measures the gross rate of labor-augmenting technological process.  $Z_t$  is the technology shock, which is exogenously evolving according to the law of motion:

$$\log Z_t = (1 - \psi) \log \bar{Z} + \psi \log Z_{t-1} + \epsilon_t, \quad \epsilon_t \sim i.i.d.(0, \sigma^2) \quad (3)$$

where  $\psi$  and  $\bar{Z}$  are parameters, and  $0 < \psi < 1$ . The innovation  $\epsilon_t$  is normally distributed.

As in a neoclassical growth model, capital stock depreciates at a constant rate of  $\delta$ , and households invest a fraction of income in capital stock in each period. This amount of investment forms part of productive capital in current period. Therefore the law of motion for aggregate capital stock is

$$K_{t+1} = (1 - \delta)K_t + I_t, \quad 0 < \delta < 1 \quad (4)$$

The model economy is a closed economy, where  $Y_t = C_t + I_t$ . In equilibrium the representative consumer maximizes his or her utility function (1) subject to the aggregate constraints

$$Y_t = C_t + I_t$$

$$Y_t = Z_t K_{t-1}^\rho (\eta^t H_t)^{1-\rho}$$

$$K_{t+1} = (1 - \delta)K_t + I_t$$

$$\log Z_t = (1 - \psi)\log \bar{Z} + \psi \log Z_{t-1} + \epsilon_t, \quad \epsilon_t \sim i.i.d.(0, \sigma_\epsilon^2)$$

### 3.3 The Hybrid Model: A DSGE-VAR Approach

Kydland and Prescott (1982) argue that in the basic RBC framework, the U.S. business cycle fluctuations are purely driven by real technology shocks. This one-shock assumption makes real business cycle models stochastically singular. Using a version of the King *et al.* (1988) model, Ingram *et al.* (1994) point out that it is impossible to derive the realizations of the productivity shocks using a singular model if the variance-covariance matrix of the observable variables is actually nonsingular. In order to overcome this singularity problem, Ingram *et al.* (1994), DeJong *et al.* (2000a, b), Ireland (2001 and 2002), and Kim (2000) elaborate the DSGE model to a more elaborate model by including as many shocks as there are endogenous variables in the model. This approach, in addition, can be served to identify sources of output variation<sup>1</sup>.

Recently, Ingram and Whiteman (1994), DeJong *et al.* (2000a, b), and Schorfheide (2000) have used a Bayesian framework to estimate and evaluate

---

<sup>1</sup>The literature suggest that the technology shocks are primarily responsible for the postwar U.S. business cycle fluctuations.

DSGE models. The principle underlying a Bayesian analysis of DSGE models is to combine prior and likelihood functions in order to obtain posterior distributions of the variables of interest. However, different methods have been applied to this kind of research. Ingram and Whiteman (1994) use the King *et al.* (1988) real business cycle model as a source of priors in Bayesian VAR (BVAR) forecasting exercises, whereas, the method pursued by DeJong *et al.* (2000a, b) and Schorfheide (2000) lies between calibration and maximum likelihood estimation exclusively within the DSGE model. Moreover, there is a significant progress in the development of DSGE models that deliver acceptable forecasts (Smets and Wouters, 2003a, b, 2004; Del Negro and Schorfheide, 2004, Del Negro *et al.*, 2005). The authors use prior information derived from DSGE models in the estimation of the VARs. The hybrid models are then used to perform forecasting exercises. The empirical results suggest that the out-of-sample forecasts from the estimated DSGE models outperform the VARs estimated with simple least squares methods.

The approach proposed in this paper is based on Ireland (2004), which is different from the ones discussed above. We augment the linearized solution of the model with unobservable errors that have a VAR representation. This approach was developed originally by Sargent (1989) and pursued by Altug (1989), Watson (1993), Hall (1996), and McGrattan *et al.* (1997). The hybrid DSGE-VAR model is constructed as follows.



The approximated solution is applied to the log-linearized model<sup>2</sup>, where a serially correlated residual is augmented to each equation as in (5)

$$\hat{\pi}_t = A\hat{x}_t + \mu_t \quad (5)$$

and

$$\hat{x}_t = B\hat{x}_{t-1} + C\epsilon_t \quad (6)$$

$$\mu_t = D\mu_{t-1} + \xi_t \quad \xi_t \sim i.i.d.(0, \sigma_\xi^2) \quad (7)$$

where  $\hat{\pi}_t$  is the vector of all de-trended endogenous variables in log-deviations,  $\hat{\pi}_t = [\hat{y}_t \ \hat{c}_t \ \hat{i}_t \ \hat{h}_t]'$ , and  $\hat{x}_t$  is the vector of de-trended state variables in log-deviations,  $\hat{x}_t = [\hat{k}_t \ \hat{z}_t]'$ . The matrix  $D$  is governing the persistence of the VAR residuals. The covariance matrix of the residuals in (7),  $E\xi_t\xi_t' = V$ , is uncorrelated with the innovation to technology,  $\epsilon_t$ . The covariance matrix  $V$  is also constrained to be positive definite (Hamilton, 1994: 147).

Sargent (1989) assumes the measurement errors are uncorrelated with the data generated from the model by restricting  $D$  and  $V$  matrices as diagonal. In this paper, however, we estimate the DSGE model both with and without the restrictions on  $D$  and  $V$  matrices. The advantage of imposing no restrictions on  $D$  and  $V$  matrices is that the residuals in  $\mu_t$  can capture not only the measurement errors, but also the movements and co-movements in the data that the stylized real business cycle model cannot explain (Ireland, 2004: 1210). Furthermore, in

---

<sup>2</sup>Appendix B describes the steady state of the model as well as the the log-linearized model

order to guarantee the residuals in  $\mu_t$  are stationary, the eigenvalues of the matrix  $D$ , which govern the persistence of the VAR residuals, are constrained to be less than one.

The hybrid model is estimated based on quarterly data on real Gross National Product (GNP), consumption, investment and hours worked, for the South African economy, over the period of 1970:1-2000:4. The model economy is a closed economy (i.e.  $Y_t = C_t + I_t$ ), where  $C_t$  and  $I_t$  are defined as final consumption expenditure by households and gross investment respectively<sup>3</sup>. The series are then converted into per-capita form by dividing them with the population aged by 15-64. Since there is no data for hours worked, we generate the series as follows. We assume employees work 40 hours per week and multiply it by the ratio of employment to the labor force.<sup>4</sup>

The hybrid model consisting of (5), (6), and (7) is in state-space form and can be estimated via a maximum likelihood approach. In our real business cycle model, output, consumption, and investment grow at the same rate of  $\eta$  in steady state. Before estimation, the series for output, consumption, and investment are de-trended by dividing with  $\eta$ . In addition, series for  $I_t$  is redundant in estimation since the resource constraint holds by construction in the data. Therefore,  $\hat{\pi}_t, \mu_t$ , and  $\xi_t$  is reduced to  $3 \times 1$  vector:

---

<sup>3</sup>Data are obtained from South African Reserve Bank Quarterly Bulletin, seasonally adjusted at constant price (2000 = 100).

<sup>4</sup>Data for employment is obtained from Statistics South Africa. Population aged 15-64 obtained from World Bank database is used as the proxy of labor force.

$$\hat{\pi}_t = [\hat{y}_t \ \hat{c}_t \ \hat{h}_t]'$$

$$\mu_t = [\mu_{yt} \ \mu_{ct} \ \mu_{ht}]'$$

$$\xi_t = [\xi_{yt} \ \xi_{ct} \ \xi_{ht}]'$$

and for all  $t = 1, 2, 3, \dots$ , the matrices  $D$  and  $V$  are:

$$D = \begin{bmatrix} d_{yy} & d_{yc} & d_{yh} \\ d_{cy} & d_{cc} & d_{ch} \\ d_{hy} & d_{hc} & d_{hh} \end{bmatrix}; \quad V = \begin{bmatrix} v_y^2 & v_{yc} & v_{yh} \\ v_{cy} & v_c^2 & v_{ch} \\ v_{hy} & v_{hc} & v_h^2 \end{bmatrix}$$

The structural parameters,  $\beta, \rho, \eta, \delta$ , and  $\psi$ , are constrained to satisfy the theoretical restrictions discussed in Section 2. The discount factor  $\beta$  and capital depreciation rate  $\delta$  are fixed in the estimation. The discount factor  $\beta$  is set equal to 0.99, as in Hansen (1985), which implies an annual real interest rate of four percent in steady state. The annual aggregate capital depreciation rate  $\delta$  is obtained from annual averaged values of  $\frac{I}{Y}$  and  $\frac{K}{Y}$ . This yields an annual depreciation rate of 0.076, or a quarterly rate of 0.019. The fixed  $\beta$  and  $\delta$  together with the estimated  $\rho, \eta, \gamma$ , and  $z$  help match the steady state values of  $y, c, h$  in the model with those in the data, whereas  $\psi$  and  $\sigma$  only affect the model's dynamics.

### 3.4 Results

In this section, we compare the out-of-sample forecasting performance of the hybrid DSGE-VAR model with the VARs, both Classical and Bayesian, in terms of the Root Mean Squared Errors (RMSEs). At this stage, a few words need to be said regarding the choice of the evaluation criterion for the out-of-sample forecasts generated from Bayesian models. As Zellner (1986: 494) points out “the optimal Bayesian forecasts will differ depending upon the loss function employed and the form of predictive probability density function”. In other words, Bayesian forecasts are sensitive to the choice of the measure used to evaluate the out-of-sample forecast errors. This fact was also observed in a recent study by Gupta (2006). However, Zellner (1986) points out that the use of the mean of the predictive probability density function for a series, is optimal relative to a squared error loss function and the Mean Squared Error (MSE), and, hence, the RMSE is an appropriate measure to evaluate performance of forecasts, when the mean of the predictive probability density function is used. This is exactly what we do below in Tables 8 through 11, when we use the average RMSEs over the one- to four-quarter-ahead forecasting horizon.

But, before we proceed to the discussion of the forecasting performance of the alternative models, it is important to lay out the basic structural differences and advantages of using BVARs over traditional VARs for forecasting.

### 3.4.1 Classical and Bayesian VARs

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

$$\chi_t = C + \lambda(L)\chi_t + \varepsilon_t \quad (8)$$

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

A crucial drawback of the VAR forecasts is “overfitting” due to the inclusion too many lags and too many variables, some of which may be insignificant. The problem of “overfitting” results in multicollinearity and loss of degrees of freedom, leads to inefficient estimates and large out-of-sample forecasting errors. Thus, it can be argued the performance of VAR forecasts will deteriorate rapidly as the forecasting horizon becomes longer.

A forecaster can overcome this “overfitting” problem by using Bayesian techniques. The motivation for the Bayesian analysis is based on the knowledge that more recent values of a variable are more likely to contain useful information about its future movements than older values. From a Bayesian perspective, the

exclusion restriction in the VAR is an inclusion of a coefficient without a prior probability distribution (Litterman, 1986a).

