DP2008/Preliminary Draft

Real-time conditional forecasts with Bayesian VARs: An application to New Zealand

Chris Bloor and Troy Matheson

December 2008

JEL classification: C11, C13, C33, C53

www.rbnz.govt.nz/research/discusspapers/

Discussion Paper Series

ISSN 1177-7567

DP2008/Preliminary Draft

Real-time conditional forecasts with Bayesian VARs: An application to New Zealand*

Chris Bloor and Troy Matheson†

Abstract

We examine the real-time forecasting performance of Bayesian VARs (BVARs) of different sizes using an unbalanced data panel. In a real-time out-of-sample forecasting exercise, we find that our BVAR methodology outperforms univariate and VAR benchmarks, and produces comparable forecast accuracy to the judgementally-adjusted forecasts produced internally at the Reserve Bank of New Zealand. We analyse forecast performance and find that, while there are trade offs across different variables, a 35 variable BVAR generally performs better than smaller or larger specifications. Finally, we demonstrate some techniques for imposing judgement and for forming a semi-structural interpretation of the BVAR forecasts.

* The views expressed in this paper are those of the author(s) and do not necessarily reflect the views of the Reserve Bank of New Zealand. All errors and omissions are ours and the views expressed are not necessarily those of the Reserve Bank of New Zealand.

† Address: Economics Department, Reserve Bank of New Zealand, 2 The Terrace, PO Box 2498, Wellington, New Zealand. *email address:* Troy.Matheson@rbnz.govt.nz.
ISSN 1177-7567

1 Introduction

An important part of policy making in real time is forming an up to date picture of the near-term outlook for the economy. Typically central bank forecasters analyse a large number of data series to form a judgemental forecast of the current position of the economy and of the very near future. These forecasts have critical importance in the policy process for an inflation targeting central bank such as the Reserve Bank of New Zealand (RBNZ).

A challenge facing model builders is how to incorporate the information from a large number of data series in a systematic fashion in real time. Data tend to be released incrementally throughout the quarter, so that not all series of interest are available to the forecaster when they wish to make their forecasts. In addition, data tend to be revised over time, so that forecast evaluation conducted on revised data may misrepresent forecast accuracy in real time (Croushore 2006).

The literature on factor models has provided one avenue for using large data sets to produce forecasts using unbalanced panels of data in real time (Gianonne *et al* 2008). This approach has been applied to the New Zealand case in Matheson (2007).

Recently, another approach has been developed for forecasting using large data panels. De Mol *et al* (2008) show that if the data is characterised by an approximate factor structure, a Bayesian forecast based on point estimates converges to the optimal forecast as long as the prior is imposed more tightly as the model size increases. Building on this, Banbura *et al* (2008) develop a large Bayesian VAR (BVAR) with Litterman (1986) and sums of coefficients (Doan, Litterman, and Sims 1984) priors containing 108 US variables. They find that the forecasting performance of this model compares favourably to that of competing specifications.

An advantage of BVARs is that, in contrast to factor models that typically work with transformed data, they can be estimated on non-stationary levels. As a result, information from long-run cointegrating relationships is retained in the BVAR forecast. In addition, the process of combining a large quantity of data into factors means that it is often difficult to disentangle an economic story from factor model forecasts. As each data series enters the BVAR individually, it is easier to identify the effect that each variable is having on the forecasts, and shock decompositions can be more meaningfully calculated.

Bloor and Matheson (2008) extended the model of Banbura *et al* (2008) to the New Zealand case, and calculated impulse responses for a range of shocks.

The model of Banbura *et al* (2008) was extended by incorporating the co-persistence prior of Sims (1993). In addition, restrictions were imposed on lagged variables, and the priors were implemented on a block-by-block basis with the Zha (1999) estimation methodology. Using a balanced panel of revised data, a large BVAR containing 95 variables was able to produce more accurate forecasts than most benchmarks in an out-of-sample forecasting exercise.

