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 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 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 Baysian 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 excludes, 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 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.[1]

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)[2] . Hansen and Prescott (1993)

[1] 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.

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

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 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 Baysian 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:

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

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

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

$$Y_t = Z_t K_{t-1}^{\rho} N_t^{1-\rho} \tag{2}$$

where ρ 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:

$$logZ_t = (1-\psi)log\overline{Z} + \psi logZ_{t-1} + \epsilon_t, \qquad \epsilon_t \sim i.i.d.(0, \sigma^2) \tag{3}$$

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

As in a neoclassical growth model, capital stock depreciates at the rate δ, 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

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

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

$$Y_t = C_t + I_t \tag{5}$$

$$Y_t = Z_t K_{t-1}^{\rho} N_t^{1-\rho} \tag{6}$$

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

$$logZ_t = (1-\psi)log\overline{Z} + \psi logZ_{t-1} + \epsilon_t, \qquad \epsilon_t \sim i.i.d.(0, \sigma^2) \tag{8}$$

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 [3] 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 δ 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 (ρ)

[3] The log-linearized equilibrium conditions are presented in Appendix A.

is equal to 0.26[4] , 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:

$$logZ_t - logZ_{t-1} = (logY_t - logY_{t-1}) - (1-\rho)(logN_t - logN_{t-1}) \tag{9}$$

Omitting the capital part of the expression [5] 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:

$$logZ_t = \psi logZ_{t-1} + \epsilon_t, \qquad \epsilon_t \sim i.i.d.(0, \sigma^2) \tag{10}$$

The persistence parameter ψ 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 ϵ_t. These innovations have a standard deviation of 0.0083.

The discount factor β 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 η, is set equal to one. The parameter A, in the utility

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

[5] There is no quarterly capital stock data available.

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

ρ	$1-\rho$	A	$\overline{Z}$	δ	σ_ϵ	β	ψ	η
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[6] (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

[6]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 [7] . 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

[7] 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**

	Baseline Model			RSA data		
Variable	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 [8] . 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

[8] 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)**

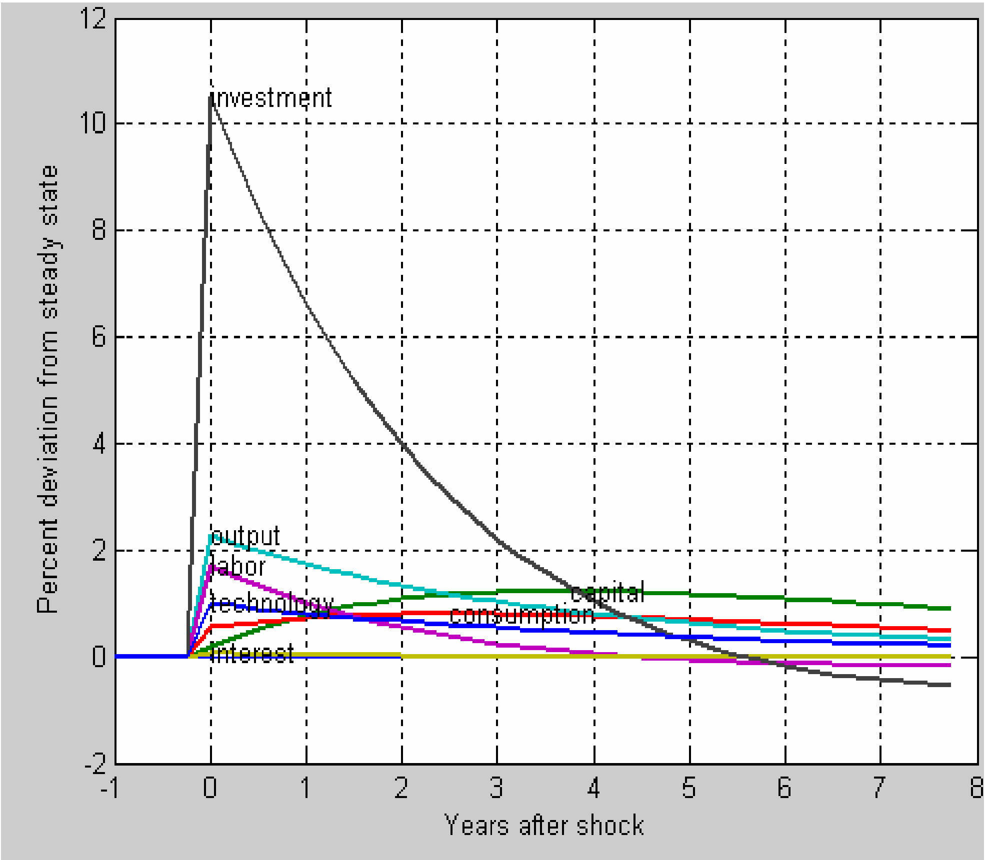
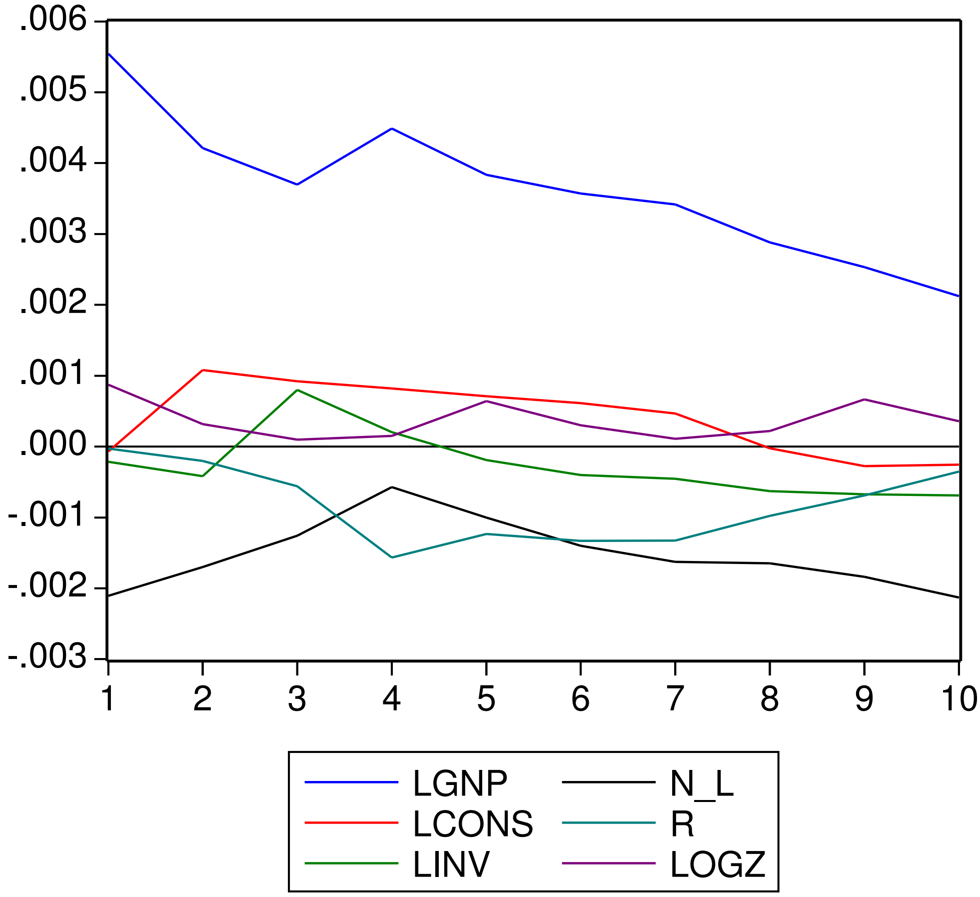