The Bayesian model proposed by Litterman (1981), Doan, *et al.* (1984), and Litterman (1986b), imposes restrictions on those coefficients by assuming they are more likely to be near zero. The restrictions are imposed by specifying normal prior<sup>5</sup> distributions with zero means and small standard deviations for all the coefficients with standard deviation decreasing as lag increases. One exception is that the mean of the first own lag of a variable is set equal to unity to reflect the assumption that own lags account for most of the variation of the given variable. To illustrate the Bayesian technique, suppose the “Minnesota prior” means and variances take the following form:

$$\begin{aligned}
 \beta_i &\sim N(1, \sigma_{\beta_i}^2) \\
 \beta_j &\sim N(0, \sigma_{\beta_j}^2)
 \end{aligned}
 \tag{9}$$

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

---

<sup>5</sup>Note Litterman (1981) uses a diffuse prior for the constant, which is popularly referred to as the “Minnesota prior” due to its development at the University of Minnesota and the Federal Reserve bank at Minneapolis.

Doan *et al.* (1984) propose a formula to generate standard deviations as a function of a small number of *hyperparameters*<sup>6</sup> :  $w$ ,  $d$ , and a weighting matrix  $f(i, j)$ . This approach allows the forecaster to specify individual prior variances for a large number of coefficients based on only a few *hyperparameters*. The specification of the standard deviation of the distribution of the prior imposed on variable  $j$  in equation  $i$  at lag  $m$ , for all  $i, j$  and  $m$ , defined as  $S(i, j, m)$ :

$$S(i, j, m) = [w \times g(m) \times f(i, j)] \frac{\hat{\sigma}_i}{\hat{\sigma}_j} \quad (10)$$

where:

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

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

The term  $w$  is the measurement of standard deviation on the first own lag, which indicates the overall tightness. A decrease in the value of  $w$  results a tighter prior. The parameter  $g(m)$  measures the tightness on lag  $m$  relative to lag 1, and is assumed to have a harmonic shape with a decay of  $d$ . An increasing in  $d$ , tightens the prior as lag increases.<sup>7</sup> The parameter  $f(i, j)$  represents the tightness of variable  $j$  in equation  $i$  relative to variable  $i$ . Reducing the interaction parameter

---

<sup>6</sup>The name of *hyperparameter* is to distinguish it from the estimated coefficients, the parameters of the model itself.

<sup>7</sup>In this paper, we set the overall tightness parameter ( $w$ ) equal to 0.3, 0.2, and 0.1, and the harmonic lag decay parameter ( $d$ ) equal to 0.5, 1, and 2. These parameter values are chosen so that they are consistent with the ones that used by Liu and Gupta (2007).

$k_{ij}$  tightens the prior.  $\hat{\sigma}_i$  and  $\hat{\sigma}_j$  are the estimated standard errors of the univariate autoregression for variable  $i$  and  $j$  respectively. In the case of  $i \neq j$ , the standard deviations of the coefficients on lags are not scale invariant (Litterman, 1986b: 30). The ratio,  $\frac{\hat{\sigma}_i}{\hat{\sigma}_j}$  in (10), scales the variables so as to account for differences in the units of magnitudes of the variables.

The BVAR model is estimated using Theil’s (1971) mixed estimation technique, which involves supplementing the data with prior information on the distribution of the coefficients. For each restriction imposed on the parameter estimated, the number of observations and degrees of freedom are increased by one in an artificial way. Therefore, the loss of degrees of freedom associated with the unrestricted VAR is not a concern in the BVAR.

### 3.4.2 Forecast accuracy

Table 8 to 11 report the RMSEs from the hybrid DSGE-VAR model along with the VARs. The hybrid model does better job in predicting output and its components than it does in predicting hours worked.<sup>8</sup> To be more precise, for output and consumption the unconstrained hybrid model does better than the constrained hybrid model and the unrestricted VAR. However, for hours worked

---

<sup>8</sup>The hybrid model has 21 parameters, the six structural parameters  $\gamma, \rho, \eta, z, \psi$ , and  $\sigma$  from the real business cycle model, the fifteen elements from matrix D and V governing the the behavior of the VAR residuals. For the constrained hybrid model, the number of parameters is reduced to 12. The VAR(1) model that we use to judge the hybrid model’s out-of-sample forecasting performance also has 21 parameters, output, consumption, and hours worked together with a constant and a linear time trend.



Table 8: **RMSE (2001Q1-2005Q4): Output**

QA	1	2	3	4	AVE
DSGE-VAR (Uncon)	<i>1.2432</i>	<i>1.7841</i>	<i>1.8214</i>	<i>1.7765</i>	<i>1.6563</i>
DSGE-VAR (Con)	1.3671	2.0595	2.4382	2.8072	2.1680
VAR (1)	1.4611	2.3092	2.8747	3.4087	2.5134
BVAR ( $w=.3, d=.5$ )	<b>0.6698</b>	<b>1.0454</b>	<b>1.3164</b>	<b>1.5712</b>	<b>1.1507</b>

QA: quarter ahead; RMSE: root mean squared error (%).

the constrained hybrid model outperforms the unconstrained one but not the unrestricted VAR. The scenario for investment is a bit different. The unconstrained hybrid model does better than the constrained one for only the one-quarter and two-quarters ahead out-of-sample forecasts, whereas for the three-quarters and four-quarters ahead forecasts the constrained hybrid model outperforms the unconstrained one.

As far as the forecasting performances of the BVARs are concerned, it is clear that the BVARs improve the out-of-sample forecast performance significantly. The RMSEs<sup>9</sup> generated from the BVARs are much smaller than those generated from both the hybrid model and the unrestricted VARs. In addition, the result suggests that a BVAR with a relatively loose prior produces smaller forecast errors. For all variables, output, consumption, investment and hours worked, a BVAR with the most loose prior ( $w = 0.3, d = 0.5$ ) performs the best.

---

<sup>9</sup>Here we only report the BVAR with the prior that does the best in terms of the out-of-sample forecasting performance.

Table 9: **RMSE (2001Q1-2005Q4): Consumption**

QA	1	2	3	4	AVE
DSGE-VAR (Uncon)	<i>1.2001</i>	<i>1.7884</i>	<i>2.1229</i>	<i>2.2967</i>	<i>1.8520</i>
DSGE-VAR (Con)	1.2287	1.9548	2.5158	3.0207	2.1800
VAR (1)	1.2029	1.7833	2.181	2.4643	1.9079
BVAR (w=.3, d=.5)	<b>0.5215</b>	<b>0.7080</b>	<b>0.8293</b>	<b>0.8570</b>	<b>0.7290</b>

QA: quarter ahead; RMSE: root mean squared error (%).

Table 10: **RMSE (2001Q1-2005Q4): Investment**

QA	1	2	3	4	AVE
DSGE-VAR (Uncon)	<i>2.8404</i>	<i>3.5985</i>	4.1179	4.1522	3.6773
DSGE-VAR (Con)	2.9518	3.6293	<i>3.9484</i>	<i>4.0228</i>	<i>3.6381</i>
VAR (1)	3.0437	4.3241	5.5072	6.4486	4.8309
BVAR (w=.3, d=5)	<b>1.1230</b>	<b>1.4757</b>	<b>1.8097</b>	<b>2.0608</b>	<b>1.6173</b>

QA: quarter ahead; RMSE: root mean squared error (%).

Table 11: **RMSE (2001Q1-2005Q4): Hours worked**

QA	1	2	3	4	AVE
DSGE-VAR (Uncon)	2.5066	3.3475	4.0577	4.5857	3.6244
DSGE-VAR (Con)	<i>2.4477</i>	<i>2.9966</i>	<i>3.5075</i>	<i>3.7018</i>	<i>3.1634</i>
VAR (1)	2.3913	2.941	3.2884	3.2920	2.9782
BVAR5 (w=.3, d=.5)	<b>1.2420</b>	<b>1.6435</b>	<b>1.8927</b>	<b>1.9342</b>	<b>1.6781</b>

QA: quarter ahead; RMSE: root mean squared error (%).

In order to evaluate the models' forecast accuracy, we perform the across-model test between the hybrid model and the VAR(1), as well as the BVAR model. The across-model test is based on the statistic proposed by Diebold and Mariano (1995).<sup>10</sup> The across-model test results are reported in Table 12. The results indicate that, in general, the hybrid models outperform the unrestricted VAR(1) model for forecasting output and its components. One exception is consumption, the constrained hybrid model does not outperform the unrestricted VAR(1) model. However, most of these test statistics are not significant at 5% level. As far as the forecasting performance of the BVAR is concerned, the BVAR with the most loose prior ( $w = 0.3$ ,  $d = 0.5$ ) outperforms the hybrid models and the unrestricted VAR(1) model. In addition, most of these test statistics are significant either at 5% or 10% level. Finally, for hours worked, both constrained and unconstrained hybrid model do not outperform either the unrestricted VAR(1) model or the BVAR model, although few of the statistics are significant at 10% level.

---

<sup>10</sup>The test statistic is defined as the following. For instance, let  $\{e_t^v\}_{t=1}^T$  denote the associated forecast errors from the unrestricted VAR(1) model and  $\{e_t^h\}_{t=1}^T$  denote the forecast errors from the hybrid model. The test statistic is then defined as  $s = \frac{l}{\sigma_l}$ , where  $l$  is the sample mean of the "loss differentials",  $\{l_t\}_{t=1}^T$ , using  $l_t = (e_t^v)^2 - (e_t^h)^2$  for all  $t = 1, 2, 3, \dots, T$ , and where  $\sigma_l$  is the standard error of  $l$ . The  $s$  statistic is asymptotically distributed as a standard normal random variable and can be estimated under the null hypothesis of equal forecast accuracy, i.e.  $l = 0$ . Therefore, in this case, a positive value of  $s$  suggests that the hybrid model outperforms the unrestricted VAR(1) model in terms of out-of-sample forecasting.

Table 12: **Across-Model Test Statistics**

Quarters Ahead	1	2	3	4
(A) Output				
DSGE-VAR (Uncon) vs. VAR(1)	1.898*	1.823*	1.848*	1.616
DSGE-VAR (Con) vs. VAR(1)	0.888	1.718	1.592	1.576
DSGE-VAR (Uncon) vs. BVAR	-2.579**	-2.287**	-1.819*	-1.501
DSGE-VAR (Con) vs. BVAR	-2.740**	-2.310**	-1.916*	-1.598
VAR(1) vs. BVAR	-2.566**	-2.222**	-1.907*	-1.657
(B) Consumption				
DSGE-VAR (Uncon) vs. VAR(1)	0.126	-0.103	0.581	0.762
DSGE-VAR (Con) vs. VAR(1)	-0.418	-0.842	-1.055	-1.149
DSGE-VAR (Uncon) vs. BVAR	-3.267**	-1.935*	-1.643	-1.408
DSGE-VAR (Con) vs. BVAR	-3.760**	-2.035**	-1.859*	-1.499
VAR(1) vs. BVAR	-3.324**	-1.765*	-1.472	-1.197
(C) Investment				
DSGE-VAR (Uncon) vs. VAR(1)	0.604	0.985	1.093	1.166
DSGE-VAR (Con) vs. VAR(1)	0.329	1.568	1.633	1.466
DSGE-VAR (Uncon) vs. BVAR	-2.716**	-2.035**	-1.404	-1.148
DSGE-VAR (Con) vs. BVAR	-2.762**	-2.444**	-1.733*	-1.283
VAR(1) vs. BVAR	-2.394**	-2.086**	-1.605	-1.383
(D) Hours Worked				
DSGE-VAR (Uncon) vs. VAR(1)	-0.522	-0.915	-1.014	-0.976
DSGE-VAR (Con) vs. VAR(1)	-1.117	-0.717	-1.024	-1.208
DSGE-VAR (Uncon) vs. BVAR	-1.933*	-1.535	-1.321	-1.132
DSGE-VAR (Con) vs. BVAR	-1.947*	-1.686	-1.461	-1.345
VAR(1) vs. BVAR	-1.968*	-1.727*	-1.490	-1.329

Note: \* and \*\* indicate 10% and 5% significant respectively. BVAR is the optimal one with  $w = 0.3$  and  $d = 0.5$ .