Since Bloor and Matheson (2008) focussed on impulse responses, the model was not designed to forecast with unbalanced panels. In this paper, we use similar model specifications to Bloor and Matheson (2008). In addition, we employ the Waggoner and Zha (1999) conditional forecasting estimation techniques to forecast series that are missing at the end of the data panel. These techniques allow us to apply exogenous paths or impose shocks to any variable in the model in a model-consistent manner.

We conduct a real-time out-of-sample forecasting exercise and find that our BVAR methodology produces more accurate forecasts than a range of univariate and VAR forecasts. Moreover, the BVAR forecasts show comparable forecast accuracy to the judgementally-adjusted forecasts produced internally at the RBNZ. We analyse forecast performance as the size of the BVAR model increases. While the results are not conclusive, and differ across variables, we find that a 35 variable specification generally performs better than smaller or larger models.

To demonstrate how this model can be used in the policy environment we consider an alternative scenario. In this scenario we consider a counterfactual experiment in which a forecaster in 2006Q4 was able to perfectly predict the sharp run-up in commodity prices over the following two years. Also, to further highlight the usefulness of the BVAR methodology, we outline techniques for analysing shock decompositions on a block-by-block basis, and use this to interpret the BVAR forecasts at a single point of time. These techniques avoid the need to make strong identifying assumptions, and allow a semi-structural interpretation of the BVAR forecasts to be made.

The remainder of the paper is organised as follows. Section 2 outlines the BVAR methodology and the Waggoner and Zha (1999) conditional forecasting algorithm. Section 3 outlines the data and model specifications used, and section 4 outlines the forecasting exercise. Section 5 describes the forecasting results. Section 6 discusses tools that can be used to interpret the forecasts and we conclude in section 7.

2 Methodology

2.1 The Bayesian VAR

Let $Y_t = (y_{1,t}, y_{2,t}, \ldots, y_{N,t})'$ be a set of time series with a reduced-form VAR(p) representation:

$$Y_t = c + \sum_{k=1}^{p} B_k Y_{t-k} + u_t \tag{1}$$

where $c = (c_1, \ldots, c_N)'$ is an n-dimensional vector of constants, B_k is an $N \times N$ autoregressive matrix, and u_t is an N-dimensional white noise process with covariance matrix $\mathrm{E}u_t u_t' = \Psi$.

The Litterman (1986) prior, often referred to as the Minnesota prior, shrinks the diagonal elements of B_1 towards one and the other coefficients $(B_1, \ldots, B_p)$ towards zero:

$$Y_t = c + Y_{t-1} + u_t \tag{2}$$

The moments for the prior distribution of the coefficients are:[1]

$$\mathrm{E}[(B_k)_{ij}] = \begin{cases} \delta_i, & j = i, k = 1 \\ 0, & \text{otherwise} \end{cases} \quad \text{and} \quad \mathrm{V}[(B_k)_{ij}] = \frac{\lambda^2}{k^2} \frac{\sigma_i^2}{\sigma_j^2} \tag{3}$$

The Minnesota prior thus embodies the belief that more recent lags provide more useful information than more distant ones. The coefficients $B_1, \ldots, B_p$ are assumed to be independent and normally distributed, and the covariance matrix of the residuals is assumed to be diagonal, fixed and known (ie $\Psi = \Sigma$, where $\Sigma = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_N^2)$). The prior on the intercept is diffuse. Note that the random walk prior, $\delta_i = 1$ for all i, reflects a belief that all the variables are highly persistent. However, the researcher can also incorporate priors where some variables are characterised by a degree of mean-reversion, $0 \leq \delta < 1$.

The overall tightness of the prior distribution around δ_i is governed by the hyperparameter λ: $\lambda = 0$ imposes the prior exactly so that the data do not inform the parameter estimates, and $\lambda = \infty$ removes the influence of the prior

[1] Note that Litterman's original assumption that the residual covariance matrix is fixed and diagonal has been removed from 3 by imposing a normal prior distribution for the coefficients and an inverse Wishart prior distribution for the covariance matrix of the residuals Ψ, the so-called Inverse-Wishart prior (see Kadiyala and Karlsson 1997 and Sims and Zha 1998).

altogether. The factor $1/k^2$ is the rate at which the prior variance decreases with the lag length of the VAR, and σ_i^2/σ_j^2 accounts for the different scale and variability of the data.

