CENTRE FOR APPLIED MACROECONOMIC ANALYSIS

The Australian National University

OPEN ECONOMY DSGE-VAR FORECASTING AND POLICY ANALYSIS: HEAD TO HEAD WITH THE RBNZ PUBLISHED FORECASTS

Kirdan Lees
Reserve Bank of New Zealand

Troy Matheson
Reserve Bank of New Zealand

Christie Smith
Reserve Bank of New Zealand

Open economy DSGE-VAR forecasting and policy analysis: Head to head with the RBNZ published forecasts

Kirdan Lees, Troy Matheson and Christie Smith

January 2007

Discussion Paper Series

DP2007/01

Open economy DSGE-VAR forecasting and policy analysis:
Head to head with the RBNZ published forecasts[*]

Kirdan Lees, Troy Matheson and Christie Smith[†]

Abstract

We evaluate the performance of an open economy DSGE-VAR model for New Zealand along both forecasting and policy dimensions. We show that forecasts from a DSGE-VAR and a 'vanilla' DSGE model are competitive with, and in some dimensions superior to, the Reserve Bank of New Zealand's official forecasts. We also use the estimated DSGE-VAR structure to identify optimal policy rules that are consistent with the Reserve Bank's Policy Targets Agreement. Optimal policy rules under parameter uncertainty prove to be relatively similar to the certainty case. The optimal policies react aggressively to inflation and contain a large degree of interest rate smoothing, but place a low weight on responding to output or the change in the nominal exchange rate.

[†] Corresponding Address: Economics Department, Reserve Bank of New Zealand, 2 The Terrace, Wellington, 6011, New Zealand. Tel: +64 4 471 3666, Fax: +64 4 473 1209. *email address: leesk@rbnz.govt.nz.*

1 Introduction

We apply the DSGE-VAR methodology developed by Del Negro and Schorfheide (2004) to model a small open economy with an explicit inflation target – New Zealand. We assess the DSGE-VAR's forecasting performance and we use the structural model to investigate optimal policy rules.

Forecasting performance is a key criterion for evaluating model performance within central banks; models are typically required to forecast well before they become an established part of the policy process. Since the Reserve Bank's forecasts are predicated on endogenous policy, the published forecasts of the Reserve Bank of New Zealand provide a unique benchmark against which to compare our DSGE-VAR forecasts. We go head-to-head with these official forecasts using real-time data.

One of the advantages of the DSGE-VAR technology over reduced form statistical models is that the policy-maker learns about the structure of the economy (Del Negro and Schorfheide 2004). We adopt the model estimated by Lubik and Schorfheide (2005), which serves as a minimal set of theory for a small open economy model. This structure enables us to conduct policy experiments: we apply the DSGE-VAR model to identify how effective different policy rules are at achieving the inflation targeting objectives specified in the Reserve Bank's Policy Targets Agreement (PTA).[1]

The rest of the paper proceeds as follows. Section 2 discusses the DSGE-VAR technology and outlines the Del Negro-Schorfheide algorithm we adopt as our estimation procedure. Section 3 outlines the Lubik and Schorfheide model, our parameter estimates, and the impulse responses implied by the model. Section 4 compares out-of-sample forecasts of the DSGE-VAR to other VAR alternatives and to the official forecasts of the Reserve Bank of New Zealand. Section 5 details our policy experiment and concluding comments are made in section 6.

---

[1] The Reserve Bank of New Zealand Act (1989) specifies price stability as the primary objective of the Bank. The PTA, negotiated agreement between the Minister of Finance and the Governor of the Reserve Bank of New Zealand, clarifies the Reserve Bank's objective.

2   DSGE-VARs

Developing plausible empirical models of the macroeconomy has been a focus of policy research ever since national accounts data became available in the 1940s. Considerable progress in this endeavour was made in the 1950s and 1960s under the auspices of the Cowles Commission research programme. The Cowles Commission developed techniques to identify and estimate simultaneous equation models (SEMs). Unfortunately, subsequent theoretical developments called into question the assumptions used to identify SEMs, and their performance as forecasting tools was also subject to criticism.

Empirical research bifurcated following the decline in popularity of the Cowles Commission approach (Heckman 1999). One strand of the literature focused on developing structural DSGE models of the macroeconomy that were tightly defined by theory, while another strand placed emphasis on statistical models, often VARs, that were tightly tied to the data.

While unrestricted VARs largely avoid false parameter restrictions that bias parameter estimates, the parameters in macroeconomic VAR models are often not very precisely estimated. In principle, it is of course highly desirable to allow the data to guide views regarding the data generating process. In practice, however, the data often do not speak very clearly because of comparatively small data samples, collinearity in the lagged data, and a tendency to overfit the data due to the proliferation of parameters.

In response to this latter problem, Bayesian techniques have been applied to VAR models. These Bayesian techniques seek to shrink the parameters towards particular parts of the parameter space. The application of such restrictions helps to sharpen inference, and has led to improved forecasting performance in practice.

The original Bayesian VARs (BVARs) were motivated by statistical prior beliefs regarding the unpredictability of data. The Minnesota prior associated with Doan, Litterman, and Sims (1984) and Litterman (1986), for example, shrinks the parameter values on higher lags towards zero. With the Minnesota prior, the further into the past one goes, the tighter is the prior that the associated coefficients are near zero.[2]

Rather than use statistical priors, Del Negro and Schorfheide (2004) develop an estimation methodology that uses DSGE theory to motivate one's prior beliefs over the VAR parameters. Their methodology thus enables researchers to reunite

---

[2] See chapter 12 of Hamilton (1994) for a more elaborate discussion.

the two literatures that evolved in the wake of the Cowles Commission SEMs. Ideally, this framework yields models that can be used both for forecasting and for policy experiments.

2.1   Estimating a DSGE-VAR

Wold (1938) demonstrated that covariance-stationary processes have an infinite order moving average (MA) representation. If suitable restrictions prevail, infinite order MA processes can be represented using either autoregressive moving average models (ARMAs) or indeed autoregressions (ARs). Multivariate analogues for vector-valued stochastic processes parallel the univariate relationships between MAs, ARMAs, and ARs.

It has long been realised that theoretical models simply imply restricted forms for statistical models such as vector autoregressions or vector autoregressive-moving average (VARMA) models. The correspondence between theoretical and statistical models has prompted interest in using theoretical models as the source of priors for their statistical counterparts. Ingram and Whiteman (1994) show that the prior from an RBC model can help forecast key US macroeconomic variables. DeJong, Ingram, and Whiteman (2000) emphasize that Bayesian methods can be used to learn about the theoretical model.

Working in this vein, Del Negro and Schorfheide (2004) develop an estimation methodology that allows researchers to learn about theoretical models from statistical counterparts.[3] Specifically, Del Negro and Schorfheide (2004) use a small dynamic stochastic general equilibrium (DSGE) model to provide priors for a VAR. The DSGE model incorporates rational, forward-looking agents who maximise their welfare subject to the constraints they face. By confronting the DSGE prior with the VAR, one can obtain a posterior distribution for the parameters of the DSGE model.

Del Negro and Schorfheide's approach can be thought of as generating artificial data using the DSGE model to extend the sample of actual data. The VAR is then applied to this augmented data sample. The number of data observations generated by the DSGE model determines the influence that the DSGE model will have on the VAR. If more data is simulated from the DSGE model, then it will have greater influence on the parameter estimates obtained from the VAR.

---

[3] See Del Negro and Schorfheide (2003) for an overview of the methods, and an application to US data.

As Del Negro and Schorfheide (2004) describe, one can envisage a hierarchical process that begins by generating a prior for the DSGE parameter vector, denoted θ. Conditional on the DSGE prior one forms a prior for the VAR parameters. The prior for the VAR parameters can then be confronted with the data to form a posterior for the parameters of the VAR and the DSGE model.

Suppose that we have the following VAR model:

$$y_t = \Phi_0 + \Phi_1 y_{t-1} + ... + \Phi_p y_{t-p} + u_t \tag{1}$$

where y_t is an $n \times 1$ vector of variables at time t, Φ_i are coefficients for $i = 0, 1, ...p$, and $u_t \sim N(0, \Sigma_u)$. Such a system can be represented more parsimoniously as:

$$Y = X\Phi + U \tag{2}$$

where y_t' is the tth row of Y, $[y_{t-1}' ... y_{t-p}']$ is the tth row of X, $\Phi = [\Phi_0' ... \Phi_p']'$ and u_t' is the tth row of U. Conditional on some initial values, the likelihood function for this sample of data is:

$$\Pr(Y|\Phi, \Sigma_u) \propto |\Sigma_u|^{-T/2} \exp\left(-\frac{1}{2} tr\left[\Sigma_u^{-1}(Y - X\Phi)'(Y - X\Phi) \right] \right) \tag{3}$$

where $tr[.]$ denotes the trace of a matrix.

Suppose that λT artificial observations are generated, and let these artificial observations be denoted with superscript $*$. Del Negro and Schorfheide (2004) show that the likelihood of this artificial sample is:

$$\Pr(Y(\theta)^*|\Phi, \Sigma_u) \propto |\Sigma_u|^{-\lambda T/2} \exp\left(-\frac{1}{2} tr\left[\Sigma_u^{-1}(Y^* - X^*\Phi)'(Y^* - X^*\Phi) \right] \right) \tag{4}$$

The joint likelihood of the sample of actual and artificial data is then:

$$\Pr(Y^*(\theta), Y|\Phi, \Sigma_u) \propto \Pr(Y|\Phi, \Sigma_u)\Pr(Y(\theta)^*|\Phi, \Sigma_u) \tag{5}$$

The usual Bayesian approach is to specify a prior and to update that prior with the likelihood of the data using Bayes' rule to obtain the posterior. Applying such an interpretation to equation (5), one can regard $\Pr(Y^*(\theta)|\Phi, \Sigma_u)$ as representing $\Pr(\Phi, \Sigma_u|\theta)$, ie as a prior for Φ and Σ_u.

Del Negro and Schorfheide (2004) make a slight modification to this probability. $\Pr(\Phi, \Sigma_u|\theta)$ is equated to

$$\Pr(Y^*(\theta)|\Phi, \Sigma_u)\Pr(\Phi, \Sigma_u) \tag{6}$$

For analytical convenience, let $\Pr(\Phi, \Sigma_u) \propto |\Sigma_u|^{-(n+1)/2}$; this is an improper prior for θ and Σ_u. The probability $\Pr(\Phi, \Sigma_u | \theta)$ can then be calculated as:

$$
\begin{aligned}
\Pr(\Phi, \Sigma_u | \theta) = {} & c^{-1}(\theta) |\Sigma_u|^{-\frac{\lambda T + n + 1}{2}} \times \\
& \exp\left(-\frac{1}{2} tr\left[\lambda T \Sigma_u^{-1}(\Gamma_{yy}^*(\theta) - \Phi'\Gamma_{xy}^* - \Gamma_{yx}^*\Phi + \Phi'\Gamma_{xx}^*(\theta)\Phi) \right] \right)
\end{aligned}
\tag{7}
$$

where Γ_{yy}^*, Γ_{xy}^*, Γ_{yx}^*, Γ_{xx}^* are the implied population moments from the DSGE model.