### 3.5 Conclusion

In this paper, we develop an estimable DSGE model, in which we augment the linearized equations with a vector of residuals that follow a AR(1) process. The hybrid model, thus, combines the micro-founded DSGE model with the flexibility of the atheoretical VAR model, and hence, the name — DSGE-VAR. We then employ the hybrid model to measure the out-of-sample forecasting performance for output, consumption, investment, and hours worked for the South African economy over 2001:1-2005:4. The results indicate that, in general, the estimated hybrid DSGE model outperforms the Classical VAR, but not the Bayesian VARs. Moreover, the results suggest that a BVAR with a relatively loose prior produces smaller out-of-sample forecast errors.

The Hansen's (1985) version real business cycle model used in this paper is singular in the sense that the technology shock is the only shock to the system. Therefore, it is necessary to study the importance of various shocks in accounting for the dynamic behaviour of output and its main components. In this regard, future research aims to estimate a New Keynesian DSGE model, which will allow us to incorporate nominal shocks. Further, we also aim to estimate the current model using Bayesian techniques. The ultimate goal of all these future extensions will be to analyze whether the DSGE model can outperform the BVARs, as far as forecasting is concerned.

## A. Optimization

In our model economy, the representative consumer problem is to maximize the utility function (1) by choosing  $\{C_t, H_t, K_{t+1}\}_{t=0}^{\infty}$

$$U(C_t, H_t) = E_t \sum_{t=0}^{\infty} \beta^t (\ln C_t - \gamma H_t), \quad 0 < \beta < 1 \quad \gamma > 0$$

subject to the resource constraint:

$$Z_t K_t^\rho (\eta^t H_t)^{1-\rho} \geq C_t + K_{t+1} - (1 - \delta)K_t \quad (\text{A.1})$$

From the (A.1), we have:

$$C_t = Z_t K_t^\rho (\eta^t H_t)^{1-\rho} + (1 - \delta)K_t - K_{t+1} \quad (\text{A.2})$$

The Bellman equation for this problem:

$$V(K_t, Z_t) = \max_{H_t, K_{t+1}} \{ \ln [Z_t K_t^\rho (\eta^t H_t)^{1-\rho} + (1 - \delta)K_t - K_{t+1}] - \gamma H_t \} + \beta E_t V(K_{t+1}, Z_{t+1}) \quad (\text{A.3})$$

The first order condition (FOC) for hours worked:

$$\frac{\partial V(K_t, Z_t)}{\partial H_t} = 0 \quad (\text{A.4})$$

$$\frac{1}{C_t} (1 - \rho) Z_t K_t^\rho \eta^{t(1-\rho)} H_t^{-\rho} - \gamma = 0 \quad (\text{A.5})$$

$$\gamma = \frac{Y_t}{C_t} (1 - \rho) \frac{1}{H_t} \quad (\text{A.6})$$

and the FOC for capital stock:

$$\frac{\partial V(K_t, Z_t)}{\partial K_{t+1}} = 0 \quad (\text{A.7})$$

$$\frac{1}{C_t}(-1) + \beta E_t V(K_{t+1}, Z_{t+1}) = 0 \quad (\text{A.8})$$

The envelope condition is:

$$\frac{\partial V(K_t, Z_t)}{\partial K_t} = \frac{1}{C_t} Z_t \rho K_t^{\rho-1} (\eta^t H_t)^{1-\rho} + (1 - \delta) \quad (\text{A.9})$$

Updating (A.9) and combining with (A.8) yields the Euler equation for capital stock:

$$\frac{1}{C_t} = \beta E_t \left[ \frac{1}{C_{t+1}} Z_{t+1} \rho K_{t+1}^{\rho-1} (\eta^{t+1} H_{t+1})^{1-\rho} + (1 - \delta) \right] \quad (\text{A.10})$$

$$\frac{1}{C_t} = \beta E_t \left\{ \frac{1}{C_{t+1}} \left[ \frac{Y_{t+1}}{K_{t+1}} \rho + (1 - \delta) \right] \right\} \quad (\text{A.11})$$

## B. The steady state and log-linearization

### B.1 The steady state

The complete model economy:

$$Y_t = C_t + I_t \quad (\text{B.1})$$

$$Y_t = Z_t K_t^\rho (\eta^t H_t)^{1-\rho} \quad (\text{B.2})$$

$$K_{t+1} = (1 - \delta)K_t + I_t \quad (\text{B.3})$$

$$\gamma = \frac{Y_t}{C_t} (1 - \rho) \frac{1}{H_t} \quad (\text{B.4})$$

$$\frac{1}{C_t} = \beta E_t \left\{ \frac{1}{C_{t+1}} \left[ \left( \frac{Y_{t+1}}{K_{t+1}} \right)^\rho + (1 - \delta) \right] \right\} \quad (\text{B.5})$$

$$\log Z_t = (1 - \psi) \log \bar{Z} + \psi \log Z_{t-1} + \epsilon_t, \quad \epsilon_t \sim i.i.d.(0, \sigma_\epsilon^2) \quad (\text{B.6})$$

In equilibrium,  $y_t = Y_t/\eta^t$ ,  $c_t = C_t/\eta^t$ ,  $i_t = I_t/\eta^t$ ,  $h_t = H_t$ ,  $k_t = K_t/\eta^t$ , and  $z_t = Z_t$ , therefore we can rewrite the model as:

$$y_t = c_t + i_t$$

$$y_t = z_t k_t^\rho h_t^{1-\rho}$$

$$\eta k_{t+1} = (1 - \delta)k_t + i_t$$

$$\gamma = \frac{y_t}{c_t} (1 - \rho) \frac{1}{h_t}$$

$$\frac{\eta}{c_t} = \beta E_t \left\{ \frac{1}{c_{t+1}} \left[ \left( \frac{y_{t+1}}{k_{t+1}} \right)^\rho + (1 - \delta) \right] \right\}$$

$$\log z_t = (1 - \psi) \log \bar{Z} + \psi \log z_{t-1} + \epsilon_t, \quad \epsilon_t \sim i.i.d.(0, \sigma_\epsilon^2)$$

In steady state we have  $y_t = y$ ,  $c_t = c$ ,  $i_t = i$ ,  $h_t = h$ ,  $k_t = k$ , and  $z_t = z$  for all  $t = 0, 1, 2, \dots$ . Solving for the steady state values of the six variables:



$$a = A \tag{B.7}$$

$$k = \left( \frac{\rho}{\eta/\beta - 1 + \delta} \right) y \tag{B.8}$$

$$i = \left[ \frac{\rho(\eta - 1 + \delta)}{\eta/\beta - 1 + \delta} \right] y \tag{B.9}$$

$$c = \left\{ 1 - \left[ \frac{\rho(\eta - 1 + \delta)}{\eta/\beta - 1 + \delta} \right] \right\} y \tag{B.10}$$

$$h = \left( \frac{1 - \rho}{\gamma} \right) \left\{ 1 - \left[ \frac{\rho(\eta - 1 + \delta)}{\eta/\beta - 1 + \delta} \right] \right\}^{-1} \tag{B.11}$$

$$y = z^{1/(1-\rho)} \left( \frac{\rho}{\eta/\beta - 1 + \delta} \right)^{\rho/(1-\rho)} \left( \frac{1 - \rho}{\gamma} \right) \left\{ 1 - \left[ \frac{\rho(\eta - 1 + \delta)}{\eta/\beta - 1 + \delta} \right] \right\}^{-1} \tag{B.12}$$

## B.2 Log-linearization

This section presents the log-linearized DSGE model. The principle of log-linearization is to replace all equations by Taylor approximation around the steady state, which are linear functions in the log-deviations of the variables (Uhlig, 1995:4). Suppose  $\Pi_t$  be the vector of variables,  $\pi$  their steady state, and  $\hat{\pi}_t$  the vector of log-deviations:

$$\hat{\pi}_t = \log \Pi_t - \log \pi \tag{B.13}$$

in other words,  $\hat{\pi}_t$  denote the percentage deviations from their steady state levels.

Using first-order Taylor approximations to rewrite all the equations of the model:

$$\hat{y}_t = \hat{z}_t + \rho \hat{k}_t + (1 - \rho) \hat{h}_t \quad (\text{B.14})$$

$$\hat{z}_t = \psi \hat{z}_{t-1} + \varepsilon_t \quad (\text{B.15})$$

$$\left(\frac{\eta}{\beta} - 1 + \delta\right) \hat{y}_t = \left[\left(\frac{\eta}{\beta} - 1 + \delta\right) - \rho(\eta - 1 + \delta)\right] \hat{c}_t + \rho(\eta - 1 + \delta) \hat{i}_t \quad (\text{B.16})$$

$$\eta \hat{k}_{t+1} = (1 - \delta) \hat{k}_t + (\eta - 1 + \delta) \hat{i}_t \quad (\text{B.17})$$

$$\hat{c}_t + \hat{h}_t = \hat{y}_t \quad (\text{B.18})$$

$$\begin{aligned}
 0 = & \frac{\eta}{\beta} \hat{c}_t - \frac{\eta}{\beta} E_t \hat{c}_{t+1} + \left(\frac{\eta}{\beta} - 1 + \delta\right) E_t \hat{y}_{t+1} \\
 & - \left(\frac{\eta}{\beta} - 1 + \delta\right) \hat{k}_{t+1}
 \end{aligned} \quad (\text{B.19})$$



## Chapter 4

# A New-Keynesian DSGE Model for Forecasting the South African Economy

### 4.1 Introduction

The objective of this paper is to develop a New-Keynesian Dynamic Stochastic General Equilibrium (NKDSGE) Model for forecasting growth rate of output, inflation, and a measure of nominal short-term interest rate, in our case the 91-days Treasury Bills rate, for South African economy. The model is estimated via maximum likelihood technique for quarterly data over the period of 1970:1-2000:4. Based on a recursive estimation using the Kalman filter algorithm, the out-of-sample forecasts from the NKDSGE model are then compared with the same generated from the Classical and Bayesian variants of the VAR models for the period 2001:1-2006:4.

During the last three decades, lot of work has gone into developing well-structured New-Keynesian-Macroeconomic (NKM) models in response to criticisms on the traditional, once-dominant, IS-LM framework of macroeconomic analysis. The NKM models incorporate the nominal (price and/or wage) rigidities into the traditional IS-LM framework to capture the time series properties of the data. More recently, the so called new generation NKM models (Goodfriend and King, 1997; Rotemberg and Woodford, 1997; McCallum and Nelson, 1999, Smets and Wouters, 2003) that are built on a dynamic stochastic general equilibrium framework, based on optimizing behavior of agents, has also gained tremendous prominence. However, this type of micro-founded NKM models have generally been used for policy analysis, few being used for forecasting purposes. One exception in this regard is the study by Smets and Wouters (2004). The authors develop and estimate a micro-founded NKM model with sticky prices and wages for the Euro area. The results indicate that the forecasting performance of NKM model is reasonably well comparable to the atheoretical VAR.