The sums of coefficients prior of Doan *et al* (1984) is a modification of the Minnesota prior that is motivated by the frequent practice of specifying a VAR in first differences. The sums of coefficients prior is best described by writing the VAR in error correction form:

$$\Delta Y_t = c - (I_N - B_1 - \ldots - B_p)Y_{t-1} + C_1 \Delta Y_{t-1} + \ldots + C_{p-1}\Delta Y_{t-p+1} + u_t \quad (4)$$

The sums of coefficients prior shrinks $(I_N - B_1 - \ldots - B_p)$ towards zero, where a hyperparameter τ controls the degree of shrinkage. As $\tau \to 0$ the VAR will increasingly satisfy the prior, while higher values of τ will loosen the prior until, when $\tau = \infty$, the prior has no influence on VAR estimates. The sums of coefficients restriction implies that there are as many stochastic trends in the VAR as there are $I(1)$ variables. Sims (1993) introduced a prior that makes some allowance for stable, long-run cointegrating relationships amongst the variables in the system. This 'co-persistence' prior is governed by the hyperparameter θ. As $\theta \to 0$, the VAR will increasingly satisfy the prior, while as $\theta \to \infty$ there will be increasingly more stochastic trends in the system.

The priors described above outline are what Robertson and Tallman (1999) call the modified Litterman prior.[2]

Writing the VAR in matrix notation yields:

$$Y = XB + U \quad (5)$$

where $Y = (y_1, \ldots, y_T)'$, $X = (X_1, \ldots, X_T)'$, $X_t = (Y'_{t-1}, \ldots, Y'_{t-p}, 1)$, $U = (u_1, \ldots, u_T)'$, and $B = (B_1, \ldots, B_p, c)'$ is the $k \times N$ matrix of coefficients with $k = Np + 1$. The form of the prior is then:

$$\Psi \sim iW(S_0, \alpha_0) \quad \text{and} \quad B|\Psi \sim N(B_0, \Psi \otimes \Omega_0) \quad (6)$$

where the parameters B_0, Ω_0, S_0, and α_0 satisfy the prior expectations for B and Ψ.

We implement the modified Litterman prior by adding dummy observations to the system (5). It can be shown that adding T_d dummy observations

[2] Robertson and Tallman (1999) find that the modified Litterman prior produces relatively good forecasts of unemployment, inflation, and GDP growth in the US compared to the Litterman (1986) prior and the Sims and Zha (1998) prior.

Y_d and X_d is equivalent to imposing the Inverse-Wishart prior with $B_0 = (X_d'X_d)^{-1}X_d'Y_d$, $\Omega = (X_d'X_d)^{-1}$, $S_0 = (Y_d - X_dB_0)'(Y_d - X_dB_0)$, and $\alpha_0 = T_d - k - N - 1$. The following dummy observations match our prior moments:

$$Y_d = \begin{pmatrix} \text{diag}(\delta_1\sigma_1, \ldots, \delta_N\sigma_N)/\lambda \\ 0_{N(p-1)\times N} \\ \cdots\cdots\cdots \\ \text{diag}(\delta_1\mu_1, \ldots, \delta_N\mu_N)/\tau \\ \cdots\cdots\cdots \\ J \\ \cdots\cdots\cdots \\ \text{diag}(\sigma_1, \ldots, \sigma_N) \\ \cdots\cdots\cdots \\ 0_{1\times N} \end{pmatrix} \quad X_d = \begin{pmatrix} K_d \otimes \text{diag}(\sigma_1, \ldots, \sigma_N)/\lambda & 0_{Np\times 1} \\ \cdots\cdots\cdots & \cdots \\ K \otimes \text{diag}(\delta_1\mu_1, \ldots, \delta_N\mu_N)/\tau & 0_{N\times 1} \\ \cdots\cdots\cdots & \cdots \\ (J_1, \ldots, J_p)_{1\times Np} & 1/\theta \\ \cdots\cdots\cdots & \cdots \\ 0_{N\times Np} & 0_{N\times 1} \\ \cdots\cdots\cdots & \cdots \\ 0_{1\times N} & \epsilon \end{pmatrix} \tag{7}$$