Given two conditions, the above process yields a proper prior for the VAR parameters after suitable normalisation. Conditional on the vector of DSGE parameters (θ), the VAR parameters have a conjugate, inverted-Wishart-Normal prior. That is, the variance covariance matrix Σ_u conditional on θ has an inverted Wishart distribution and Φ conditional on Σ_u and θ has a normal distribution. The conjugate prior reduces the computational burden of the algorithm significantly.

To fully specify this in Bayesian terms, the prior $\Pr(\Phi, \Sigma_u, \theta)$ is formed hierarchically: one forms a prior for the DSGE model and then conditional on that prior one forms a prior view for the VAR parameters. For example,

$$
\Pr(\Phi, \Sigma_u, \theta) = \Pr(\Phi, \Sigma_u | \theta) \Pr(\theta)
\tag{8}
$$

In its entirety, we have

$$
\Pr(Y, \Phi, \Sigma_u, \theta) = \Pr(Y | \Phi, \Sigma_u, \theta) \Pr(\Phi, \Sigma_u | \theta) \Pr(\theta)
\tag{9}
$$

But $\Pr(Y | \Phi, \Sigma_u, \theta)$ is simply $\Pr(Y | \Phi, \Sigma_u)$, and $\Pr(\Phi, \Sigma_u | \theta)$ is as in equation (7), which harks back to (4). The probability of the intersection of data and parameters is of course proportional to the posterior probability of the parameters given the data, that is:

$$
\Pr(\Phi, \Sigma_u, \theta | Y) = \frac{\Pr(Y, \Phi, \Sigma_u, \theta)}{\Pr(Y)}
\tag{10}
$$

Thus, one can maximise the right hand side of (9) to find the parameters that maximise the posterior probability, since $\Pr(Y)$ is a constant.

The posterior distribution of the parameters is explored by using the following factorisation:

$$
\Pr(\Phi, \Sigma_u, \theta | Y) = \Pr(\Phi, \Sigma_u | Y, \theta) \Pr(\theta | Y)
\tag{11}
$$

Conditional on θ, the posterior distributions of Φ and Σ_u are once again conjugate inverted-Wishart-Normal.

Rather than literally simulating the artificial data, the expected moments of the DSGE model are used instead of moments from simulated data to avoid sampling variation. The algorithm used to weight a VAR together with a DSGE model thus rests on appropriately weighting the moments of the two models, rather than on generating simulated data samples.

Thus,

$$\Pr(Y^*(\theta), Y(\theta), \Phi, \Sigma_u) \propto \Pr(Y|\Phi, \Sigma_u)\Pr(Y^*|\Phi, \Sigma_u)\Pr(\Phi, \Sigma_u) \qquad (12)$$

2.2 The Del Negro-Schorfheide algorithm

The exposition thus far implicitly conditions on the choice of the hyper-parameter λ. The hyper-parameter λ is chosen to maximise the marginal data density:

$$\max_{\lambda} Pr_{\lambda}(Y) = \int Pr(Y|\Phi, \Sigma_u)Pr_{\lambda}(\Phi, \Sigma_u|\theta)Pr(\theta)d\theta \qquad (13)$$

As Del Negro and Schorfheide (2004) note, it is conceptually possible to average results over the λ hyper-parameter, but they (and we) instead concentrate on the value of λ that maximises the function. As can be seen in equation (13), the marginal data density reflects both the likelihood and the prior, and the choice of hyper-parameter.

Here we briefly summarise the Del Negro-Schorfheide algorithm used to obtain the DSGE-VAR results.

1. The first step is to specify the prior for the DSGE model parameters. This involves determining the prior distributions of the DSGE parameters and key parameters of those distributions (such as measures of location and dispersion).

2. Once the DSGE prior has been specified, the model needs to be transformed into a state space form, linking the theoretical model to the observation equations. Restrictions on the admissible parameter space for the estimation also need to be specified. Using the csminwel procedure from Chris Sims, one estimates the DSGE parameters with the highest posterior probability. The rational expec-

tations solution from csminwel provides the (DSGE-restricted) reduced form for the rational expectations model.

3. Once the posterior mode is available for the DSGE parameters, the Metropolis-Hastings algorithm can be used to explore the posterior distribution of θ. Since the VAR parameters – conditional on both θ and λ – are conjugate, it is straightforward to determine the posterior distribution of the VAR parameters.[4]

4. The VAR parameters that maximise the posterior distribution are a weighted function of the expected moments from the DSGE model and the moments of the unrestricted VAR. The VAR parameters at the posterior mode are thus readily obtainable from these DSGE and unrestricted VAR moments.

5. Searching over a grid of λ values, one can find the optimal λ value that maximises the marginal data density $\Pr_\lambda(Y)$. This step requires integration of the expression: $\int \Pr(Y|\Phi,\Sigma_u)d(\Phi,\Sigma_u)$. The integral can be approximated using the simulated observations for Φ and Σ_u.

6. Once the optimal value of λ is determined one can examine the properties of the DSGE-VAR model, including the impulse responses, variance decompositions and other summary statistics.

7. The DSGE-VAR model can also be used to forecast future realisations of the variables of interest.

3   The model

We use the model of Lubik and Schorfheide (2005), for which the primary antecedent is Gali and Monacelli (2005). Both papers build on the 'new open economy macroeconomics' (NOEM) literature. The new Keynesian models in the NOEM literature are natural points of reference for policy institutions, since the rigidities in these models mean that there is a substantive stabilisation role for policy. Understanding how policy operates in such models and identifying good policies are natural objectives for central banks. The behaviour of actual and optimal policy in these models has thus been a key focus of papers, such as Gali

---

[4] The joint posterior for θ, Φ, and Σ_u can be estimated by using the Metropolis-Hastings algorithm to simulate a data sample from the posterior of θ, and then for each θ realisation drawing from the conditional distributions for Σ_u and then Φ. See Koop (2003) or Geweke (2005) for introductions to the Metropolis-Hastings algorithm.

and Monacelli (2005), Lubik and Schorfheide (2005), Del Negro and Schorfheide (2004), Benigno (2004), and many others.

In many ways, Lubik and Schorfheide (2005) can be regarded as a minimal set of theory for modelling an inflation targeting open economy. The model sacrifices complexity in the interests of tractability.

The model has a continuum of countries with a continuum of firms producing differentiated goods. Each firm operates in a monopolistically competitive environment. Firms set prices according to Calvo staggered pricing. The production function is linear in labour, and abstracts from capital accumulation entirely. Technology is assumed to follow a unit root process and is common to both the domestic and world economies.

Consumers have constant relative risk aversion preferences and they aggregate consumption goods using Dixit-Stiglitz aggregation. Consumers also have a preference for home-produced goods, or at least this is an estimable quantity.

Monetary policy is specified by a flexible Taylor rule, with the lagged interest rate, inflation, output, and the change in the exchange rate as arguments in the policy rule.

International financial markets are assumed to be perfect, enabling risk-sharing between domestic and foreign consumers. World output reflects production in both the international and domestic economies. The exchange rate is introduced into the model via purchasing power parity (PPP). Terms of trade effects also have an effect on output. The model treats the terms of trade, world output, and world inflation as exogenous AR(1) processes.

The linearised version of the model has a forward-looking IS curve (reflecting consumers' intertemporal optimisation) and a Phillips curve governing inflation behaviour. The latter relates inflation to a notion of marginal cost.

The linearised equations are provided below.

$$\tilde{y}_t = E_t\tilde{y}_{t+1} - \chi(\tilde{R}_t - E_t\tilde{\pi}_{t+1}) - \rho_z\tilde{z}_t$$
$$- \alpha\chi E_t\Delta\tilde{q}_{t+1} + \alpha(2-\alpha)\frac{1-\tau}{\tau}E_t\Delta\tilde{y}_{t+1}^* \tag{14}$$

$$\tilde{\pi}_t = \beta E_t\tilde{\pi}_{t+1} + \alpha\beta E_t\Delta\tilde{q}_{t+1} - \alpha\Delta\tilde{q}_t + \frac{\kappa}{\chi}(\tilde{y}_t - \tilde{\bar{y}}_t) \tag{15}$$

$$\tilde{R}_t = \rho_R\tilde{R}_{t-1} + (1-\rho_R)[\psi_\pi\tilde{\pi}_t + \psi_y\tilde{y}_t + \psi_{\Delta e}\Delta\tilde{e}_t] + \varepsilon_t^R \tag{16}$$

$$A_t = A_{t-1} + \varepsilon_{z,t} \tag{17}$$

$$\Delta\tilde{q}_t = \rho_q\Delta\tilde{q}_{t-1} + \varepsilon_{q,t} \tag{18}$$

$$\tilde{y}_t^* = \rho_{y^*}\tilde{y}_{t-1}^* + \varepsilon_{y_t^*} \tag{19}$$

$$\tilde{\pi}_t^* = \rho_{\pi^*}\tilde{\pi}_{t-1}^* + \varepsilon_{\pi_t^*} \tag{20}$$

$$\Delta\tilde{e}_t = \tilde{\pi}_t - (1-\alpha)\Delta\tilde{q}_t - \tilde{\pi}_t^* \tag{21}$$

where $\chi = [\tau + \alpha(2-\alpha)(1-\tau)]$; $\tilde{\bar{y}}_t = -\alpha(2-\alpha)\frac{1-\tau}{\tau}\tilde{y}_t^*$; and $z_t = \ln A_t - \ln A_{t-1}$. Output is denoted y_t; inflation π_t; the nominal interest rate R_t; technological growth z_t; potential output in the absence of nominal rigidities is $\bar{y}_t$; e_t is the nominal exchange rate; and q_t is the terms of trade. Tildes denote deviations from steady state values and asterisks denote foreign variables.

The policy parameters ψ_π, ψ_y, $\psi_{\Delta e}$, and ρ_R indicate the strength of the response to inflation, deviations of output from steady-state, the change in the nominal exchange rate and the lag of the interest rate respectively. α is the import share of domestic consumption, β is the discount factor, τ is the intertemporal elasticisty of substitution and κ gives the output slope in the Phillips curve. The coefficients ρ_q, ρ_z, ρ_{y^*}, ρ_{π^*}, drive the AR(1) processes for the terms of trade, technology, foreign output and foreign inflation respectively. The magnitudes of the shocks are parameterized by σ_R, σ_q, σ_z, σ_{y^*}, and σ_{π^*}, the standard deviations of the shocks to the interest rate, terms of trade, technology, foreign output, and foreign inflation respectively.