In a recent paper, Liu *et al.* (2007) develop and estimate a Hansen(1985)-type hybrid model for forecasting the South African economy. The hybrid model is based on a real business cycle (RBC) framework. Kydland and Prescott (1982) argue that in the basic RBC framework, the U.S. business cycle fluctuations are purely driven by real technology shocks. This one-shock assumption makes RBC models stochastically singular. In order to overcome this singularity problem, the authors augment the theoretical model with unobservable errors having a

VAR representation. This allows one to combine the theoretical rigor of a DSGE model with the flexibility of an atheoretical VAR model. The results indicate that the estimated hybrid DSGE model outperforms the Classical VAR, but not the Bayesian VARs in terms of out-of-sample forecasting performances. Having resorted to a RBC framework, prevents Liu *et al.* (2007) from analyzing the role of nominal shocks. This is, in our opinion, inappropriate for the South African economy, since South African economy, just as other developing economies, is subject to nominal shocks.

In this paper, following Rotemberg and Woodford (1997) and Ireland (2004), we develop and estimate a NKDSGE model with sticky prices. The model consists of three equations, an expectational IS curve, a forward-looking version of the Phillips curve, and a Taylor-type monetary policy rule. Furthermore, the model is characterized by four shocks: a preference shock; a technology shock; a cost-push shock; and a monetary policy shock. Essentially, by incorporating four shocks, that generally tends to affect a macroeconomy, we attempt to model the empirical stochastics and dynamics in the data better, and hence, improve the predictions. In addition, using a NKDSGE model, allows us to model product market rigidities, which is also an important feature of the South African economy. Further allowing for explicit interest rate rules also helps in modelling the inflation targeting frame regime of the South African economy, understanding better in comparison to the RBC model for obvious reason.

The rest of the paper is structured as follows. Section 2 lays out the theoretical model, while Section 3 shows the solution of the model. Results are presented in Section 4 and Section 5 concludes.

## 4.2 The Model

### 4.2.1 The Representative Household

The economy consists of a continuum of infinitely-lived households. In each period  $t = 0, 1, 2, \dots$ , a representative household makes a sequence of decisions to maximize the expected utility over a composite consumption good  $C_t$ , real money balance  $M_t/P_t$ , and leisure  $1 - h_t$ :

$$E \sum_{t=0}^{\infty} \beta^t [a_t \log(C_t) + \log\left(\frac{M_t}{P_t}\right) - \left(\frac{1}{\eta}\right) h_t^\eta], \quad 0 < \beta < 1, \quad \eta \geq 1, \quad (1)$$

where  $\beta$  is the subjective discount factor and  $a_t$  is the preference shock which follows an AR(1) process as in Ireland (2004):

$$\log(a_t) = \rho_a \log(a_{t-1}) + \varepsilon_{at}, \quad 0 \leq \rho_a < 1, \quad \varepsilon_{at} \sim i.i.d.(0, \sigma_a^2), \quad (2)$$

The representative household carries money  $M_{t-1}$  and bonds  $B_{t-1}$  from the previous period into the current period  $t$ . In time period  $t$ , the household receives a lump-sum transfer  $T_t$  from the monetary authority and the nominal profit or dividend payment  $D_t$  from the intermediate good firms. In addition, the household also receives its usual labor income  $W_t h_t$ , where  $W_t$  denotes the nominal

wage. Therefore, in each time period the representative household maximizes its expected utility (1) by choosing consumption, labor supply, money and bond, subject to the following budget constraint:

$$C_t + \frac{B_t}{r_t P_t} = \frac{W_t}{P_t} h_t + \frac{B_{t-1}}{P_t} + D_t + T_t - \frac{M_t - M_{t-1}}{P_t} \quad (3)$$

where  $r_t$  denotes the gross nominal interest rate and  $P_t$  denotes the nominal price.

In this version of NKDSGE model, capital accumulation decision is ignored. Christiano *et al.* (2005) assume that the household owns capital stock and makes capital accumulation and utilization decisions in each time period<sup>1</sup>. However, as noted that there exists little relationship between capital stock and output at business cycle frequencies (McCallum and Nelson, 1999; Cogley and Nason, 1995), the role of capital has been ignored here.

Given (1) and (2), the representative household's first order conditions are as follows:

$$\frac{W_t}{P_t} = a_t^{-1} C_t h_t^{\eta-1} \quad (4)$$

$$\frac{a_t}{C_t} = r_t \beta E_t \left[ a_{t+1} \left( \frac{1}{C_{t+1}} \right) \left( \frac{P_t}{P_{t+1}} \right) \right] \quad (5)$$

$$\frac{M_t}{P_t} = a_t^{-1} C_t [r_t / (r_t - 1)] \quad (6)$$

---

<sup>1</sup>For further details on capital accumulation and utilization in NKDSGE models see Dostey and King (2001), and Smets and Wouters (2003).

where (4) is the intratemporal optimality condition, capturing the consumption and leisure trade-off, i.e. the marginal rate of substitution between consumption and leisure equals to the real wage. Equation (5) represents the intertemporal allocation of consumption, whereas (6) is the money demand equation. It shows that the optimal condition of money holding requires that the marginal rate of substitution between money and consumption must equalize with the opportunity cost of holding money.

#### 4.2.2 Final-Goods Production

In the final-goods sector, a representative firm produces the final good  $Y_t$  according to a constant elasticity of substitution (CES) production function as suggested by Dixit and Stiglitz (1977):

$$Y_t = \left( \int_0^1 Y_{jt}^{\frac{\theta_t-1}{\theta_t}} dj \right)^{\frac{\theta_t}{\theta_t-1}}, \quad \theta_t > 1, \quad (7)$$

where  $Y_{jt}$  denotes the output of intermediate good  $j$  which the representative final-goods firm uses as input to produce  $Y_t$  units of final goods.

Given the intermediate-goods price  $P_{jt}$ ,<sup>2</sup> firm maximizes its profits:

$$\max_{Y_{jt}} \left\{ Y_t - \frac{1}{P_t} \int_0^1 P_{jt} Y_{jt} \right\} \quad (8)$$

Solving the firm's profits maximization problem (8), yields:

---

<sup>2</sup>As explained in section 2.3, the representative intermediate-goods firm is assumed to sell its output in a monopolistically competitive market.



$$Y_{jt} = \left( \frac{P_{jt}}{P_t} \right)^{-\theta_t} Y_t \quad (9)$$

Since the final-goods firms operate in a perfectly competition, in equilibrium the representative firm's profit should equal to zero. Hence, the equilibrium market price for final good is given as follows:

$$P_t = \left( \int_0^1 P_{jt}^{1-\theta_t} dj \right)^{\frac{1}{1-\theta_t}} \quad (10)$$

It is important to point out that the production function (7) implies a constant elasticity of substitution between intermediate goods.  $\theta_t$  is a stochastic parameter determining the time-varying mark-up in the goods market (Smets and Wouters, 2003; Ireland, 2004). This is a convenient way to introduce the so called mark-up or cost-push shocks into the NKDSGE model as proposed by Clarida *et al.* (1999). The cost-push shock follows the following autoregressive process:

$$\log\theta_t = (1-\rho_\theta)\log\theta + \rho_\theta\log\theta_{t-1} + \varepsilon_{\theta t}, \quad 0 \leq \rho_\theta < 1, \quad \varepsilon_{\theta t} \sim i.i.d.(0, \sigma_\theta^2), \quad (11)$$

where the serially uncorrelated innovation  $\varepsilon_{\theta t}$  is normally distributed.

### 4.2.3 Intermediate-Goods Production

In the intermediate-goods sector, firms are monopolistically competitive and face a quadratic cost of price adjustment. In each time period, the representative intermediate-goods firm hires  $h_{jt}$  units of labor and produces  $Y_{jt}$  units of intermediate good  $j$ , according to the following technology:

$$Y_{jt} = Z_t h_{jt} \quad (12)$$

This is a standard constant-return-to-scale production function, but without capital.  $Z_t$  is the aggregate technology shock, which is assumed follow a random walk with a positive drift:

$$\log Z_t = \log \bar{Z} + \log Z_{t-1} + \varepsilon_{zt}, \quad \varepsilon_{zt} \sim i.i.d.(0, \sigma_z^2), \quad (13)$$

where  $\bar{Z} > 1$  and the serially uncorrelated innovation  $\varepsilon_{zt}$  is normally distributed. In equilibrium, this supply-side disturbance acts as a shock to the Phillips curve in the NKDSGE model (Ireland, 2001).

As stated above, the representative intermediate-goods firm faces a quadratic cost of nominal price adjustment along the line of Rotemberg (1982). Mathematically, we have:

$$\frac{\phi}{2} \left[ \frac{P_{jt}}{\pi P_{jt-1}} \right]^2 Y_t, \quad \phi > 0, \quad \pi > 1, \quad (14)$$

where  $\phi$  is the parameter that governs the magnitude of the cost of price adjustment and  $\pi$  is the steady-state gross rate of inflation.

Since the representative intermediate-goods firm operates in a monopolistically competitive market, it chooses its own sale price  $P_{jt}$  taking as given a downward sloping demand curve in order to maximize its market value:

$$E \sum_{t=0}^{\infty} \beta^t (a_t/C_t) \left\{ \left[ \frac{P_{jt}}{P_t} \right]^{1-\theta} Y_t - \left[ \frac{P_{jt}}{P_t} \right]^{-\theta} \left( \frac{W_t}{P_t} \right) \left( \frac{Y_t}{Z_t} \right) - \frac{\phi}{2} \left[ \frac{P_{jt}}{\pi P_{jt-1}} \right]^2 Y_t \right\} \quad (15)$$

where  $\beta^t (a_t/C_t)$  measures the representative household's marginal utility of an additional unit of real profit generated in time period  $t$ . The first order condition is:

$$\begin{aligned} (\theta_t - 1) \left( \frac{P_{jt}}{P_t} \right)^{-\theta_t} \left( \frac{Y_t}{P_t} \right) &= \beta \phi E_t \left[ \left( \frac{a_{t+1}}{a_t} \right) \left( \frac{C_t}{C_{t+1}} \right) \left( \frac{P_{jt+1}}{\pi P_{jt}} - 1 \right) \left( \frac{Y_{t+1}}{P_{jt}} \right) \left( \frac{P_{jt+1}}{\pi P_{jt}} \right) \right] \\ &+ \left[ \theta_t \left( \frac{P_{jt}}{P_t} \right)^{-\theta_t-1} \left( \frac{W_t}{P_t} \right) \left( \frac{Y_t}{Z_t} \right) \left( \frac{1}{P_t} \right) \right] - \left[ \phi \left( \frac{P_{jt}}{\pi P_{jt-1}} - 1 \right) \left( \frac{Y_t}{\pi P_{jt-1}} \right) \right] \end{aligned} \quad (16)$$

The representative intermediate-goods firm sets its markup price  $P_{jt}$  in such a way that the actual markup price will differ from, but tend to gravitate towards, the desired markup overtime (Ireland, 2004: 9).