where $J = (\delta_1\mu_1, \ldots, \delta_N\mu_N)/\theta$, $K = 1, \ldots, p$, $K_d = \text{diag}(K)$, and ϵ is a very small number.[3] Generally speaking, the first block of dummies impose prior beliefs on the autoregressive coefficients, the second block of dummies impose the sums of coefficients prior, the third block of dummies impose the co-persistence prior, and the fourth and fifth blocks impose the priors for the covariance matrix and the intercepts, respectively. Following common practice, we set the prior for the scale parameter σ_i equal to the residual standard deviation from a univariate autoregressive regression with p lags for variable $y_{i,t}$. Likewise, the parameter μ_i (the prior for the average level of variable $y_{i,t}$) is set equal to the sample average of variable $y_{i,t}$.

Augmenting the system with dummy observations yields:

$$Y^* = X^*B + U^* \tag{8}$$

where $Y^* = (Y', Y_d')'$, $X^* = (X', X_d')'$ and $U^* = (U', U_d')'$. After adding the diffuse prior $\Psi \propto |\Psi|^{-(N+3)/2}$, which ensures the existence of the prior expectation of Ψ, the posterior has the form:

$$\Psi|Y \sim iW(\hat{\Sigma}, T_d + 2 + T - k) \text{ and } B|\Psi, Y \sim N(\hat{B}, \Psi \otimes (X^{*\prime}X^*)^{-1}) \tag{9}$$

where $\hat{B} = (X^{*\prime}X^*)^{-1}X^{*\prime}Y^*$ and $\hat{\Sigma} = (Y^* - X^*\hat{B})'(Y^* - X^*\hat{B})$ (Banbura *et al* 2008). Thus, the posterior expectation of the parameters coincide with the OLS estimates of the dummy-augmented system (8). The dummy observations (7) also make it clear that as λ, τ, and θ tend to infinity the Minnesota, sums of coefficients, and co-persistence dummies will tend to zero, and the posterior parameter estimates will tend to the OLS estimates from the original, un-augmented system (5).

[3] Note: if v is a vector of dimension $1 \times v_N$, the operation diag(v) yields a $v_N \times v_N$ matrix with v on the diagonal and zeros elsewhere.

2.2 Tailoring the prior to penalise over-fitting

Adding more variables to a classical regression leads to a deterioration in the parameter estimates – over-fitting. However, in the context of Bayesian regression, De Mol *et al* (2008) show that a forecast based on point estimates converges to the optimal forecast as long as the tightness of the prior (the degree of shrinkage) increases as the number of time series N becomes larger. Using a similar algorithm to Banbura *et al* (2008), the tightness of the prior can be increased as N increases by:

1. Selecting N^* (where $N^* < N$) benchmark variables for which in-sample fit will be evaluated;

2. Evaluating the in-sample fit of a VAR estimated with OLS on the N^* benchmark variables;

3. Setting the sums of coefficients hyperparameter τ and the co-persistence hyperparameter θ to be proportionate to the overall tightness hyperparameter λ ($\tau = \phi_1\lambda$ and $\theta = \phi_2\lambda$, where $\phi_1 \geq 0$ and $\phi_2 \geq 0$);

4. Choosing the overall tightness hyperparameter λ to have the same in-sample fit as the benchmark VAR.

We follow Banbura *et al* (2008) by defining in-sample fit as a measure of relative 1-step-ahead mean squared error (MSE) evaluated using the training sample $t = 1, \ldots, T-1$. The MSE for variable i for a given λ is:

$$MSE_i^\lambda = \frac{1}{T-p-1} \sum_{t=p}^{T-2} (y_{i,t+1|t}^\lambda - y_{i,t+1})^2 \tag{10}$$