The model is specified in terms of stationary variables to enable the Kalman filter to be used to estimate the state space form of the model. Since technology is the integrated process that drives the trending behaviour of series such as output and consumption, the model is made stationary by taking the ratio of the key variables to the level of technology.

The observed variables are output growth, annualised inflation, interest rates, the change in the terms of trade and the change in the exchange rate (defined according to the US convention that an appreciation of the domestic currency corre-

sponds to a decline in magnitude). The foreign variables and the level of technology are not observed directly but are inferred using the Kalman filter.

3.1 Data

We use New Zealand quarterly data for real output growth, inflation, the nominal interest rate, exchange rate changes, and terms of trade changes. The sample is from 1990Q1 to 2005Q4. Real output growth is computed as the log first difference in seasonally-adjusted (production) real gross domestic product and is scaled by 100 to convert it into percentage changes. Inflation is defined as the log first difference in the consumers price index and is scaled by 400 to convert it into annualised percentage changes. The nominal interest rate is the level of the 90 day bank bill yield. Exchange rate changes are defined to be 100 times the log first differences in the trade-weighted nominal exchange rate index (TWI), but inverted so that an increase reflects a depreciation. The terms of trade is 100 times the log first difference of the merchandise terms of trade (export prices over import prices).

3.2 Estimated model

We iterate over a grid that contains the values of λ=0.6, 0.8, 1, 1.2, 1.4, 1.6, 1.8, 2.0, 2.5, 3.0, 5.0, 101 and find the optimal λ is 1.0, implying a weight of 50 percent on the DSGE model and 50 percent on the VAR. Figure 2 below shows that the likelihood is relatively flat in a region near the optimal lambda.

Before presenting draws from the posterior, we test that the Markov chain Monte Carlo has converged. The appendix presents results that indicate convergence based on three sets of tests for convergence in the literature. We draw 4,000,000 draws from the posterior with $\lambda = 1.0$, burn the first 3,500,000 and test whether the parameters have converged. We also thin the post-burn set of parameters, retaining every 20th draw to leave 25,000 draws from the posterior.[5] The parameter estimates based on the simulated posterior are presented in table 1 below.

The policy parameters are presented in the first four rows of the table. The data shifts the response to inflation to 3.719. To compare this with traditional empirical Taylor rule coefficients one needs to multiply $1 - \rho_R$ by ψ_π, which yields 2.29.

---

[5] We work with a thinned posterior simply for computational reasons.

Figure 1
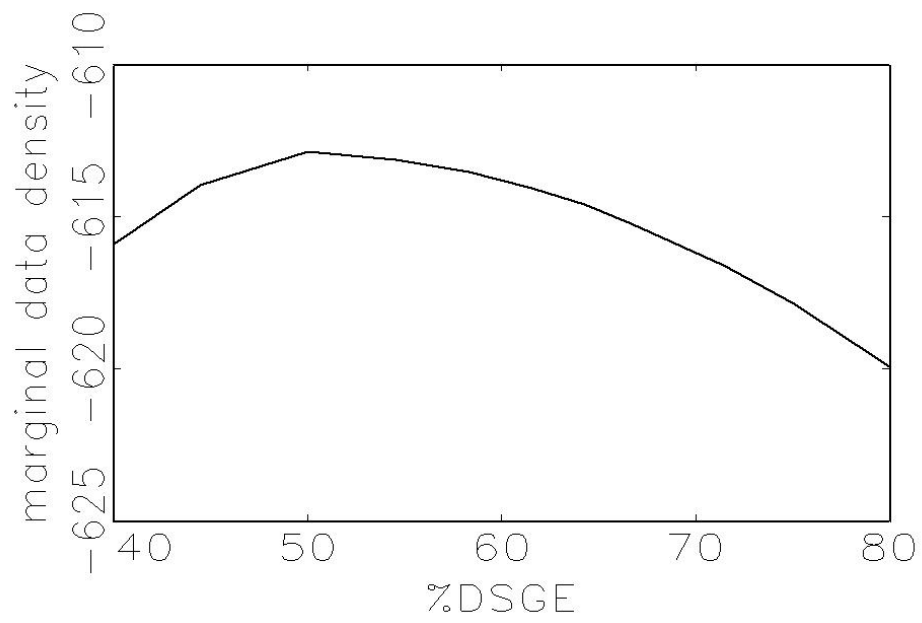**Marginal data density as a function of the λ hyperparameters**

Table 1
Prior and posterior distributions: New Zealand 1990Q1 – 2005Q4

Para.	Dist.	Prior				Posterior			
		Mean	Stdev	ci(L)	ci(H)	Mean	Stdev	ci(L)	ci(H)
ψ_π	$\mathfrak{G}$	1.500	0.5	0.784	2.396	3.719	0.597	2.792	4.729
ψ_y	$\mathfrak{G}$	0.250	0.125	0.087	0.485	0.189	0.074	0.081	0.323
$\psi_{\Delta e}$	$\mathfrak{G}$	0.250	0.125	0.087	0.485	0.075	0.037	0.027	0.138
ρ_R	$\mathfrak{B}$	0.500	0.2	0.168	0.824	0.383	0.121	0.164	0.573
α	$\mathfrak{B}$	0.300	0.05	0.221	0.386	0.211	0.036	0.157	0.274
ρ^*	$\mathfrak{G}$	0.500	0.25	0.171	0.968	0.670	0.322	0.231	1.262
κ	$\mathfrak{G}$	0.500	0.25	0.171	0.968	1.578	0.450	0.881	2.370
τ	$\mathfrak{B}$	0.500	0.2	0.178	0.826	0.531	0.088	0.387	0.679
ρ_q	$\mathfrak{B}$	0.500	0.2	0.175	0.825	0.203	0.099	0.064	0.386
ρ_a	$\mathfrak{B}$	0.200	0.1	0.061	0.382	0.312	0.059	0.220	0.413
ρ_{y^*}	$\mathfrak{B}$	0.900	0.05	0.806	0.968	0.907	0.039	0.840	0.965
ρ_{π^*}	$\mathfrak{B}$	0.800	0.1	0.615	0.940	0.619	0.108	0.441	0.801
$\tilde{\pi}$	$\mathfrak{G}$	2.000	0.5	1.252	2.893	2.144	0.301	1.638	2.645
$\tilde{\gamma}$	$\mathfrak{N}$	0.500	0.2	0.172	0.827	0.781	0.096	0.616	0.936
σ_q	$\mathfrak{G}^{-1}$	2.000	4	1.278	4.757	0.857	0.130	0.680	1.100
σ_z	$\mathfrak{G}^{-1}$	2.000	4	1.286	4.620	1.647	0.206	1.346	2.019
σ_r	$\mathfrak{G}^{-1}$	2.000	4	1.308	4.801	0.838	0.105	0.685	1.019
σ_{y^*}	$\mathfrak{G}^{-1}$	2.000	4	1.301	4.743	2.339	0.868	1.323	3.814
σ_{π^*}	$\mathfrak{G}^{-1}$	2.000	4	1.285	4.796	2.376	0.302	1.904	2.900

$\mathfrak{B}$, denotes the Beta distribution, $\mathfrak{G}$ denotes the Gamma distribution, and $\mathfrak{G}^{-1}$ denotes the Inverse Gamma distribution. Stdev denotes standard deviation. The Inverse Gamma priors are of the form $p(\sigma|v,s) \propto \sigma^{-v-1}e^{-s/2\sigma^2}$. For the Inverse Gamma priors, we report the parameters s and v. The priors are truncated at the boundary of the indeterminacy region of the parameter space.

This estimate is somewhat higher than the standard Taylor coefficient of 1.5, but is fairly comparable to the estimate in Justiniano and Preston (2006). The response to the deviation of output from steady-state, ψ_y falls slightly to 0.189, somewhat below the parameter estimate of $\psi_y = 0.25$ that Lubik and Schorfheide (2005) find over a longer sample period, but close to the coefficient estimated by Justiniano and Preston (2006), though Justiniano and Preston include output growth as an additional argument in their policy rule. The posterior draws suggest a small coefficient on the response to the change in the nominal exchange rate, consistent with Lubik and Schorfheide's conclusion that central banks do not respond to the change in the nominal exchange rate.

In addition, the data shrinks the dispersion on the interest rate smoothing parameter returning a coefficient of 0.383, lower than the corresponding value 0.63 in Lubik and Schorfheide (2005) but similar to the coefficient 0.364 reported in Santacreu (2005). In general, the literature suggests much lower interest rate smoothing in New Zealand than in other countries, though Justiniano and Preston (2006) are a NZ exception. For the United States Lubik and Schorfheide (2004) report a parameter of 0.84 in post-1982 data, close to the parameter of 0.91 that Dennis (2005) reports. In the euro area Adolfson, Lindé, and Villani (2005) report 0.874 for the period 1994Q1 to 2002Q4, while Lubik and Schorfheide (2005) report coefficients of 0.76 and 0.74 for Australia and the United Kingdom respectively.

With regard to the structural parameters, the import share α falls from the prior 0.3 to the posterior 0.211. In contrast, the coefficient κ on deviations of output from steady state increases markedly, although the confidence interval shows wide dispersion. The intertemporal elasticity of substitution τ is relatively tightly estimated with a mean of 0.531; much higher than the 0.31 Lubik and Schorfheide (2005) report for Canadian data, for example. Finally, foreign output appears particularly persistent with 95 percent of the draws from the posterior falling above 0.840. The data return a relatively tightly defined steady-state annualised inflation rate of 2.144 and steady-state quarter on quarter growth of 0.781. The steady-state real interest rate $\tilde{r}$ is $4 \times \tilde{\gamma} + \rho^* = 4 \times 0.781 + 0.670 = 3.794$.

To provide a cross-check on the system estimation methods and on the importance of the priors in generating posterior results, Fukač, Pagan, and Pavlov (2006) suggest using single equation methods to estimate model parameters, when such methods are consistent with the assumptions of the model. In the Lubik and Schorfheide model, Fukač, Pagan, and Pavlov note that it is possible to estimate α, ρ_{π^*}, and ρ_q in single equations.

As for the UK results presented by Fukač *et al*, our single equation estimate sug-

gests that the growth rate of the NZ terms of trade is actually slightly negatively correlated. This result contrasts with our Bayesian posterior, which implies that this correlation is positive (reflecting the beta prior which places zero probability on negative values). Somewhat similarly, the persistence in foreign inflation from a single equation regression is only about 0.29, so it is apparent that the prior mean of 0.8 used in systems estimation has a fairly sizeable effect on the posterior mode estimate 0.619. Lastly, single equation methods imply that $\alpha = 0.92$, much higher than the posterior estimate from the Bayesian estimation. These single equation estimates suggest to us that the relationship between domestic and foreign prices in Lubik and Schorfheide's model should be revisited in future work, to improve the match to the data.