#### 4.2.4 The Monetary Authority

The model is closed by assuming that the monetary authority follows a modified Taylor (1993) rule. That is, the monetary authority adjusts its instrument, the nominal short-term interest rate, in response to deviations of inflation and

output from their steady-state levels, as well as lagged deviations of interest rate and deviations of current growth rate.

$$\hat{r}_t = \rho_r \hat{r}_{t-1} + \rho_\pi \hat{\pi}_t + \rho_g \hat{g}_t + \rho_x \hat{x}_t + \varepsilon_{rt}, \quad \varepsilon_{rt} \sim i.i.d.(0, \sigma_r^2), \quad (17)$$

$r_t$  is the nominal short-term interest rate,  $g_t$  output growth, and  $x_t$  output gap<sup>3</sup>. The  $\varepsilon_{rt}$ 's represent exogenous monetary policy shocks, which are assumed to be serially uncorrelated.

Monetary policy rules are often preferred over discretionary decisions. A formal rule is the desire for governance “by laws, not by means”, as well as, the way to overcome “dynamic inconsistency” (Barro and Gordon, 1983; Rogoff, 1985). From a monetary transmission mechanism point of view, monetary policy affects the target variable(s) and the economy mainly through the private-sector expectations of the future interest rates, inflation, and output. Since growth rate of output is public knowledge, besides output gap, we include output growth in our interest rate rule as well. Moreover, output growth can be one of the most important and observable indicator, as apposed to the more elaborated output gap, that the monetary authority responds to.

The measure of output gap associated with NKM model differs from the empirical (statistical) approach. The empirical approach essentially involves detrending output from its smooth trend. It requires using either a univariate technique like the Hodrick-Prescott filter or a multivariate technique like adapted

---

<sup>3</sup>A letter with a hat above indicates its deviation.

multivariate filter to determine the smooth trend – potential output <sup>4</sup> . However, the main properties of the resulting series, the potential output, do not seem to hinge critically on the exact techniques used. Moreover, the use of detrended output as a proxy for the output gap has been criticized due to the lack of theoretical justification (Gali, 2002). Using a simple estimated linear model, Smets (1998) shows that output gap uncertainty can have a significant effect on the efficient response coefficients in Taylor-type rules for the US economy.

We define the output gap in the following way as proposed by Ireland (2004). Under the structure of our model, suppose there is a benevolent government that seeks to maximize the representative household’s welfare:

$$E \sum_{t=0}^{\infty} \beta^t \left[ a_t \log Y_t - \frac{1}{\eta} \left( \int_0^1 N_{jt} dj \right)^\eta \right] \quad (18)$$

that is, in each time period  $N_{jt}$  units of labor are allocated to the representative intermediate firm to produce  $Y_{jt}$  units of intermediate good  $j$ , which will then be used as input goods to produce  $Y_t$  units of final goods.

This optimization problem is subject to the following economy-wide constraint:

$$Y_t = Z_t \left( \int_0^1 N_{jt}^{\frac{\theta_t-1}{\theta_t}} dj \right)^{\frac{\theta_t}{\theta_t-1}} \quad (19)$$

---

<sup>4</sup>For detailed discussion of different techniques for computing potential output, see Nelson and Plosser (1972), Hodrick and Prescott (1997), Laxton and Tetlow (1992).

The first order condition implies that the optimal level of output in the final-goods sector is given by<sup>5</sup> :

$$Y_t = a_t^{\frac{1}{\eta}} Z_t \quad (20)$$

The model's output gap  $x_t$  is then defined by dividing the actual output by the optimal level of output:

$$x_t = \left(\frac{1}{a_t}\right)^{\frac{1}{\eta}} \frac{Y_t}{Z_t} \quad (21)$$

### 4.3 Solution of the Model

In equilibrium, markets must clear. A symmetric equilibrium is characterized by the following conditions:  $Y_{jt} = Y_t$ ,  $P_{jt} = P_t$ ,  $h_{jt} = h_t$ , for all  $j \in [0, 1]$  and  $t = 0, 1, 2, \dots$ . In addition, market clearing conditions require  $M_t = M_{t-1} + T_t$ ,  $B_t = B_{t-1} = 0$ .

These market clearing conditions imply that  $Y_t = C_t$ ; households are homogeneous with respect to consumption and bond holdings (Woodford, 1996; Erceg et al., 2000); intermediate-goods firms are identical with respect to price and production decisions, and; money and asset markets are clearing for all  $t = 0, 1, 2, \dots$

---

<sup>5</sup>It is clear that the optimal level of output responds positively to the preference shock  $a_t$  and the technology shock  $Z_t$ .

We then log-linearize the model around its steady-state. The log-linearized model contains two main equations of our NKDSGE model, the expectational IS curve (B.12) and the New Keynesian Phillips curve (B.13): <sup>6</sup>

$$\hat{x}_t = E_t \hat{x}_{t+1} - (\hat{r}_t - E_t \hat{\pi}_{t+1}) + \left(1 - \frac{1}{\eta}\right)(1 - \rho_a) \hat{a}_t \quad ((B.12))$$

$$\hat{\pi} = \beta E_t \hat{\pi}_{t+1} + \psi \hat{x}_t - \hat{\theta}_t / \phi \quad ((B.13))$$

These two main equations (B.12) and (B.13) imply that in a NKDSGE model the presence of nominal rigidities (the cost-push shock  $\hat{\theta}_t / \phi$  here) is a potential source of nontrivial real effects of monetary policy shocks (Gali, 2002). Without the cost-push shock, the monetary authority can simply set the real interest rate equal to its natural rate  $(1 - \frac{1}{\eta})(1 - \rho_a) \hat{a}_t$  in order to stabilize both the inflation rate and the output gap.

To estimate the model, we apply the method proposed by Blanchard-Kahn (1980) to the log-linearized model. Specifically:

$$f_t = A s_t \quad (22)$$

and

$$s_{t+1} = B s_t + C \varepsilon_{t+1} \quad (23)$$

---

<sup>6</sup>Appendix B describes the symmetric equilibrium and the log-linearization of the model.

where

$$f_t = [\hat{g}_t, \hat{\pi}_t, \hat{r}_t]' \quad (24)$$

$$s_t = [\hat{y}_{t-1}, \hat{\pi}_{t-1}, \hat{r}_{t-1}, \hat{x}_{t-1}, \hat{g}_{t-1}, \hat{a}_t, \hat{e}_t, \hat{z}_t, \hat{\varepsilon}_{rt}]' \quad (25)$$

$$\varepsilon_{t+1} = [\varepsilon_{at+1}, \varepsilon_{et+1}, \varepsilon_{zt+1}, \varepsilon_{rt+1}]' \quad (26)$$

The empirical model consisting of (22) and (23) has three observable variables, output growth, inflation, and the nominal short-term interest rate, and two unobservable variables namely the de-trended output and the output gap. The model also consists of four different shocks, the preference shock  $\hat{a}_t$ , the cost-push shock<sup>7</sup>  $\hat{e}_t$ , the technology shock  $\hat{z}_t$ , and the monetary policy shock  $\varepsilon_{rt}$ . All the shocks are assumed to be serially uncorrelated. In other words, the covariance matrix of  $\varepsilon_{t+1}$  is diagonal:

$$E\varepsilon_{t+1}\varepsilon_{t+1}' = \begin{bmatrix} \sigma_a & 0 & 0 & 0 \\ 0 & \sigma_e & 0 & 0 \\ 0 & 0 & \sigma_z & 0 \\ 0 & 0 & 0 & \sigma_r \end{bmatrix} \quad (27)$$

The empirical model is in state-space form and can be estimated via maximum likelihood approach. The model is estimated based on quarterly data on real Gross Domestic Product (GDP), GDP deflator, and 91-day Treasury Bills rate (TBILL) as the nominal short-term interest rate over the period of 1970:1-2000:4.

---

<sup>7</sup> $\hat{e}_t = \hat{\theta}_t/\phi$  is the transformed cost-push cost.



Before calculating the output (GDP) growth, GDP is converted into per-capita form by dividing it with the size of population aged between 15-64. The data for seasonally adjusted real GDP, GDP deflator, and the 91-days TBILL rate are obtained from the South African Reserve Bank Quarterly Bulletin. Note the base year is the year of 2000. Series for population aged between 15-64 is obtained from World Bank database.

#### 4.4 Results

In this section, we compare the out-of-sample forecasting performance of the NKDSGE model with the VARs, both Classical and Bayesian, in terms of the Root Mean Squared Errors (RMSEs). At this stage, a few words need to be said regarding the choice of the evaluation criterion for the out-of-sample forecasts generated from Bayesian models. As Zellner (1986: 494) points out “the optimal Bayesian forecasts will differ depending upon the loss function employed and the form of predictive probability density function”. In other words, Bayesian forecasts are sensitive to the choice of the measure used to evaluate the out-of-sample forecast errors. This fact was also observed in a recent study by Gupta (2006). However, Zellner (1986) points out that the use of the mean of the predictive probability density function for a series, is optimal relative to a squared error loss function and the Mean Squared Error (MSE), and, hence, the RMSE is an appropriate measure to evaluate performance of forecasts, when the mean of the predictive probability density function is used.

But, before we proceed to the discussion of the forecasting performance of the alternative models, it is important to lay out the basic structural differences and advantages of using BVARs over traditional VARs for forecasting.

#### 4.4.1 Classical and Bayesian VARs

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

$$\chi_t = C + \lambda(L)\chi_t + \varepsilon_t \quad (28)$$

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

A crucial drawback of the VAR forecasts is “overfitting” due to the inclusion too many lags and too many variables, some of which may be insignificant. The problem of “overfitting” results in multicollinearity and loss of degrees of freedom, leads to inefficient estimates and large out-of-sample forecasting errors. Thus, it can be argued the performance of VAR forecasts will deteriorate rapidly as the forecasting horizon becomes longer.

A forecaster can overcome this “overfitting” problem by using Bayesian techniques. The motivation for the Bayesian analysis is based on the knowledge that more recent values of a variable are more likely to contain useful information about its future movements than older values. From a Bayesian perspective, the exclusion restriction in the VAR is an inclusion of a coefficient without a prior probability distribution (Litterman, 1986a).

The Bayesian model proposed by Litterman (1981), Doan, *et al.* (1984), and Litterman (1986b), imposes restrictions on those coefficients by assuming they are more likely to be near zero. The restrictions are imposed by specifying normal prior<sup>8</sup> distributions with zero means and small standard deviations for all the coefficients with standard deviation decreasing as lag increases. One exception is that the mean of the first own lag of a variable is set equal to unity to reflect the assumption that own lags account for most of the variation of the given variable. To illustrate the Bayesian technique, suppose the “Minnesota prior” means and variances take the following form:

$$\begin{aligned}\beta_i &\sim N(1, \sigma_{\beta_i}^2) \\ \beta_j &\sim N(0, \sigma_{\beta_j}^2)\end{aligned}\tag{29}$$

where  $\beta_i$  represents the coefficients associated with the lagged dependent variables in each equation of the VAR, while  $\beta_j$  represents coefficients other than  $\beta_i$ . The

---

<sup>8</sup>Note Litterman (1981) uses a diffuse prior for the constant, which is popularly referred to as the “Minnesota prior” due to its development at the University of Minnesota and the Federal Reserve bank at Minneapolis.

prior variances  $\sigma_{\beta_i}^2$  and  $\sigma_{\beta_j}^2$ , specify the uncertainty of the prior means,  $\beta_i = 1$  and  $\beta_j = 0$ , respectively.