The variables can then ordered so that the N^* baseline variables are ordered first. The overall tightness hyperparameter (λ) for a given measure of baseline fit (FIT) is can then be found by conducting a grid search over λ:

$$\lambda(\text{FIT}) = \arg\min_\lambda \left| \text{FIT} - \frac{1}{N^*} \sum_{i=1}^{N^*} \frac{MSE_i^\lambda}{MSE_i^0} \right| \tag{11}$$

where MSE_i^0 is the MSE of variable i with the prior restriction imposed exactly ($\lambda = 0$), and baseline fit is defined as the average relative MSE from an OLS-estimated VAR containing the N^* baseline variables:

$$\text{FIT} = \frac{1}{N^*} \sum_{i=1}^{N^*} \frac{MSE_i^\infty}{MSE_i^0} \tag{12}$$

It is clear that there are a multitude of ways to increase the tightness of the prior as the number of variables increases. A researcher, for example, could choose λ such that the average fit across all N variables matched the average fit over the N^* baseline variables in the unrestricted VAR. Or, as analysed in Banbura *et al* (2008), the researcher could choose λ such that the average fit on the N^* baseline variables in the Bayesian VAR is lower than that implied by the unrestricted VAR.

Indeed, due to the short sample period available in our forecasting excercise (section 4), we found that the Banbura *et al* (2008) methodology tended to show signs of overfitting the data, and hence produced poor forecasts. We found better results by imposing FIT=0.5. For all models λ has been set to achieve this fit over the baseline variables. In addition, the forecasting performance for alternative values of ϕ_1 and ϕ_2, the tightness with which the sums of coefficients and co-persistence priors are imposed, was investigated. While the forecasting results are generally robust to different values, we found that ϕ_1=10 and ϕ_2=100 produced the best results. Thus, we use this specification in the remainder of the paper.[4]

2.3 A structural VAR

We have outlined a BVAR methodology in which each variable is a linear function of lags of all variables in the system – the VAR is symmetric. However, in a small open economy like New Zealand, foreign variables are key determinants of the business cycle, while domestic variables are not likely to have much influence on the foreign variables. It thus makes economic sense to make the lags of foreign variables exogenous to the domestic variables (see, for example, Cushman and Zha 1997 and Zha 1999). Bayesian inference lag restrictions can readily be made using the estimation methods laid out in Zha (1999) or Waggoner and Zha (2003).

In this paper, we impose lag restrictions and implement what Zha (1999) calls strong recursive blocks in the contemporaneous matrix.[5] An important feature of the strongly recursive identification scheme is that it can be readily applied to large systems. This scheme allows different lag assumptions across different blocks of equations and, perhaps more importantly, a forecast from

[4] Forecast results under a range of different prior specifications are available from the authors on request.

[5] We use identification assumptions for estimation when lag restrictions are imposed: OLS estimates are no longer the most efficient.

a given block of equations is invariant to the ordering of the variables within that block.[6]

To better describe a VAR with strong recursive blocks, consider the structural form of our VAR:

$$A_0 Y_t = C + \sum_{k=1}^{p} A_k Y_{t-k} + \epsilon_t \tag{13}$$

where A_0 is the $N \times N$ contemporaneous coefficient matrix and ϵ_t is an N-dimensional vector of structural disturbances.

Now, we partition the system into n blocks of equations, where each block i has the form:

$$A_{i,j,0} Y_t = C_i + \sum_{k=1}^{p} A_{i,j,k} Y_{t-k} + \epsilon_{i,t}, \quad i = 1, \ldots, n \tag{14}$$

where $i = 1, \ldots, n$ and $A_{i,j,k}$ is an $m_i \times m_j$ matrix with $m_1 + \ldots + m_n = N$. Notice that $i = j$ implies that all variables enter all equations (as in 13), while when $i \neq j$ the VAR is asymmetric in the sense that not all variables enter into all equations. The model 13 has strong recursive blocks in the contemporaneous coefficient matrix A_0 if $A_{i,j,0} = 0$ for $i > j$ and $A_{i,j,0}$ is unrestricted for $j \geq i$ (Zha 1999). Clearly, when lag structures do not differ across blocks, a VAR identified using a Cholesky decomposition of the estimated covariance matrix is equivalent to 14, where each equation forms its own block.