Figure 2: **Impulse responses to technology shock(actual data)**



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)[9]

. 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:

$$\chi_t = C + \lambda(L)\chi_t + \varepsilon_t \tag{11}$$

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

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

[9]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.

$MAPE = [\frac{1}{n}\sum_{t=1}^{n}\frac{|F_t - \hat{F}_t|}{F_t}] \times 100$, 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 Beyesian 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[10] 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:

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

[10]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 β_i represents the coefficients associated with the lagged dependent variables in each equation of the VAR, while β_j represents coefficients other than β_i. The prior variances $\sigma^2_{\beta_i}$ and $\sigma^2_{\beta_j}$, 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*[11] : 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 $S(i,j,m)$:

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

where:

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

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

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,

[11]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[12] 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

[12]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**

			BVARs				
QA	VAR	DSGE	(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	**0.0032**	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**

			BVARs				
QA	VAR	DSGE	(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	**0.0092**	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**

			BVARs				
QA	VAR	DSGE	(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	**0.0423**	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**

			BVARs				
QA	VAR	DSGE	(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	**0.0295**

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

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

			BVARs				
QA	VAR	DSGE	(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	**0.3372**

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, $\overline{X}$ their steady state, and x_t the vector of log-deviations:

$$x_t = logX_t - log\overline{X} \tag{A.1}$$

in other words, x_t denote the percentage deviations from their steady state levels. (A.1) can be written alternatively:

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

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[13] .

The complete model economy:

[13] For details of log-linearization, see Uhlig (1995).

$$Y_t = C_t + I_t \tag{A.3}$$

$$Y_t = Z_t K_{t-1}^{\rho} N_t^{1-\rho} \tag{A.4}$$

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

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

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

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

$$logZ_t = (1-\psi)log\overline{Z} + \psi logZ_{t-1} + \epsilon_t, \qquad \epsilon_t \sim i.i.d.(0,\sigma^2) \tag{A.9}$$

In steady state, we have:

$$\overline{Y} = \overline{C} + \overline{I} \tag{A.10}$$

$$\overline{Y} = \bar{Z}\bar{K}^{\rho}\bar{N}^{1-\rho} \tag{A.11}$$

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

$$\overline{I} = \delta\overline{K} \tag{A.13}$$

$$A = \frac{1}{\overline{N}}(1-\rho)\frac{\overline{Y}}{\overline{C}^{\eta}} \tag{A.14}$$

$$\overline{R} = \frac{1}{\beta} \tag{A.15}$$

The log-linearized equations:

$$\overline{Y}y_t = \overline{C}c_t + \overline{I}i_t \tag{A.16}$$

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

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

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

$$0 = -\eta c_t + y_t - n_t \tag{A.20}$$

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

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

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:

$$0 = Ax_t + Bx_{t-1} + Cy_t + Dz_t \tag{B.1}$$

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

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

The recursive law of motion is derived using Uhlig's MATLAB program:[14]

$$y_t = Px_{t-1} + Qz_t \tag{B.4}$$

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

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

[14]See Uhlig (1995) for details of solving recursive stochastic linear systems with the method of undetermined coefficients.