Doan *et al.* (1984) propose a formula to generate standard deviations as a function of a small number of *hyperparameters*<sup>9</sup> :  $w, d$ , and a weighting matrix  $f(i, j)$ . This approach allows the forecaster to specify individual prior variances for a large number of coefficients based on only a few *hyperparameters*. The specification of the standard deviation of the distribution of the prior imposed on variable  $j$  in equation  $i$  at lag  $m$ , for all  $i, j$  and  $m$ , defined as  $S(i, j, m)$ :

$$S(i, j, m) = [w \times g(m) \times f(i, j)] \frac{\hat{\sigma}_i}{\hat{\sigma}_j} \quad (30)$$

where:

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

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

The term  $w$  is the measurement of standard deviation on the first own lag, which indicates the overall tightness. A decrease in the value of  $w$  results a tighter prior.

The parameter  $g(m)$  measures the tightness on lag  $m$  relative to lag 1, and is assumed to have a harmonic shape with a decay of  $d$ . An increasing in  $d$ , tightens

---

<sup>9</sup>The name of *hyperparameter* is to distinguish it from the estimated coefficients, the parameters of the model itself.

the prior as lag increases.<sup>10</sup> The parameter  $f(i, j)$  represents the tightness of variable  $j$  in equation  $i$  relative to variable  $i$ . Reducing the interaction parameter  $k_{ij}$  tightens the prior.  $\hat{\sigma}_i$  and  $\hat{\sigma}_j$  are the estimated standard errors of the univariate autoregression for variable  $i$  and  $j$  respectively. In the case of  $i \neq j$ , the standard deviations of the coefficients on lags are not scale invariant (Litterman, 1986b: 30). The ratio,  $\frac{\hat{\sigma}_i}{\hat{\sigma}_j}$  in (30), scales the variables so as to account for differences in the units of magnitudes of the variables.

The BVAR model is estimated using Theil's (1971) mixed estimation technique, which involves supplementing the data with prior information on the distribution of the coefficients. For each restriction imposed on the parameter estimated, the number of observations and degrees of freedom are increased by one in an artificial way. Therefore, the loss of degrees of freedom associated with the unrestricted VAR is not a concern in the BVAR.

#### 4.4.2 Forecast accuracy

Table 13 to 15 report the RMSEs from the NKDSGE model along with the VARs. When compared to the VAR and BVAR, the NKDSGE model does a better job in predicting inflation than it does in predicting output growth and the nominal short-term interest rate (TBILL). To be more precise, for inflation, the NKDSGE model outperforms both the unrestricted VAR and the optimal

---

<sup>10</sup>In this paper, we set the overall tightness parameter ( $w$ ) equal to 0.3, 0.2, and 0.1, and the harmonic lag decay parameter ( $d$ ) equal to 0.5, 1, and 2. These parameter values are chosen so that they are consistent with the ones that used by Liu and Gupta (2007), and Liu *et al.* (2007).

Table 13: **RMSE (2001Q1-2006Q4): Output Growth**

QA	1	2	3	4	AVE
NKDSGE	0.726	0.787	0.888	0.961	0.840
VAR (1)	0.756	0.700	0.797	0.851	0.776
BVAR ( $w=.1, d=1$ )	0.633	0.701	0.797	0.863	<b>0.748</b>

QA: quarter ahead; RMSE: root mean squared error (%).

BVAR<sup>11</sup>, while for output growth and TBILL the RMSEs generated from the NKDSGE model are larger than those generated from the unrestricted VAR and the BVAR.

As far as the forecasting performances of the BVARs are concerned, except for inflation, the optimal BVAR outperforms both the NKDSGE model and the unrestricted VAR. For inflation, the optimal BVAR only outperforms the unrestricted VAR. As shown in Table 13 to 15, for output growth and inflation a BVAR with a relatively tighter prior ( $w = 0.1, d = 1$ ) produces smaller forecast errors, whereas for TBILL the opposite holds. Interestingly, this finding is different from Liu *et al.* (2007), in which a BVAR with a relatively loose prior produces smaller forecast errors. Specifically, Liu *et al.* show that for all four variables forecasted, namely output, consumption, investment and hours worked, a BVAR with the most loose prior ( $w = 0.3, d = 0.5$ ) outperforms the estimated Hansen(1985)-type DSGE model and a Classical VAR.

<sup>11</sup>Here we only report the RMSEs from the optimal BVAR, i.e. a BVAR with a specific set of “hyperparameters” for which we obtain the lowest RMSEs for each quarter.

Table 14: **RMSE (2001Q1-2006Q4): Inflation**

QA	1	2	3	4	AVE
NKDSGE	0.280	0.349	0.422	0.439	<b>0.373</b>
VAR (1)	0.364	0.409	0.462	0.519	0.439
BVAR (w=.1, d=1)	0.312	0.402	0.467	0.520	0.425

QA: quarter ahead; RMSE: root mean squared error (%).

Table 15: **RMSE (2001Q1-2006Q4): TBILL**

QA	1	2	3	4	AVE
NKDSGE	0.914	1.586	2.067	2.406	1.743
VAR (1)	0.813	1.464	1.962	2.334	1.643
BVAR (w=.3, d=.5)	0.688	1.365	1.901	2.306	<b>1.565</b>

QA: quarter ahead; RMSE: root mean squared error (%).

In order to evaluate the models' forecast accuracy, we perform the across-model test between the NKDSGE model and the VAR and BVAR models in pairs. The across-model test is based on the statistic proposed by Diebold and Mariano (1995). The test statistic is defined as the following. For instance, let  $\{e_t^v\}_{t=1}^T$  denote the associated forecast errors from the unrestricted VAR(1) model and  $\{e_t^k\}_{t=1}^T$  denote the forecast errors from the NKDSGE model. The test statistic is then defined as  $s = \frac{l}{\sigma_l}$ , where  $l$  is the sample mean of the "loss differentials" with  $\{l_t\}_{t=1}^T$  obtained by using  $l_t = (e_t^v)^2 - (e_t^k)^2$  for all  $t = 1, 2, 3, \dots, T$ , and where  $\sigma_l$  is the standard error of  $l$ . The  $s$  statistic is asymptotically distributed as a standard normal random variable and can be estimated under the null hypothesis of equal forecast accuracy, i.e.  $l = 0$ . Therefore, in this case, a positive value of  $s$  would suggest that the NKDSGE model outperforms the unrestricted VAR(1) model in terms of out-of-sample forecasting. Results are reported in Table 16. In

Table 16: **Across-Model Test Statistics**

Quarters Ahead	1	2	3	4
(A) Output Growth				
BVAR vs. NKDSGE	-1.573	-1.271	-1.332	-1.257
BVAR vs. VAR(1)	-0.913	3.143*	0.002	1.433
NKDSGE vs. VAR(1)	0.976	1.710	1.310	1.270
(B) Inflation				
BVAR vs. NKDSGE	0.760	0.541	0.358	1.078
BVAR vs. VAR(1)	-1.145	-0.533	0.052	0.747
NKDSGE vs. VAR(1)	-0.889	-0.588	-0.355	-1.019
(C) TBILL				
BVAR vs. NKDSGE	-2.226*	-1.542	-0.896	-0.547
BVAR vs. VAR(1)	-1.377	-1.009	-0.769	-0.576
NKDSGE vs. VAR(1)	2.463*	1.010	0.577	0.371

Note:\* indicates at the 5% level significant.

general, the NKDSGE model does a better job in predicting inflation than it does in predicting output growth and the nominal short-term interest rate (TBILL). The differences between RMSEs generated from the NKDSGE model and the VARs are minor, since most of the test statistics are insignificant.

## 4.5 Conclusion

In this paper, we show that, besides its usual usage for policy analysis, a small-scale NKDSGE model has a future for forecasting. We show that the NKDSGE model outperforms both the Classical and Bayesian variants of the VARs in forecasting inflation, but not for output growth and the nominal short-term interest rate. However, the differences of the forecast errors are minor. The



indicated success of the NKDSGE model for predicting inflation is important, especially in the context of South Africa — an economy targeting inflation.

As suggested by Smets and Wouters (2004), a NKDSGE model estimated by Bayesian techniques can become an useful tool in the forecasting kit for central banks. In this backdrop, further research will concentrate on developing an estimated NKDSGE model based on Bayesian techniques. In addition, future research in this area will aim to extend the current framework into that of a small open economy.

## Optimization

### Household

In the NKDSGE model, the representative household chooses  $\{C_t, h_t, \frac{M_t}{P_t}, \frac{B_t}{P_t}\}$  to maximize the utility function (1):

$$E \sum_{t=0}^{\infty} \beta^t [a_t \log(C_t) + \log(\frac{M_t}{P_t}) - (\frac{1}{\eta}) h_t^\eta], \quad 0 < \beta < 1, \quad \eta \geq 1, \quad ((1))$$

subject to the budget constraint (3):

$$C_t + \frac{B_t}{r_t P_t} = \frac{W_t}{P_t} h_t + \frac{B_{t-1}}{P_t} + D_t + T_t - \frac{M_t - M_{t-1}}{P_t} \quad ((3))$$

The resulting Bellman's equation is as follows:

$$V(M_{t-1}, B_{t-1}, a_t, Z_t, \varepsilon_{rt}) = \max [a_t \log(C_t) + \log(\frac{M_t}{P_t}) - (\frac{1}{\eta}) h_t^\eta + \beta E_t V(M_t, B_t, a_{t+1}, Z_{t+1}, \varepsilon_{rt+1})] \quad (A.1)$$

Substituting  $C_t$  from (3) into (A.1) and solving this problem yields the following first order condition (FOC) for hours worked:

$$\frac{\partial V(M_{t-1}, B_{t-1}, a_t, Z_t, \varepsilon_{rt})}{\partial h_t} = 0 \quad (A.2)$$

$$\frac{a_t}{C_t} \frac{W_t}{P_t} - h_t^{\eta-1} = 0 \quad (A.3)$$

$$\frac{W_t}{P_t} = a_t^{-1} C_t h_t^{\eta-1} \quad (A.4)$$

The FOC for bond holdings is given as follows:

$$\frac{\partial V(M_{t-1}, B_{t-1}, a_t, Z_t, \varepsilon_{rt})}{\partial B_t} = 0 \quad (\text{A.5})$$

$$\frac{a_t}{C_t} \left( -\frac{1}{r_t P_t} \right) + \frac{\partial \beta E_t V(M_t, B_t, a_{t+1}, Z_{t+1}, \varepsilon_{rt+1})}{\partial B_t} = 0 \quad (\text{A.6})$$

The associated envelope condition is:

$$\frac{\partial V(M_{t-1}, B_{t-1}, a_t, Z_t, \varepsilon_{rt})}{\partial B_{t-1}} = \frac{a_t}{C_t} \frac{1}{P_t} \quad (\text{A.7})$$

Updating (A.7) and combining with (A.6) yields

$$\frac{a_t}{C_t} = r_t \beta E_t \left[ a_{t+1} \left( \frac{1}{C_{t+1}} \right) \left( \frac{P_t}{P_{t+1}} \right) \right] \quad (\text{A.8})$$

FOC for money holdings can be derived as follows:

$$\frac{\partial V(M_{t-1}, B_{t-1}, a_t, Z_t, \varepsilon_{rt})}{\partial M_t} = 0 \quad (\text{A.9})$$

$$\frac{a_t}{C_t} (-1) \frac{1}{P_t} + \frac{1}{M_t} + \frac{\partial \beta E_t V(M_t, B_t, a_{t+1}, Z_{t+1}, \varepsilon_{rt+1})}{\partial M_t} = 0 \quad (\text{A.10})$$