Figure 2 shows the match of the model to key selected moments in the data. The first column of figure 2 shows that the model matches the data mean of the growth rate of output, inflation, the change in the exchange rate, and the change in the terms of trade. However, the model understates the standard deviation of the nominal interest rate. Nominal interest rates were double digit for the first six quarters of the data sample, in part because inflation was between $4-5$ percent but the neutral real rate for New Zealand may also have been higher at this point in time (Basdevant, Björksten, and Karagedikli 2004). Subsequent interest rates are lower: the ninety interest rate averaged 7.5 percent in March 2006 with several analysts suggesting rates had reached the top of the cycle. The posterior appears – appropriately – to have difficulty matching the volatility of interest rates that arises from the earliest quarters in the sample period.

The second column of figure 2 shows the distributions of the population standard deviations of the model variables relative to their respective sample counterparts. The model concentrates most of the mass of the distribution of the standard deviation of output growth between 0.95 and 1.2, slightly higher than the standard deviation of output growth in the data. The model matches the standard deviation of inflation but underestimates the standard deviations for interest rates, the change in the exchange rate, and the terms of trade.

The distribution of the first order autocorrelation statistics for the model is displayed relative to the sample autocorrelations in the third column of the table. The model matches the data autocorrelations, with the exception of the nominal interest rate. This is clearly associated with the surprisingly low degree of interest rate smoothing (0.383) in the estimated policy rule. Fukač and Pagan (2006) compare correlations in UK data to correlations implied by 20,000 draws from the posterior of parameter estimates of the Lubik and Schorfheide (2005) model.

Figure 2
Matching selected moments: New Zealand 1990Q1 - 2005Q4
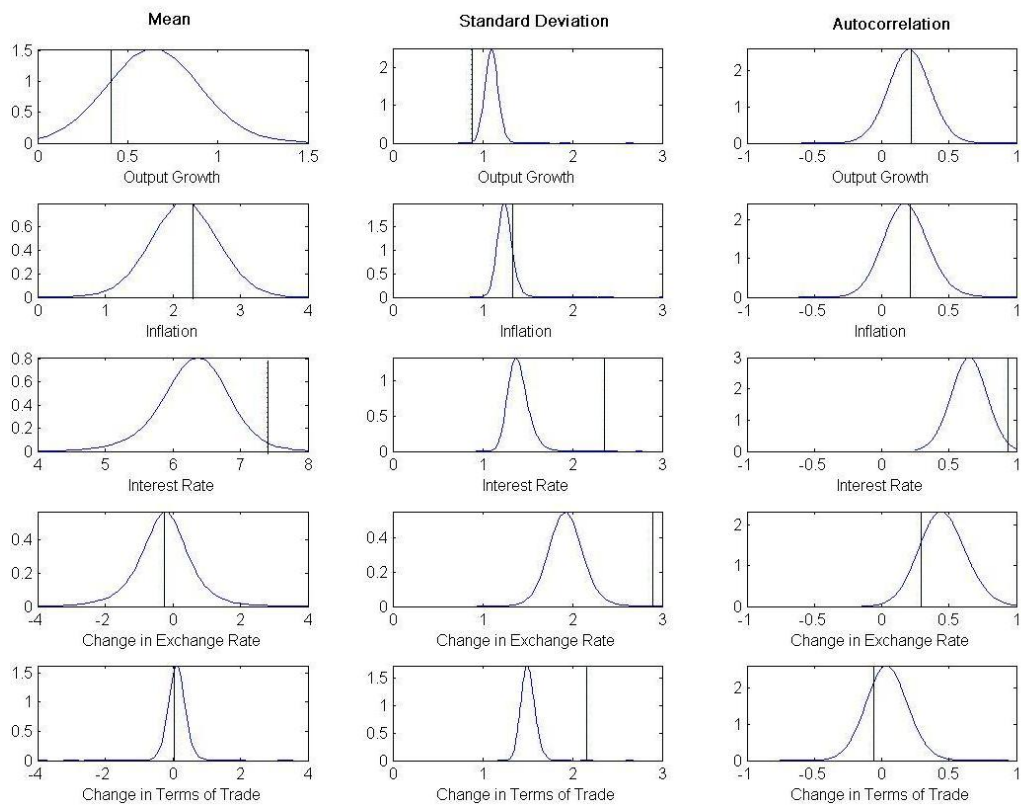
The distributions are the implied moment distributions from the DSGE-VAR; the vertical lines represent the moments in the data.

In marked contrast to our results, they report too little autocorrelation in output growth and too much correlation in exchange rates.

3.3 Identifying impulses from structural shocks

Equation (1) can be thought of as a reduced form of the 'structural' VAR:

$$\begin{align}
B_0 y_t &= B_c + B_1 y_{t-1} + \ldots + B_p y_{t-p} + \varepsilon_t \tag{22}\\
\Rightarrow y_t &= \Phi_0 + \Phi_1 y_{t-1} + \ldots + \Phi_p y_{t-p} + u_t
\end{align}$$

where $\Phi_i = B_0^{-1} B_i$ for $i = c, 1, \ldots p$ and $u_t = B_0^{-1} \varepsilon_t$.

If the elements in the vector of structural shocks ε_t are orthogonal to each other, one can think of identifying the individual effects of each structural shock. If one knows B_0 (and hence its inverse assuming B_0 is non-singular), then it is straightforward to establish the impact of structural shocks. However, B_0 or B_0^{-1} cannot simply be backed out from the reduced form regression; additional restrictions need to be imposed to identify B_0.

In principle, the QR decomposition can be applied to the unknown B_0^{-1}. This implies there exists an orthonormal basis Ω and a lower triangular matrix Σ_{tr} such that $\Sigma_{tr} \Omega = B_0^{-1}$. The contemporaneous effect of a structural shock ε_t on y_t is

$$\left(\frac{\partial y_t}{\partial \varepsilon_t'} \right) = \Sigma_{tr} \Omega = B_0^{-1} \tag{23}$$

Σ_{tr} can be identified from the variance-covariance matrix of the reduced form errors using the Cholesky decomposition, since $\Sigma_{tr} \Sigma_{tr}' = \Sigma_u$, where Σ_u is the variance-covariance matrix of u_t. However, to identify the impact of the structural shocks one also needs to know Ω. Unfortunately, the likelihood function (or the variance-covariance matrix of the reduced form residuals) cannot by itself identify Ω, since $\Omega \Omega' = I_n$, where I_n is the identity matrix. (The orthonormality of Ω implies that $\Sigma_{tr} \Omega \Omega' \Sigma_{tr}' = \Sigma_u$.) Consequently, additional prior or external information must be applied to the model to identify Ω.

Following Del Negro and Schorfheide (2004), it is natural to use the theoretical DSGE model to provide the prior information that enables the identification of Ω. Equivalently, one could think of the DSGE model as providing information about the matrix B_0. In essence, Del Negro and Schorfheide (2004) replace Ω with $\Omega(\theta^*)$ from the DSGE model.

For a given θ, the posterior and prior distribution will be the same (since the reduced form variance-covariance matrix Σ_u which determines Σ_{tr} is invariant with respect to θ). However, θ is updated in light of the data, and so the impulse responses will also change *a posteriori*. The degree to which the impulse responses will evolve depends on the weight that is being applied to the data in the DSGE-VAR.

3.4 The estimated structural IRFs

Figure 3 shows the impulse response functions from $\hat{\lambda}$ estimated with data up to 2005Q3. The plot shows the DSGE impulse responses (solid lines) and the DSGE-VAR impulse responses (dashed lines), along with the corresponding 90 percent confidence bands. Notice that the sign and magnitude of the DSGE and DSGE-VAR impulse responses are quite similar. However, along some dimensions, such as the impact of technology shocks on inflation, there is substantial uncertainty both about the initial impact of the shock and about how it propagates through the system.

Nevertheless, the model dynamics can be broadly described using the estimated impulse responses. A contraction in monetary policy initially reduces output growth and appreciates the exchange rate, lowering inflation. A terms of trade shock lowers inflation and increases output via an appreciation of the currency. Since technology is assumed to be difference stationary, productivity shocks increase output permanently. This leads to a fall in inflation and an easing in monetary policy. While there is an appreciation of the exchange rate predicted by the model, this is subject to much uncertainty according to the impulse responses from the DSGE-VAR. A shock to foreign output leads to a fall in domestic potential output. The subsequent excess demand is met by rising inflation and a contraction in monetary policy. Again, the overall impact on the exchange rate is quite uncertain according to the impulse responses from the DSGE-VAR.

4 Evaluating forecasting performance

Forecast accuracy is typically viewed as a metric to assess both the credibility of a model and the credibility of the policy-makers who use it. Forecasting the macro-economy as accurately as possible is an important policy task, since it helps to explain and justify current policy actions.