The associated envelope condition is:

$$\frac{\partial V(M_{t-1}, B_{t-1}, a_t, Z_t, \varepsilon_{rt})}{\partial M_{t-1}} = \frac{a_t}{C_t} \frac{1}{P_t} \quad (\text{A.11})$$

Updating (A.11) and combining with (A.10), we have:

$$\frac{a_t}{C_t} = \frac{P_t}{M_t} + \beta E_t \left( \frac{a_{t+1}}{C_{t+1}} \frac{P_t}{P_{t+1}} \right) \quad (\text{A.12})$$

$$\frac{P_t}{M_t} = \beta E_t \left( \frac{P_t}{P_{t+1}} \right) - \frac{a_t}{C_t} \quad (\text{A.13})$$

Using (A.8):

$$\frac{M_t}{P_t} = a_t^{-1} C_t [r_t / (r_t - 1)] \quad (\text{A.14})$$

### Final goods firm

A representative firm produces the final good  $Y_t$  using intermediate goods  $Y_{jt}$  according to the CES production function:

$$Y_t = \left( \int_0^1 Y_{jt}^{\frac{\theta_t-1}{\theta_t}} dj \right)^{\frac{\theta_t}{\theta_t-1}} \quad (\text{A.15})$$

The firm maximizes its profit:

$$\max_{Y_{jt}} \left\{ Y_t - \frac{1}{P_t} \int_0^1 P_{jt} Y_{jt} \right\} \quad (\text{A.16})$$

Alternatively, the firm minimizes its expenditure given the production constraint. The Lagrangean for the firm is given by the following expression:

$$L = \int_0^1 P_{jt} Y_{jt} dj - P_t \left[ Y_t - \left( \int_0^1 Y_{jt}^{\frac{\theta_t-1}{\theta_t}} dj \right)^{\frac{\theta_t}{\theta_t-1}} \right] \quad (\text{A.17})$$

Setting  $\frac{\partial L}{\partial Y_{jt}} = 0$ , yields:

$$P_{jt} = P_t \frac{\partial Y_t}{\partial Y_{jt}} \quad (\text{A.18})$$

where:

$$\frac{\partial Y_t}{\partial Y_{jt}} = \frac{\theta_t}{\theta_t - 1} \left( \int_0^1 Y_{jt}^{\frac{\theta_t - 1}{\theta_t}} dj \right)^{\frac{1}{\theta_t - 1}} \left( \frac{\theta_t - 1}{\theta_t} \right) Y_{jt}^{-1/\theta_t} \quad (\text{A.19})$$

$$= \left( \frac{Y_t}{Y_{jt}} \right)^{1/\theta_t} \quad (\text{A.20})$$

Substituting (A.20) into (A.18), yields:

$$Y_{jt} = \left( \frac{P_{jt}}{P_t} \right)^{-\theta_t} Y_t \quad (\text{A.21})$$

Here, given Euler's theorem, profits in this sector must equal to zero in equilibrium:

$$P_t Y_t = \int_0^1 P_{jt} Y_{jt} dj \quad (\text{A.22})$$

Solving for the optimal price of the final-goods  $Y_t$ , yields:

$$P_t = \left( \int_0^1 P_{jt}^{1-\theta_t} dj \right)^{\frac{1}{1-\theta_t}} \quad (\text{A.23})$$

### Intermediate goods firm

A representative firm produces  $Y_{jt}$  according to the following production function:

$$Y_{jt} = Z_t h_{jt} \quad (\text{A.24})$$

given a quadratic cost of price adjustment:

$$\frac{\phi}{2} \left[ \frac{P_{jt}}{\pi P_{jt-1}} \right]^2 Y_t, \quad \phi > 0, \quad \pi > 1, \quad (\text{A.25})$$

The firm maximizes its market value:

$$E \sum_{t=0}^{\infty} \beta^t (a_t / C_t) \left\{ \left[ \frac{P_{jt}}{P_t} \right]^{1-\theta} Y_t - \left[ \frac{P_{jt}}{P_t} \right]^{-\theta} \left( \frac{W_t}{P_t} \right) \left( \frac{Y_t}{Z_t} \right) - \frac{\phi}{2} \left[ \frac{P_{jt}}{\pi P_{jt-1}} \right]^2 Y_t \right\} \quad (\text{A.26})$$

The first order condition for this problem:

$$\begin{aligned}
 (\theta_t - 1) \left( \frac{P_{jt}}{P_t} \right)^{-\theta_t} \left( \frac{Y_t}{P_t} \right) &= \beta \phi E_t \left[ \left( \frac{a_{t+1}}{a_t} \right) \left( \frac{C_t}{C_{t+1}} \right) \left( \frac{P_{jt+1}}{\pi P_{jt}} - 1 \right) \left( \frac{Y_{t+1}}{P_{jt}} \right) \left( \frac{P_{jt+1}}{\pi P_{jt}} \right) \right] \\
 &+ \left[ \theta_t \left( \frac{P_{jt}}{P_t} \right)^{-\theta_t-1} \left( \frac{W_t}{P_t} \right) \left( \frac{Y_t}{Z_t} \right) \left( \frac{1}{P_t} \right) \right] - \left[ \phi \left( \frac{P_{jt}}{\pi P_{jt-1}} - 1 \right) \left( \frac{Y_t}{\pi P_{jt-1}} \right) \right] \quad (\text{A.27})
 \end{aligned}$$

## The Log-linear Equilibrium

### Symmetric Equilibrium

In a symmetric equilibrium, the model can be summarized as follows:

$$Y_t = Z_t h_t \tag{B.1}$$

$$\frac{W_t}{P_t} = a_t^{-1} C_t h_t^{\eta-1} \tag{B.2}$$

$$\frac{a_t}{C_t} = r_t \beta E_t \left[ a_{t+1} \left( \frac{1}{C_{t+1}} \right) \left( \frac{P_t}{P_{t+1}} \right) \right] \tag{B.3}$$

$$\frac{M_t}{P_t} = a_t^{-1} C_t [r_t / (r_t - 1)] \tag{B.4}$$

$$\log(a_t) = \rho_a \log(a_{t-1}) + \varepsilon_{at} \tag{B.5}$$

$$\log \theta_t = (1 - \rho_\theta) \log \theta + \rho_\theta \log \theta_{t-1} + \varepsilon_{\theta t} \tag{B.6}$$

$$\log Z_t = \log \bar{Z} + \log Z_{t-1} + \varepsilon_{zt} \tag{B.7}$$

$$\begin{aligned}
0 = & (1 - \theta_t) \left( \frac{P_{jt}}{P_t} \right)^{-\theta_t} \left( \frac{Y_t}{P_t} \right) + \beta \phi E_t \left[ \left( \frac{a_{t+1}}{a_t} \right) \left( \frac{C_t}{C_{t+1}} \right) \left( \frac{P_{jt+1}}{\pi P_{jt}} - 1 \right) \left( \frac{Y_{t+1}}{P_{jt}} \right) \left( \frac{P_{jt+1}}{\pi P_{jt}} \right) \right] \\
& + \left[ \theta_t \left( \frac{P_{jt}}{P_t} \right)^{-\theta_t-1} \left( \frac{W_t}{P_t} \right) \left( \frac{Y_t}{Z_t} \right) \left( \frac{1}{P_t} \right) \right] - \left[ \phi \left( \frac{P_{jt}}{\pi P_{jt-1}} - 1 \right) \left( \frac{Y_t}{\pi P_{jt-1}} \right) \right] \tag{B.8}
\end{aligned}$$

### Log-linearization

In our complete model, equations (B.1)-(B.8) together with the output gap equation (21) describe the behavior of the endogenous variables  $Y_t, C_t, \pi_t, r_t$ , and  $x_t$ , and the three exogenous shocks  $a_t, \theta_t$ , and  $Z_t$ .  $Y_t, C_t$ , and  $Z_t$  are stochastically detrended so that  $y_t = Y_t/Z_t$ ,  $c_t = C_t/Z_t$ , and  $z_t = Z_t/Z_{t-1}$  are stationary.

In the absence of shocks, the economy converges to a steady-state growth path, in which  $y_t = y$ ,  $c_t = c$ ,  $\pi_t = \pi$ ,  $r_t = r$ ,  $x_t = x$ ,  $g_t = g$ ,  $a_t = a$ ,  $\theta_t = \theta$ , and  $z_t = z$  for all  $t = 0, 1, 2, \dots$ . Therefore, in steady-state we have:

$$y = \left[ a \left( \frac{\theta - 1}{\theta} \right) \right]^{\frac{1}{\eta}} \quad (\text{B.9})$$

$$r = \left( \frac{z}{\beta} \right) \pi \quad (\text{B.10})$$

$$x = \left( \frac{\theta - 1}{\theta} \right)^{\frac{1}{\eta}} \quad (\text{B.11})$$

Using first-order Taylor approximation to rewrite all the equations of the model, we have:

$$\hat{x}_t = E_t \hat{x}_{t+1} - (\hat{r}_t - E_t \hat{\pi}_{t+1}) + \left( 1 - \frac{1}{\eta} \right) (1 - \rho_a) \hat{a}_t \quad (\text{B.12})$$

$$\hat{\pi} = \beta E_t \hat{\pi}_{t+1} + \psi \hat{x}_t - \hat{\theta}_t / \phi, \quad \psi = \eta \left( \frac{\theta - 1}{\phi} \right) \quad (\text{B.13})$$

$$\hat{x}_t = \hat{y}_t - \frac{1}{\eta} \hat{a}_t \quad (\text{B.14})$$

$$\hat{g}_t = \hat{y}_t - \hat{y}_{t-1} + \hat{z}_t \quad (\text{B.15})$$

$$\hat{a}_t = \rho_a \hat{a}_{t-1} + \varepsilon_{at} \quad (\text{B.16})$$

$$\hat{\theta}_t = \rho_{\theta} \hat{\theta}_{t-1} + \varepsilon_{\theta t} \quad (\text{B.17})$$

$$\hat{z}_t = \varepsilon_{zt} \quad (\text{B.18})$$





## Chapter 5

### Conclusions

The thesis is the first attempt in using alternative forms of Dynamic Stochastic General Equilibrium (DSGE) models for forecasting the South African economy. The out-of-sample forecast performances of these alternative forms of DSGE models are evaluated by comparing them with the same generated by the Classical and Bayesian variants of the VARs.

Compared to the VARs and the BVARs, the calibrated Hansen (1985)-type DSGE model produces large out-of-sample forecast errors. The results from the second paper suggest that the estimated hybrid DSGE (DSGE-VAR) model outperforms the Classical VAR, but not the Bayesian VARs. However, it does indicate that the forecast accuracy can be improved alarmingly by using the estimated version of the DSGE model. In the third paper, we show that, besides the usual usage for policy analysis, a small-scale NKDSGE model has a future for forecasting. The NKDSGE model outperforms both the Classical and Bayesian variants

of the VARs in the case of forecasting inflation, but not for output growth and the nominal short-term interest rate. However, the differences of the forecasts errors are minor. The indicated success of the NKDSGE model for predicting inflation is important, especially in the context of South Africa — an economy targeting inflation.

As suggested by Smets and Wouters (2004), a NKDSGE model estimated by Bayesian techniques can become a useful tool in the forecasting kit for central banks. In this backdrop, further research will concentrate on developing an estimated NKDSGE model based on Bayesian techniques. In addition, future research in this area also will aim to extend the current framework into that of a small open economy.



## Bibliography

- Backus, D. K., Patrick, J. and F. E. Kydland, 1995. International business cycles: theory and evidence. In Cooley, T. F. ed., *Frontiers in business cycle research*. Princeton: Princeton University Press.
- Barro, R. J. and D. B. Gordon, 1983. Rules, Discretion, and Reputation in a Model of Monetary Policy. *Journal of Monetary Economics*, 12: 101-121.
- Blanchard, O. J. and C. M. Kahn, 1980. The solution of linear difference models under rational expectations. *Econometrica*, 48: 1305-1311.
- Christiano, L. J. and M. Eichenbaum, 1992. Current real-business-cycle theories and aggregate labor-market fluctuations. *The American Economic Review*, 82(3): 430-450.
- Christiano, L. J., Eichenbaum, M. and C. L. Evans, 2005. Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy*, 113(11): 1-45.
- Clarida, R., Gali, J. and M. Gertler, 1999. The science of monetary policy: A new Keynesian perspective. *Journal of Economic literature*, 37: 1661-1707.
- Cooley, T. F. and S. F. LeRoy, 1985. Atheoretical macroeconometrics: A Critique. *Journal of Monetary Economics*, 16: 283-308.
- Cooley, T. and J. M. Nason, 1995. Output dynamics in real-business-cycle models. *American Economic Review*, 85(3): 492-511.
- Cooley, T. F. and E. C. Prescott, 1995. Economic growth and business cycle. In Cooley, T. F. ed., *Frontiers of business cycle research*. Princeton: Princeton University Press.
- DeJong, D. N., Ingram, B. F. and C. H. Whiteman, 2000a. Keynesian impulses versus Solow residuals: identifying sources of business cycle fluctuations. *Journal of Applied Econometrics*, 15: 311-329.

- , 2000b. A Bayesian approach to dynamic macroeconomics. *Journal of Econometrics*, 98: 203-223.
- Del Negro, M. and F. Schorfheide, 2003. Take Your Model Bowling: Forecasting with General Equilibrium Models. Federal Reserve Bank of Atlanta, Economic Review, 4th Quarter 2003.
- , 2004. Priors from general equilibrium models for VARs. *International Economic Review*, 45(2): 643-673.
- Del Negro, M., Schorfheide, F., Smets, F. and R. Wouters, 2005. On the fit and forecasting performance of New-Keynesian models. European Central Banks, Working Paper, No. 491.
- Diebold, F. X. and R. S. Mariano, 1995. Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13: 253-263.
- Dixit, A. and J. Stiglitz, 1977. Monopolistic competition and optimum product diversity. *American Economic Review*, 67(3): 297-308.
- Doan, T., Litterman, R. B. and C. Smis, 1984. Forecasting and conditional projection using realistic prior distributions. *Econometric Reviews*, 3: 1-144.
- Dotsey, M. and R. G. King, 2001. Pricing production and persistence. National Bureau of Economic Research, Working Paper, No. 8407. Cambridge.
- Dua, P. and S. C. Ray, 1995. A BVAR model for the Connecticut Economy. *Journal of Forecasting*, 14: 167-180.
- Dua, P., Miller, S. M. and D. J., Smyth, 1999. Using leading indicators to forecast U.S. home sales in a Bayesian vector autoregression framework. *Journal of Real Estate Finance and Economics*, 18: 191-205.
- Erceg, C. J., Henderson, D. W. and A. T. Levin, 2000. Optimal monetary policy with staggered wage and price contracts. *Journal of Monetary Economics*, 46(2): 281-313.
- Gali, J., 2002. New perspectives on monetary policy, inflation, and the business cycle. National Bureau of Economic Research, Working Paper, No. 8767. Cambridge.
- Goodfriend, M. and R. G. King, 1997. The new neoclassical synthesis and the role of monetary policy. National Bureau of Economic Research, Macroeconomics Annual, 231-283. Cambridge.
- Gupta, R. and M. M. Sichei, (2006). A BVAR Model for the South African Economy. *South African Journal of Economics*, 74(3): 391-409.

- Gupta, R., 2006. Forecasting the South African economy with VARs and VECMs. *South African Journal of Economics*, 74(4): 611-628.
- Hall, G. J., 1996. overtime, effort, and the propagation of business cycle shocks. *Journal of Monetary Economics*, 38: 139-160.
- Hansen, G. D., 1985. Indivisible labor and the business cycle. *Journal of Monetary Economics*, 16: 309-327.
- , and T. J. Sargent, 1988. Straight time and overtime in equilibrium. *Journal of Monetary Economics*, 21: 281-308.
- , and T. J. Prescott, 1993. Did technology shocks cause the 1990-1991 recession? *The American Economic Review*, 83(2): 280-286.
- Hamilton, J. D., 1994. Time series analysis. 2nd Edition. Princeton: Princeton University Press.
- Hodrick, R. J. and E. C. Prescott, 1997. Postwar U.S. business cycle: an empirical investigation. *Journal of Money, Credit, and Banking*, 29: 1-16.
- Ingram, B. F., Kocherlakota, N. R. and N. E. Savin, 1994. Explaining business cycle: a multiple-shock approach. *Journal of Monetary Economics*, 34: 415-428.
- Ingram, B. F. and C. Whiteman, 1994. Supplanting the ‘Minnesota’ prior : forecasting macroeconomic time series using real business cycle model priors. *Journal of Monetary Economics*, 34(3): 497-510.
- Ireland, P. N., 2001. Money’s role in the monetary business cycle. National Bureau of Economic Research, Working Paper 8115. Cambridge.
- , 2002. Endogenous money or sticky price? National Bureau of Economic Research, Working Paper 9390. Cambridge.
- , 2004. A method for taking models to the data. *Journal of Economic Dynamic & Control*, 28: 1205-1226.
- , 2004. Technology shocks in the new Keynesian model. National Bureau of Economic Research, Working Paper, No. 10309. Cambridge.
- King, R. G., Plosser, C. I. and S. T. Rebelo, 1988. Production, growth and business cycles I. The basic Neoclassical model. *Journal of Monetary Economics*, 21: 195-232.
- Kydland, F. E. and E. C. Prescott, 1982. Time to build and aggregate fluctuations. *Econometrica*, 50(6): 1345-1370.

Laxton, D. and R. Tetlow, 1992. A simple multivariate filter for the measurement of potential output. Bank of Canada, Working Paper, No. 93-7. Ottawa.

Litterman, R. B., 1981. A Bayesian procedure for forecasting with vector autoregressions. Working Paper, Federal Reserve Bank of Minneapolis.

—, 1986a. A statistical approach to economic forecasting. *Journal of Business and Statistics*, 4(1): 1-4.

—, 1986b. Forecasting with Bayesian vector autoregressions—five years of experience. *Journal of Business and Statistics*, 4(1): 25-38.

Liu G. L. and R. Gupta, 2007. A small-skill DSGE model for forecasting the South Africa economy. *South Africa Journal of Economics*, 75(2): 179-193.

Liu G. L., Gupta, R. and E. Schaling, 2007. Forecasting the South African economy: A DSGE-VAR approach. Economic Research Southern Africa, Working Paper, No. 51. Cape Town.

Long, Jr. J. B. and C. Plosser, 1983. Real business cycles. *Journal of Political Economics*, 91: 39-69.

Lucas, R. E. Jr., 1976. Econometric policy evaluation: A critique. Carnegie Rochester Conference Series on Public Policy, 1: 19-46.

McCallum, T. M. and E. Nelson, 1999. An optimizing IS-LM specification for monetary policy and business cycle analysis. *Journal of Money, Credit and Banking*, 31(3): 296-316.

McGrattan, E. R., Rogerson, R., and R. Wright, 1997. An equilibrium model of the business cycle with household production and fiscal policy. *International Economic Review*, 38: 267-290.

Nelson, C. R. and C. Plosser, 1982. Trend and random walks in macro-economic time series. Bank of Canada, Technical Report, No. 50. Ottawa.

Rogerson, R., 1988. Indivisible labor, lotteries and equilibrium. *Journal of Monetary Economics*, 21: 3-16.

Rogoff, K., 1985. The Optimal Degree of Commitment to an Intermediate Monetary Target. *Quarterly Journal of Economics*, 100: 1169-89.

Rotemberg, J. J., 1982. Sticky prices in the United States. *Journal of Political Economy*, 90: 1187-1211.

Rotemberg, J. and M. Woodford, 1995. Dynamic general equilibrium models with imperfectly competitive product markets. In Cooley, T. F. ed., *Frontiers in business cycle research*. Princeton: Princeton University Press.

—, 1997. An optimization-based econometric framework for the evaluation of monetary policy. National Bureau of Economic Research, *Macroeconomics Annual*, 297-346. Cambridge.

Sargent, T. J., 1989. Two models of measurements and the investment accelerator. *Journal of political Economy*, 97: 251-287.

Schorfheide, F., 2000. Loss function-based evaluation of DSGE models. *Journal of Applied Econometrics*, 15: 645-670.

Sims, C. A., 1980. Macroeconomics and reality. *Econometrica*, 48(January): 1-48.

—, Stock, J. H. and M. W. Watson, 1990. Inference in linear time series models with some unit roots. *Econometrica*, 58: 113-144.

Smets, F., 1998. Output gap uncertainty: does it matter for the Taylor rule? Bank for International Settlements, Working Paper, No. 60. Basle.

Smets, F. and R. Wouters, 2003. An estimated dynamic stochastic general equilibrium model of the euro area. *Journal of the European Economic Association*, 1(5): 1123-1175.

—, 2004. Forecasting with a Bayesian DSGE model: an application to the Euro area. European Central Banks, Working Paper, No. 389.

Smit, B. W. and L. R. Burrows, 2002. Estimating potential output and output gaps for the South African economy. Stellenbosch Economic Working Papers, No. 5.

Solow, R. M., 1957. Technical change and the aggregate production function. *The Review of Economics and Statistics*, 39(3): 312-320.

Stock, J. H. and M. W. Watson, 2001. Vector autoregressions. *Journal of Economic Perspectives*, 15(4): 101-105.

Taylor, J. B., 1993. Discretion versus policy rules in practice. *Carnegie-Rochester Conference Series on Public Policy*, 39: 195-214.

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

Uhlig, H., 1995. A toolkit for analyzing nonlinear dynamic stochastic models easily. Discussion Paper 101. Institute for Empirical Macroeconomics, Federal Reserve Bank of Minneapolis.

Watson, M. W., 1993. Measures of fit for calibrated models. *Journal of Political Economy*, 101(6): 1011-1040.

Whitley, J. D., 1994. A course in macroeconomic modelling and forecasting. Hertfordshire: Harvester Wheatsheaf.

Woodford, M., 1996. Control of the public debt: A requirement for price stability? National Bureau of Economic Research, Working Paper, No. 5684. Cambridge.

Zellner, A., 1979. Statistical analysis of econometric models. *Journal of the American Statistical Association*, 74: 628-643.

—, 1986. A tale of forecasting 1001 series: The Bayesian knight strikes again. *International Journal of Forecasting*, 2(4): 494-494.

Zellner, A. and F. Palm, 1974. Time series analysis and simultaneous equation econometric models. *Journal of Econometrics*, 2: 17-54.

Zimmermann, C., 2001. Forecasting with real business cycle models. Center for Research on Economic Fluctuations and Employment, Working paper, No. 131.