Figure 3
Impulse responses, DSGE and DSGE-VAR (λ)
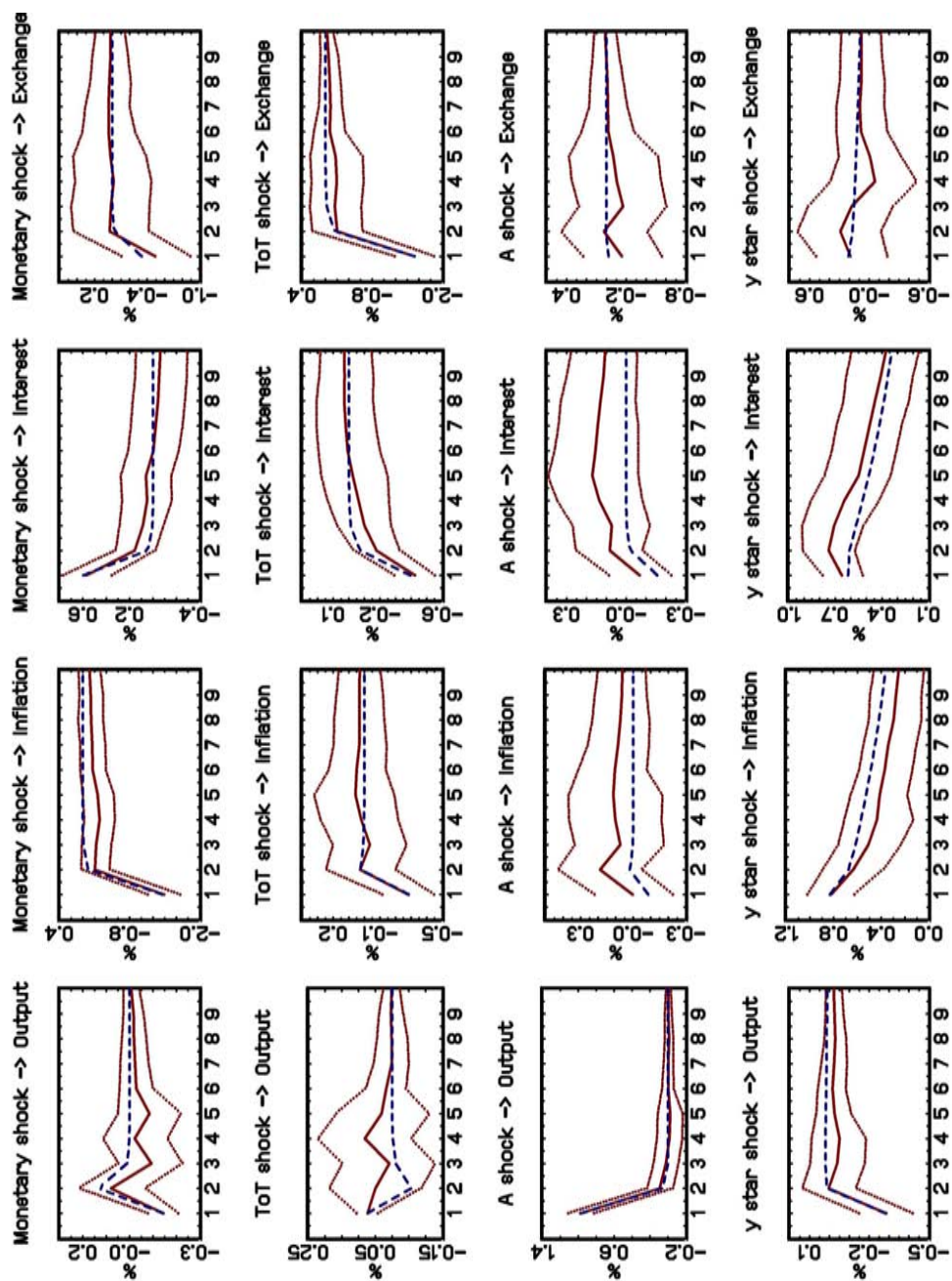
NB DSGE posterior mean (solid), DSGE-VAR posterior mean (dashed) and 90 percent probability bands (dotted).

Following Ingram and Whiteman (1994) and Del Negro and Schorfheide (2004), we test whether forecasts from the DSGE-VAR are competitive with forecasts from an unrestricted VAR and a VAR with a Minnesota prior (which shrinks the VAR coefficients towards a random walk).[6] However, we provide two extensions to the previous literature by comparing the forecasting performance of the DSGE-VAR with the real-time published forecasts of a central bank, and by considering the DSGE-VAR in the context of a small open economy. The central bank forecasts are the forecasts in the Reserve Bank's quarterly *Monetary Policy Statements* (MPS).

To simulate the forecasting performance of our models, we de-mean all data and estimate all equations recursively for 20 quarters from 1998Q4 to 2003Q3. The out-of-sample forecasting performance of the models is then evaluated at horizons *h* of 1 to 8 quarters ahead using ex-post data. The forecast errors are computed for the variables and are cumulated from quarters 1 to *h*, except interest rates which are forecast errors for the levels.[7]

Our forecasting experiment uses the data that were actually available in real time, making the forecasts directly comparable. Effectively, we have one data set for each quarter of our out-of-sample period, where each data set has one more observation of revised historical data compared with the previous data set. While financial data (exchange rates and interest rates) are typically available every minute in real time, most other data are only available at a monthly or quarterly frequency and are published with a (sometimes substantial) lag. Thus, in order to estimate a model on a symmetric data set with, say, T observations on each of the variables, the overall size of the data set that can be used in estimation is limited by the arrival of the least timely series. It may be the case that $T+1$ observations are available for some series when the Tth observation arrives for the least timely series, so that the most up-to-date information must be discarded to achieve the same number of observations for all of the series.

The Reserve Bank forecasts each of the inputs in its macroeconomic model, the *Forecasting and Policy System* (FPS), to fill gaps caused by publication lags. For

---

[6] The Minnesota prior is implemented in the same way as in Del Negro and Schorfheide (2004), where the prior mean for the first lag of log GDP, log CPI, the log exchange rate and the log terms of trade is one (implying that the prior mean for the growth rates of these variables is zero). The prior mean for the level of the interest rate is one. There is a hyperparameter ι (analogous to λ) that controls the weight of the Minnesota prior. The hyperparameter is chosen to maximise the log data density ex-ante, using a modification of the procedure used to determine λ.

[7] The results are qualitatively similar when the errors on the growth rates are not cumulated.

example, when the MPS forecasts are finalised each quarter (period T), the Reserve Bank has all but one month of financial data (two thirds of period T), the CPI observation for the previous quarter (period $T-1$), and GDP data from the quarter before that (period $T-2$). To balance the data set, sectoral experts use indicator models and judgement to forecast the key macroeconomic series up to period T, which serves as the start point for FPS forecasts.[8]

Rather than truncate our data sets in real time we use the real-time Reserve Bank forecasts to fill in the gaps in our real-time data sets caused by publication lags. We use exactly the same data as the Reserve Bank used in real time up to period T. In this way, we ensure that our models have the same information to forecast at horizons beyond period T and make the forecasting models conditional on the same information.

We begin by documenting how the performance of the DSGE-VAR changes with λ. Figure 4 displays the percentage improvement (or loss, if negative) in mean squared forecast error (MSFE) from the DSGE-VAR relative to an unrestricted VAR.

The panel in the bottom right of the figure marked 'Multivariate' shows the percentage gain in the multivariate log-determinant statistic of the DSGE-VAR over the unrestricted VAR.[9]

The grid for λ ranges from 0 to ∞, where $\lambda = \infty$ means that there is a weight of 100 percent on the data simulated from the DSGE and $\lambda = 0$ means that there is no weight on data simulated from the DSGE (the model is an unrestricted VAR).

With the exception of some mixed results for inflation and the exchange rate, the DSGE-VAR produces forecasting gains over the unrestricted VAR. These gains are reflected in a positive multivariate statistic for all horizons and λ's considered. Interestingly, the results for the multivariate statistic are similar to Del Negro and Schorfheide 2004's results for the United States: the relative statistic has an inverted U-shape as a function of the weight on the DSGE model. Likewise, Del

---

[8] The near-term forecasts that are used to complete the data set will not always be correct, meaning that there will be 'starting point' errors at the beginning of the forecast period. However, the Reserve Bank has determined that the cost of these starting point measurement errors is outweighed by the informational advantage that can be gained by using all available information to forecast.

[9] The log-determinant statistic is defined to be the negative of the natural logarithm of the determinant of the forecast error variance matrix, divided by 2 times the number of variables. The gain in this statistic can be thought of as the average gain over all variables being forecast, after accounting for the cross-correlation in the errors.

Figure 4
Percentage gain (loss) in MSFE for the DSGE-VAR over an unrestricted VAR
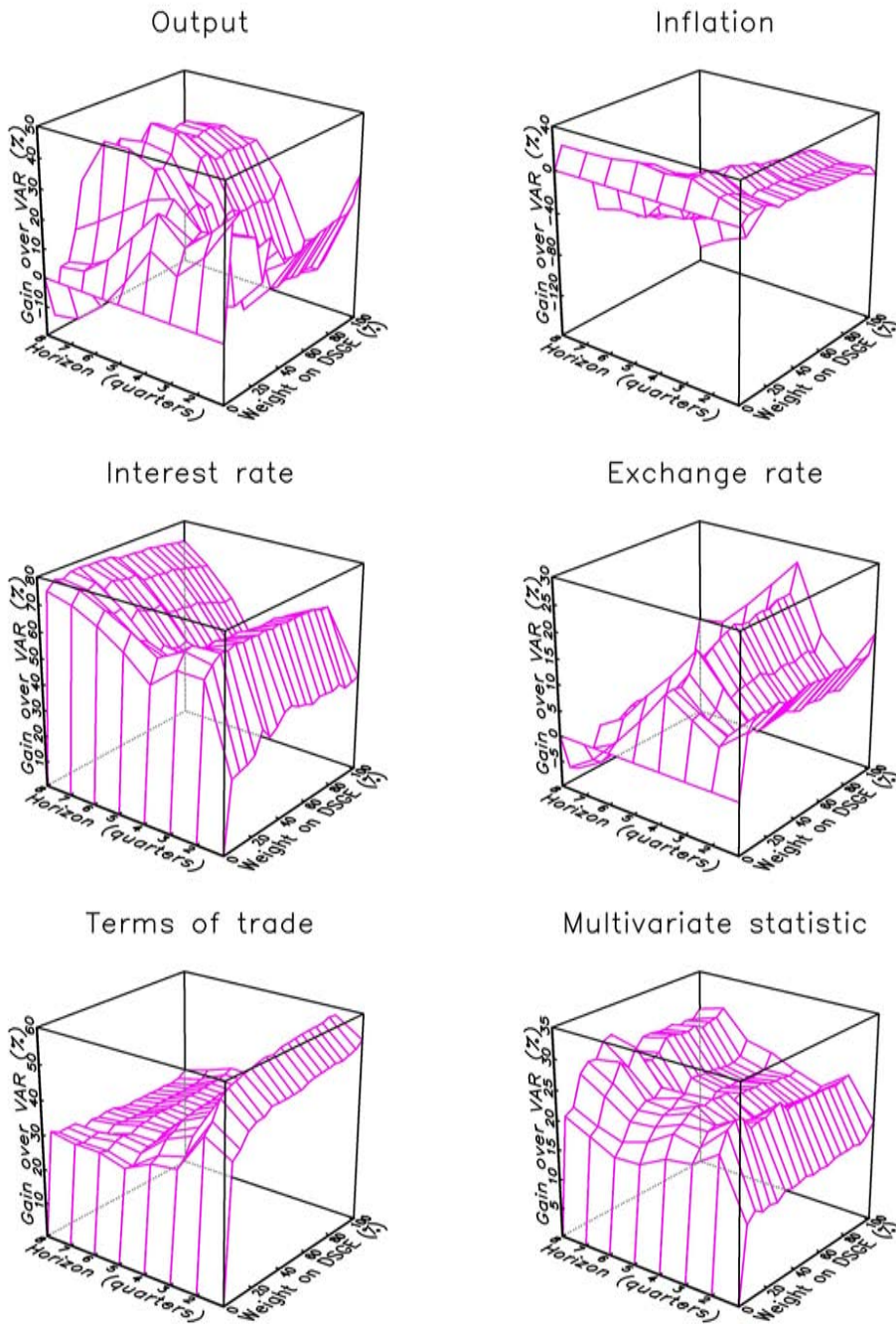
Negro and Schorfheide's (2004) forecasting results for GDP growth and inflation are broadly mirrored in New Zealand, where we find forecasting gains with higher values of λ at most horizons.

Indeed, the gains for the DSGE-VAR over the unrestricted VAR appear to be larger for New Zealand than Del Negro and Schorfheide (2004) find for the United States. This is likely because the VAR is estimated on fewer observations here than in Del Negro and Schorfheide's study.[10] The sampling variance of our estimates is reduced dramatically by increasing the weight on the prior, and it is not until the weight on the prior is large that the variance reduction is dominated by increased bias and the forecasting accuracy begins to deteriorate. This deterioration does not materialise in the terms of trade, where the DSGE-VAR improves on the unrestricted VAR regardless of the size of the artificial sample.

Del Negro and Schorfheide (2004) choose the ex-ante optimal value of $\hat{\lambda}$ to maximise the marginal data density $\Pr_\lambda(Y)$. Using this criterion, the optimal value of λ decreases over our out-of-sample period, beginning at around 3 in 1998Q4 and ending at around 2 in 2003Q3. The decreasing weight on the prior over our sample most likely reflects the decreasing sampling variance of the VAR estimates as the VAR is estimated with more data.

Table 2 details the percentage gain in MSFE over the real-time forecasts of the Reserve Bank from the VAR with the Minnesota prior (MVAR), the unrestricted VAR (UNR), the DSGE-VAR (DVAR) and the pure DSGE model with no VAR correction. In each case, we test whether the gain in the MSFE over the Reserve Bank forecasts is significant using the Diebold and Mariano (1995) test.

The variance of the mean difference in squared forecast errors is estimated using the Newey and West (1987) heteroskedasticity and autocorrelation consistent estimator, with a truncation lag of $(h-1)$. We compare the test statistic to a Student-t distribution with $(T-1)$ degrees of freedom. Note we cannot test for statistical significance of the relative multivariate statistics using the Diebold and Mariano (1995) test, so these statistics should be viewed as descriptive only.

For GDP growth the unrestricted VAR performs poorly relative to the *Monetary Policy Statement* published growth forecasts. The one quarter ahead forecast from the unrestricted VAR is 33.7 percent lower than the MSFE from the published MPS and in addition the performance of the unrestricted VAR deteriorates at

---

[10] Del Negro and Schorfheide (2004) estimate their VAR using a rolling sample of 80 quarters. Our VAR is estimated recursively beginning with 36 quarters of data and ending with 64 quarters.

Table 2
Percentage gain (loss) in MSFE over the real-time RBNZ forecasts

	GDP growth				Inflation			
h	UNR	MVAR	DVAR	DSGE	UNR	MVAR	DVAR	DSGE
1	-33.7	5.6	19.6	6.4	-19.0	0.9	-10.0	-40.1
2	12.8	23.0	36.0	18.3	-49.2	-4.7	-26.3	-83.4
3	17.0	30.2	29.1	18.3	-43.1	4.0	-9.0	-75.0
4	8.5	42.5*	29.0	28.1	-17.2	15.5	2.1	-78.4
5	-10.3	43.2*	3.7	25.2*	-9.7	21.6*	10.1	-86.3
6	-22.2	42.2	-16.3	17.0	4.0	23.9*	13.8	-110.2
7	-27.4	47.6	-26.3	2.2	10.0	24.9	15.7	-106.8
8	-20.2	43.0*	-24.4	-20.5	12.5	22.9	13.3	-114.3
	Interest rate				Exchange rate			
h	UNR	MVAR	DVAR	DSGE	UNR	MVAR	DVAR	DSGE
1	-289.5	-74.9	-165.9	-144.7	-24.9	4.7	-10.0	-3.8
2	-274.2	-21.0	-47.1	-64.0	-32.9	-0.6	-22.0	-21.8
3	-119.0	5.7	13.2	-7.6	-14.8	1.4	-0.4	1.5
4	-31.1	33.8*	14.9	24.8	-14.4	0.7	4.6	15.9
5	-23.6	43.3**	23.5	26.0	-21.0	-9.9	-9.6	4.4
6	-47.4	39.6*	32.3	29.8	-21.8	-18.9	-24.2	-8.2
7	-86.8	33.8	33.6	17.7	-19.6	-26.7	-23.3	-6.7
8	-139.6	23.9	33.3	20.5	-22.6	-38.7	-30.3	-10.4
	Terms of trade				Multivariate			
h	UNR	MVAR	DVAR	DSGE	UNR	MVAR	DVAR	DSGE
1	-10.8	46.5**	37.4*	47.8**	-19.2	-0.8	-4.9	-2.1
2	-19.8	47.8**	38.6**	49.5**	-20.8	4.0	3.2	2.9
3	-22.3	32.1*	21.5	34.8*	-16.9	3.3	2.5	3.7
4	-28.8	21.9	13.8	24.8	-11.9	10.8	8.8	10.2
5	-12.5	23.4	18.1	26.0	-11.3	11.4	8.4	12.9
6	-7.1	27.2	23.8	29.8	-12.8	10.9	5.9	12.3
7	-20.9	14.1	13.1	17.7	-18.6	7.5	3.1	11.1
8	-13.2	17.1	19.5*	20.5	-23.0	3.9	0.6	5.5

The numbers in the table reflect the percentage gain in MSFE over the real-time forecasts of the Reserve Bank: a positive number represents a gain and a negative number represents a loss. Minnesota prior (MVAR); Unrestricted VAR (UNR); VAR with DSGE prior (DVAR) and the DSGE model with no VAR correction (DSGE). The forecasts for the MVAR and DVAR are based on values of ι and λ that have highest posterior probability in each quarter. ** denotes a significant gain at the 5 per cent level. * denotes significant gain at the 10 per cent level. The relative multivariate statistics are not tested for statistical significance. All models estimated recursively in simulated real-time for 20 quarters from 1998Q4 to 2003Q3.

longer horizons. The Bayesian VAR with Minnesota prior forecasts better than the published growth forecasts across all horizons, and returns statistically significant improvements at forecast horizons four, five and eight quarters ahead.

The DSGE-VAR also forecasts well relative to the MPS forecasts, returning forecasting gains at all but longer horizons. Interestingly, the vanilla DSGE model without the VAR correction also forecasts well, returning statistically significant forecast improvement at five quarters ahead. However, this model performs poorly for forecasting inflation – forecast deterioration is reported across all horizons. In contrast, the DSGE-VAR returns forecast improvements of 10-15 percent at longer horizons while the BVAR with Minnesota prior again returns good performance.

While the unrestricted VAR returns inferior forecasts relative to MPS, the other forecasting models all return improvement in interest rate forecasts at longer horizons. The DSGE-VAR forecasts are better than the MPS forecasts at horizons three to eight quarters ahead while the DSGE model produces double digit percentage improvements, though these improvements are not statistically significant.

Both the exchange rate and the terms of trade are modelled as autoregressive processes in differences under the DSGE model. This approach returns good forecasting performance for the terms of trade.

The multivariate statistic sums the gains in forecasting performance across variables but weights the forecasts according to the variance-covariance matrix of forecast errors and allows for serial correlation in forecasts errors. The multivariate statistics show that there are overall gains from the BVAR with Minnesota prior, the DSGE-VAR and the DSGE. At longer horizons the DSGE model returns gains in forecast accuracy of 10-12 percent. The DSGE-VAR model performance is slightly worse than the DSGE model. For this data sample, the out-of-sample forecasting performance of the DSGE model is not improved with the VAR correction. However, if the central bank places significant weight on forecasting inflation over other key macroeconomic variables, the VAR correction may well be appropriate.

5 Optimal policy

5.1 Policy objectives

Forecasting the consequences of current policy behaviour is an important policy objective. However, policy-makers also need to assess how changes in their behaviour affect the economy, and to identify the welfare consequences of alternative macroeconomic policies.

As Del Negro and Schorfheide (2004) point out, policymakers may be conflicted about choosing between a pure DSGE model that may not forecast well and a VAR that forecasts well but that is not invariant to the policy regime to be considered. The DSGE-VAR framework enables policymakers to explore the trade-off between the DSGE model (which protects against the Lucas critique) and the VAR (which respects the dynamic properties of the data) by varying the weight on the DSGE model.

Our DSGE model is founded on structural parameters that describe the optimal consumption decisions of households and the optimal pricing decisions of firms. These parameters can be considered to be independent of the policy regime. Having structural models enables us to consider policy regimes as alternative rules, shifting attention away from specific interest paths tailored to a given context and towards the efficient use of the available information set (see Lucas 1976).

We direct our DSGE-VAR model to find policies that address the flexible inflation targeting objectives contained within recent Policy Targets Agreements. The PTA is the heart of the contract between the Governor of the Reserve Bank of New Zealand and the Minister of Finance. While price stability is enshrined as the primary objective of monetary policy in the Reserve Bank of New Zealand Act (1989), the 2002 PTA requires the Reserve Bank to meet the following objectives:

> 4b) In pursuing its price stability objective, the Bank shall implement monetary policy in a sustainable, consistent and transparent manner and shall seek to avoid unnecessary instability in output, interest rates and the exchange rate.

The PTA thus suggests the Reserve Bank should be *flexible* with respect to its inflation targeting objectives. To capture this flexibility, we include the variance of deviations of output from steady-state and the variance of the change in the nominal interest rate, as goal variables alongside the variance of annual inflation in the the Reserve Bank's loss function. This loss function can be represented

with the expression:

$$\min_{\psi}\{tr[\mathscr{W}\mathbf{V}(\theta,\psi)],\mathscr{B}\} \tag{24}$$

where ψ contains the policy parameters $\psi_\pi, \psi_y, \psi_{\Delta e}, \rho_R$; $\mathbf{V}$ is the variance-covariance matrix of the steady-state variables and $\mathscr{W}$ is a diagonal matrix that selects and weights variances across the goal variables (such as inflation and the change in the nominal interest rate, for example). As before $tr[\cdot]$ denotes the trace of a matrix. Since we draw from the distribution of our structural parameters, some of which have peculiar properties, we set an upper bound of 50 on our loss, $\mathscr{B}$. This bound occasionally binds for variances that are not well defined.

Since the PTA offers little direction with regard to weighting volatilities across goal variables, we consider four alternative weighting schemes. Our four weighting schemes consider permutations over two arguments: (i) inflation volatility is either equally as costly as output volatility or twice as costly; and (ii) annual inflation is either twice, or four times as costly as volatility in the change in the nominal interest rate. These schemes imply the following period loss functions:

$$
\begin{aligned}
(i) \qquad & L_t = \overline{\pi}_t^2 + y_t^2 + 0.5(i_t - i_{t-1})^2 \\
(ii) \qquad & L_t = \overline{\pi}_t^2 + 0.5y_t^2 + 0.5(i_t - i_{t-1})^2 \\
(iii) \qquad & L_t = \overline{\pi}_t^2 + y_t^2 + 0.25(i_t - i_{t-1})^2 \\
(iv) \qquad & L_t = \overline{\pi}_t^2 + 0.5y_t^2 + 0.25(i_t - i_{t-1})^2
\end{aligned}
$$

We assume the central bank does not discount future periods, such that minimising the period loss function is equivalent to minimising the intertemporal loss function over an infinite horizon.

Rather than conduct a computational search over the entire set of possible policy parameters, we restrict our search to a predefined grid of possible policy rules. In particular, we allow the following grid:

$$
\begin{aligned}
\rho_R &\in \mathbf{0.383}, 0.99, 0.95, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3 \\
\psi_\pi &\in \mathbf{3.713}, 12.0, 11.0, 10.0, 9.0, 8.0, 7.0, 6.0, 5.0, 4.0, 3.0, 2.5, 2.0, 1.5, 1.25 \\
\psi_y &\in \mathbf{0.189}, 0.5, 0.4, 0.3, 0.2, 0.1, 0.05, 0.04, 0.03, 0.02, 0.01, 0.00 \\
\psi_{\Delta e} &\in \mathbf{0.075}, 0.5, 0.4, 0.3, 0.2, 0.1, 0.05, 0.04, 0.03, 0.02, 0.01, 0.00
\end{aligned}
$$

where the estimated policy parameters are in bold.

Table 3
Macroeconomic volatility of alternative rules

	Min. loss	Est. loss	σ_y^2	$\sigma_{\bar{\pi}}^2$	$\sigma_{\Delta i}^2$	ψ_π	ψ_y	$\psi_{\Delta e}$	ρ_R
			Parameter certainty case (posterior mode)						
(i)	3.060	9.890	1.793	1.090	0.353	4.00	0.01	0.01	0.95
(ii)	2.150	8.746	1.821	0.986	0.506	5.00	0.00	0.01	0.95
(iii)	2.904	6.456	1.866	0.826	0.846	7.00	0.00	0.01	0.95
(iv)	1.962	5.312	1.906	0.705	1.215	9.00	0.00	0.01	0.95
			Parameter uncertainty case (posterior distribution)						
(i)	2.960	9.858	1.701	1.194	0.131	12.00	0.04	0.02	0.99
(ii)	2.100	8.738	1.757	0.974	0.495	5.00	0.00	0.01	0.95
(iii)	2.823	6.438	1.801	0.815	0.829	7.00	0.00	0.01	0.95
(iv)	1.913	5.319	1.839	0.696	1.191	9.00	0.00	0.01	0.95

NB. "Est. loss" gives the loss under the estimated rule with $\psi_\pi = 3.719$,
$\psi_y = 0.189, \psi_{\Delta e} = 0.075, \rho_R = 0.383$. The four loss functions considered are:
(i) $L_t = \bar{\pi}_t^2 + y_t^2 + 0.5(i_t - i_{t-1})^2$; (ii) $L_t = \bar{\pi}_t^2 + 0.5y_t^2 + 0.5(i_t - i_{t-1})^2$
(iii) $L_t = \bar{\pi}_t^2 + y_t^2 + 0.25(i_t - i_{t-1})^2$; and (iv) $L_t = \bar{\pi}_t^2 + 0.5y_t^2 + 0.25(i_t - i_{t-1})^2$

5.2   Optimal policy

We consider two cases for identifying the optimal policy rule. First, we ignore parameter uncertainty and work from the posterior mode of the model. The second case encapsulates parameter uncertainty by applying each candidate policy rule to 10,000 draws from the posterior distribution of the non-policy parameters. In the latter case, we report the mean loss for each optimal policy reaction function, where the mean loss is calculated across draws of the structural parameters. The losses are evaluated according to the four alternative weighting schemes in loss functions (i-iv).

Table 3 shows the results of the grid searches for the optimal policy rule. The first column, labelled 'Min. Loss', shows the minimum losses that can be achieved when the policy reaction function is tailored to minimise the three loss functions. The second column, 'Est. loss', shows the loss that results from the estimated policy reaction function. The remaining columns show the volatilities that result from the optimal policies, and the parameters of the optimal policies.

The first four rows of the table present the 'certainty case' that uses the posterior

Table 4
Macroeconomic volatilities from estimated rule

	σ_y^2	$\sigma_{\bar{\pi}}^2$	$\sigma_{\Delta i}^2$	ψ_π	ψ_y	$\psi_{\Delta e}$	ρ_R
Parameter certainty	2.288	0.731	13.746	0.383	3.719	0.189	0.075
Parameter uncertainty	2.217	0.769	13.686	0.383	3.719	0.189	0.075

mode. In each of the first three rows the optimal rule produces marked reductions in macroeconomic loss compared with the loss produced by the estimated rule. The optimal rule results in losses that are only $25-45$ percent of the loss incurred from the estimated rule.

Table 4 reports the macro volatilities that result from the estimated rule. The estimated rule results in slightly higher output volatility than the optimal rules, but has *lower* inflation volatility than the optimal rules from loss functions (i)-(iii). Interestingly, the deterioration in loss performance, from the estimated rule relative to the optimal rules, is driven by the greater interest rate volatility under the estimated rule.

Across all four loss functions, the optimal rules place a large weight on the lag of the nominal interest rate ($\rho_R = 0.99$) and a high response to inflation ($\psi_\pi \in \{4.00, 5.00, 7.00, 9.00\}$). Furthermore, the response to deviations of output from steady-state is essentially zero and the response to the change in the nominal exchange rate is very low across all four loss function parameterisations ($\psi_{\Delta e} = 0.01$). Lubik and Schorfheide (2005) find for a range of central banks that there is little response to the exchange rate when setting policy. For our model and specification of central banks objectives, it appears that a low response to the exchange rate is close to the optimal policy. In addition, the coefficients of the optimal rules are quite different to the prototypical Taylor rule coefficients of 1.5 on inflation and 0.5 on the output gap.

The second loss function (ii) can be thought of as a less flexible inflation-targeting regime in comparison to loss function (i). Given loss function (ii) the optimal central bank attaches less importance to stability in output, interest rates and the exchange rate. Comparing the results of these two weighting schemes in the first two rows of table 3 indicates that the optimal policy from loss function (ii) reduces inflation volatility in return for increased volatility in both output and the change in the nominal interest rate. This is achieved by responding more aggressively toward inflation.

Relative to the baseline case, loss function (iii) places a lower weight on interest rate smoothing. The optimal policy under this case generates lower inflation volatility at the expense of increased variability in the change in the nominal interest rate. Output volatility increases. The optimal rule responds even more aggressively towards inflation.

The final loss function, (iv), lowers the loss coefficients on both output volatility and the change in the nominal interest rate, in a sense combining loss functions (ii) and (iii). This loss function generates lower inflation through a more aggressive response to inflation, though at the cost of higher output and interest rate volatility. The output and interest rate volatilities appears to increase approximately linearly across the four loss function reported in each row of the table.

The second half of table 3 gives the results for the case of parameter uncertainty where we choose to work with 10,000 draws from the posterior distribution. The optimal rules produce losses similar to, or slightly smaller than, the certainty case, suggesting the parameter uncertainty embodied in the model is not particularly detrimental for monetary policy. Furthermore, the gains from employing an optimal rule rather than the estimated rule, are similar to the certainty case.

The coefficients on the optimal rule for loss function (i) look fairly different to the certainty case: the response coefficients on inflation is three times as large, there is more interest rate smoothing and stronger responses to output deviations from steady-state and the change in the nominal exchange rate not apparent in other optimal rules. However, it is unclear whether implementing the certainty rule from the first row of the table would generate large increases in loss. To this end, we map out losses from deviating from the optimal rule for loss function (i), under uncertainty, in figure 5 below.

Figure 5 contains four three-dimensional plots that maps out the change in loss that results from (possibly joint) deviations from the optimal policy parameters of loss function (i). The cell in the top left figure depicts how deviating from the optimal interest rate smoothing response and inflation response translates into higher losses for the Reserve Bank. The x-axis shows alternative inflation response setting, and alternative interest rate smoothing parameters are given on the y-axis. The vertical axis gives the loss. Dark areas on the surface correspond to areas where macroeconomic volatility is low. What is clear from the top left cell is that the model favours a high degree of interest rate smoothing. Furthermore, when interest rate smoothing is high the surface is relatively flat across the inflation response dimension, with only very low responses to inflation ($\psi_{pi} = 2$) being penalised. For example, implementing the optimal certainty equivalent rule

Figure 5
Loss accrued from deviations from the optimal rule under uncertainty
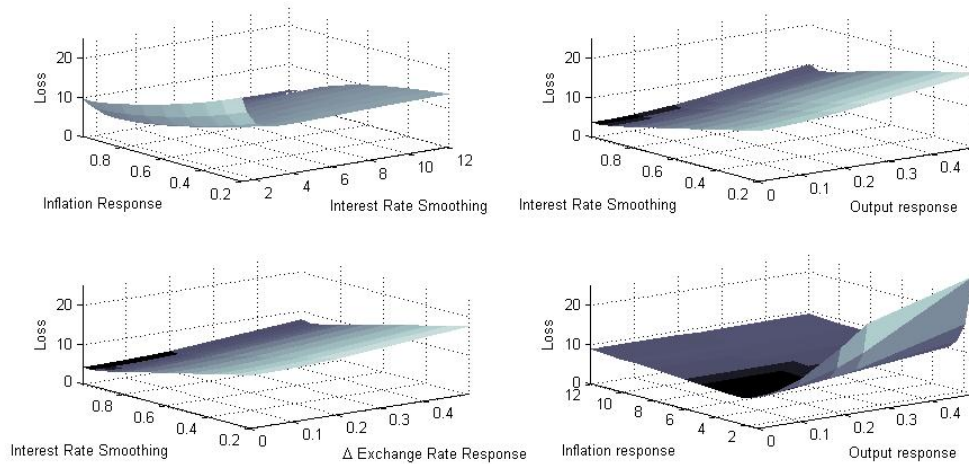
NB The figure uses loss function (i): $L_t = \overline{\pi}_t^2 + y_t^2 + 0.5(i_t - i_{t-1})^2$;

(from the first row of the table) under uncertainty generates a loss of 2.981 – only one percent larger than the loss under the rule optimised for the case of uncertainty. This suggests the differences in rules for loss function (i) are not particularly marked and one could implement either the certainty or uncertainty case without large difference in volatilities.

Figure 5 also shows the types of policies that generate particularly bad outcomes under the model. Across the cells low degrees of interest rate smoothing result in higher losses for the central bank while the picture in the top right of the figure shows that a strong response to deviation of output from steady-state is detrimental to the central bank's loss function. The figure in the bottom right of the figure retains the optimal responses to the change in the nominal exchange rate and the lag of the interest rate, but varies the output and inflation responses. Strong output responses combined with relatively low inflation responses generate particularly poor outcomes.

The similarity in optimal rules under both the certainty and uncertainty cases is borne out across the other loss function parameterisations. Returning to table 3, the rightmost columns of the table show that our grid search method settles on identical rules for loss functions (ii) to (iv) irrespective of parameter uncertainty. Thus a policymaker need not adjust the the optimal certainty equivalent rule to

incorporate the parameter uncertainty encapsulated in the posterior distribution from the DSGE model.

6   Conclusion

This paper shows the benefits of utilizing DSGE-VAR models along two dimensions: forecasting and optimal policy under uncertainty. A VAR informed by Lubik and Schorfheide's (2006) small open economy model produces forecasts comparable with, and in the case of output growth superior to, the Reserve Bank of New Zealand's judgement-adjusted published forecasts. In addition, the forecasting performance of the DSGE model with no VAR correction is competitive. At longer horizons, where monetary policy is conventionally thought to have its greatest influence, both the DSGE-VAR and the DSGE model outperform the forecasts from the Reserve Bank of New Zealand's *Monetary Policy Statement*.

The DSGE and DSGE-VAR models get close to, but do not attain, the performance of the Bayesian VAR with the Minnesota prior. However, the DSGE-VAR is informative about the structure of the economy. The DSGE structure should be robust to the Lucas critique; a policymaker that fears the Lucas critique can work with a model that places higher weight on the DSGE than is suggested by the data.

We use the DSGE-VAR to uncover the optimal policy under alternative parameterisations of flexible inflation targeting. We contrast the case of 'parameter certainty' (basing our analysis on the model defined by the posterior mode) with the case of parameter uncertainty (which uses the entire posterior distribution). We find that the optimal rules across are very similar under uncertainty and for the certainty case. The optimal rules respond aggressively to inflation and place very low weights on responding to the deviations of output from steady-state and the change in the nominal exchange rate. The optimal rules show a high degree of interest rate smoothing.

We think the DSGE-VAR is a useful modelling technology for central banks. In addition to the competitive forecasting performance reported by Del Negro and Schorfheide (2004) for the US, the DSGE-VAR produces good forecasting performance for New Zealand, a small open economy with an inflation-targeting central bank. The ability to forecast well and yet obtain economic structure suggests that the DSGE-VAR may also be a useful forecasting and policy analysis tool for other central banks.

References

Adolfson, M, J Lindé, and M Villani (2005), "Forecasting performance of an open economy dynamic stochastic general equilibrium model," *Sveriges Riksbank*, *Working Paper*, 190.

Basdevant, O, N Björksten, and O Karagedikli (2004), "Estimating a time varying neutral real interest rate for New Zealand," *Reserve Bank of New Zealand*, *Discussion Paper*, DP2004/01.

Benigno, P (2004), "Optimal monetary policy in a currency area," *Journal of International Economics*, 63(2), 293–320.

DeJong, D N, B F Ingram, and C H Whiteman (2000), "A Bayesian approach to dynamic macroeconomics," *Journal of Econometrics*, 98(2), 203–23.

Del Negro, M and F Schorfheide (2003), "Take your model bowling: Forecasting with general equilibrium models," *Federal Reserve Bank of Atlanta Economic Review*.

Del Negro, M and F Schorfheide (2004), "Priors from general equilibrium models for VARs," *International Economic Review*, 45(2), 643–73.

Dennis, R (2005), "Specifying and estimating New Keynesian models with instrument rules and optimal monetary policies," *Federal Reserve Bank of San Francisco, Working Paper*, 2004-17.

Diebold, F X and R S Mariano (1995), "Comparing predictive accuracy," *Journal of Business and Economic Statistics*, 13(3), 253–63.

Doan, T, R Litterman, and C Sims (1984), "Forecasting and conditional projection using realistic prior distributions," *Econometric Reviews*, 3, 1–100.

Fukač, M and A Pagan (2006), "Issues in adopting DSGE models for use in the policy process," *Australian National University, Centre for Applied Macroeconomic Analysis*, *CAMA Working Paper*, 10/2006.

Fukač, M, A Pagan, and V Pavlov (2006), "Econometric issues arising from DSGE models," *Mimeo*, paper presented at Reserve Bank of New Zealand, 'Macroeconomics and Model Uncertainty Conference'.

Gali, J and T Monacelli (2005), "Monetary policy and exchange rate volatility in a small open economy," *Review of Economic Studies*, 72(3), 707–34.

Gelman, A, J B Carlin, H S Stern, and D B Rubin (2004), *Bayesian Data Analysis*, Chapman and Hall/CRC, Boca Raton, Florida.

Geweke, J (1992), "Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments," in *Bayesian Statistics 4: Proceedings of the Fourth Valencia International Meeting*, eds J M

Bernardo, J O Berger, A P Dawid, and A F M Smith, 169–194, Oxford University Press, Oxford.

Geweke, J (2005), *Contemporary Bayesian Econometrics and Statistics*, John Wiley and Sons Ltd, Hoboken, New Jersey.

Hamilton, J D (1994), *Time Series Analysis*, Princeton University Press, Princeton, New Jersey.

Heckman, J L (1999), "Causal parameters and policy analysis in economcs: A twentieth century retrospective," *National Bureau of Economic Research*, *Working Paper*, 7333, available at http://ideas.repec.org/p/nbr/nberwo/7333.html.

Heidelberger, P and P Welch (1983), "Simulation run length control in the presence of an initial transient," *Operations Research*, 31, 1109–44.

Ingram, B F and C H Whiteman (1994), "Supplanting the 'Minnesota' prior: Forecasting macroeconomic time series using real business cycle model priors," *Journal of Monetary Economics*, 34(3), 497–510.

Justiniano, A and B Preston (2006), "Monetary policy and uncertainty in an empirical small open economy model," *Mimeo*, paper presented at Reserve Bank of New Zealand, 'Macroeconometrics and Model Uncertainty Conference'.

Koop, G (2003), *Bayesian Econometrics*, John Wiley and Sons Ltd, Chichester, Sussex.

Litterman, R B (1986), "Forecasting with Bayesian vector autoregressions – five years of experience," *Journal of Business and Economic Statistics*, 4(1), 25–38.

Lubik, T and F Schorfheide (2004), "Testing for indeterminacy: An application to U.S. monetary policy," *American Economic Review*, 94(1), 190–217.

Lubik, T and F Schorfheide (2005), "Do central banks respond to exchange rate movements? A structural investigation," *Mimeo*, forthcoming Journal of Monetary Economics.

Lucas, R E, Jr (1976), "Econometric policy evaluation: A critique," *Carnegie Rochester Conference Series on Public Policy*, 1, 19–46.

Newey, W K and K West (1987), "A simple, positive semidefinite, heteroskedasticity and autocorrelation consistent covariance matrix," *Econometrica*, 55, 703–708.

R Development Core Team (2006), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, URL http://www.R-project.org.

Santacreu, A M (2005), "Reaction functions in a small open economy: What

role for non-traded inflation?" *Reserve Bank of New Zealand*, *Discussion Paper*, DP2005/04.

Wold, H (1938), *A Study in the Analysis of Stationary Time Series*, Almqvist and Wiksell, Uppsala.

Appendix

A   Markov chain Monte Carlo convergence

Bayesian inference relies on accurate characterisation of the posterior. In our model we cannot derive the posterior analytically and hence numerically simulate the posterior distribution using the Metropolis-Hastings Markov chain Monte Carlo simulator (see Geweke 2005 and Gelman, Carlin, Stern, and Rubin 2004 for discussion). The table below reports the results from two tests of chain convergence, applied to $500,000$ draws, having burned off an initial 3.5 million.[11]

Column 2 uses the Heidelberger-Welch test to assess whether each parameter in the chain has converged to a stationary distribution. The test uses the Cramer-von-Mises statistic. If the entire chain does not satisfy the test then it is repeated discarding the first $10, 20, 30, \ldots, 50$ percent of the observations in the chain. If the latter 50 percent of the chain is not stationary, the chain fails the test.

Column 5 indicates whether the chain passes the half-width test, again based on the work of Heidelberger and Welch (1983). This test takes the length of chain that passes the convergence test in column 2 and calculates a 95 percent confidence interval for the mean of the parameter. Half the width of the interval is divided by the estimated of the mean. If this results in a value lower than epsilon, the chain is sufficiently long to estimate the parameter with the required accuracy. We set epsilon to 0.05 and all but the parameter ρ^* pass this test.

Finally, the right-most column of the table reports z-scores based on the convergence diagnostic proposed by Geweke (1992). This is a test based on equality of means of the first and last part of the Markov chain. We define the first part of the chain to be the initial 25 percent of the chain and the last part of the chain to be the last 50 percent of the chain. The z-scores are asymptotically standard normal and are simply the difference in the two sample means divided by the standard error (calculated using spectral methods). According to Geweke's (1992) convergence criterion, α has not converged at the 5 percent level of significance, although the parameter is similar to parameter estimates reported in other studies (see Lubik and Schorfheide 2005, for example). While some doubt remains about the posterior reported for this parameter, overall, we think the tests indicate chain convergence such that the numerically simulated posterior is a good approximation.

---

[11] These statistics were obtained using the CODA library in the statistical package R. See R Development Core Team (2006).

Table 5
Prior and posterior distributions: New Zealand 1990Q1 – 2005Q4

Para	HW-test 1	*p*-value	HW test 2	Mean	Halfwidth	z-score
ψ_π	pass	0.450	pass	3.790	0.013	0.582
ψ_y	pass	0.403	pass	0.187	0.002	-0.652
$\psi_{\Delta e}$	pass	0.505	pass	0.075	0.001	0.730
ρ_R	pass	0.955	pass	0.378	0.002	0.114
α	pass	0.634	pass	0.208	0.001	-0.657
ρ^*	pass	0.292	pass	0.668	0.009	-0.573
κ	pass	0.492	pass	1.563	0.009	-0.747
τ	pass	0.690	pass	0.545	0.004	-0.207
ρ_q	pass	0.236	pass	0.202	0.002	-0.845
ρ_a	pass	0.581	pass	0.312	0.002	-0.142
ρ_{y^*}	pass	0.098	pass	0.910	0.001	2.000
ρ_{π^*}	pass	0.589	pass	0.619	0.002	1.290
$\tilde{\pi}$	pass	0.808	pass	2.140	0.009	0.247
$\tilde{\gamma}$	pass	0.462	pass	0.777	0.002	-0.869
σ_q	pass	0.053	pass	0.869	0.005	0.248
σ_z	pass	0.664	pass	1.640	0.002	1.803
σ_r	pass	0.852	pass	0.831	0.004	1.024
σ_{y^*}	pass	0.664	pass	2.540	0.082	0.133
σ_{π^*}	pass	0.876	pass	2.378	0.006	0.230

HW-test 1 is the Heidelberger-Welch (1983) test that the sample parameter values originate from a stationary distribution; HW-test 2 is the Heidelberger-Welch (1983) test that the sample parameters values have converged to a sufficient tolerance level.