Assuming strong recursive blocks in the contemporaneous matrix, we can break the posterior distribution into blocks. Specifically, in matrix notation, each block is:

$$\underset{(T \times m_i)}{Y_i^*} = \underset{(T \times k_i)}{X_i^*} \underset{(k_i \times m_i)}{B_i} + \underset{(T \times m_i)}{U_i^*}, \quad i = 1, \ldots, n \tag{15}$$

where Y_i^* is a dummy-augmented matrix of observations of contemporaneous variables, X_i^* is a dummy-augmented matrix containing lagged variables as well as contemporaneous variables from other blocks (Y_j^*s from $j > i$), U_i^* is the matrix form of $A_{i,i,0}^{-1} \epsilon_{i,t}$, and k_i is the total number of right-hand-side variables in each equation in the ith block. The posterior estimates described above (9) become block specific, and each block can estimated separately using OLS.

[6] In general, this result does not apply when the ordering of the blocks changes.

A potential problem with a block-by-block approach here is that the tightness of the Bayesian prior (λ, τ and θ) will be the same across each block of equations, because all blocks of equations will be linked to the in-sample fit for the N^* baseline variables (section 2.2). To break this link, we define λ^{m_i}, τ^{m_i} and θ^{m_i} to be block-specific hyperparameters for each block of equations.

Essentially, for each of the n blocks of equations, we re-define the hyperparameters in (7), and select the appropriate columns from the dummy-augmented matrices (8) to construct the system 15. Notice that this method affords much flexibility in specifying the lagged relationships in each block, allowing the variables contained in any particular block to be exogenous to any other block (or subset of blocks), where the hyperparameters λ^{m_i}, τ^{m_i} and θ^{m_i} are chosen in a block-specific way. Indeed, if there is more than one large block of equations in the system, the algorithm outlined in section 2.2 can be used to set the hyperparameters for each of the large blocks.[7] In this paper, we only have one large block m_1, the endogenous block of domestic variables, and one foreign block m_2. For the foreign block, we use the hyperparameters $\lambda^{m_2} = \tau^{m_2} = \theta^{m_2} = 1$. Throughout the paper, we order each contemporaneous block to be upper triangular ($A_{i,i,0}$ is upper triangular), so that A_0 is upper triangular.

2.4 Conditional forecasting

The block-specific parameter estimates from system 15 are transformed into a reduced-form VAR for forecasting purposes:

$$Y_t = c + \sum_{k=1}^{p} B_k Y_{t-k} + A_0^{-1} \epsilon_t \tag{16}$$

where the relationships between the reduced-form parameters from system 1 and the structural parameters from system 13 are $c = A_0^{-1}C$, $B_k = A_0^{-1}A_k$, with $u_t = A_0^{-1}\epsilon_t$. Given data up to time T, the h-step out-of-sample forecast at time T can then be decomposed:

$$Y_{T+h} = D + \sum_{j=1}^{h} M_{h-j} \epsilon_{T+j}, \qquad h = 1, 2, \ldots \tag{17}$$

[7] Choose $N^{m_i,*}$ baseline variables from large block i (where $N^{m_i,*} < N^{m_i}$ and N^{m_i} is the number of endogenous variables in block i) and select the hyperparameters λ^{m_i}, τ^{m_i} and θ^{m_i} using the algorithm in section 2.2.

where:

$$M_0 = A_0^{-1}$$
$$M_i = \sum_{j=1}^{i} B_j M_{i-j}, \quad i = 1, 2, \ldots$$

$$B_j = 0 \quad \text{for} \quad j > p$$

This forecast decomposition (17) consists of two parts. The first term, D, includes the initial conditions and produces dynamic forecasts in the absence of shocks, while the second term is the dynamic impact of future structural shocks. Future shocks impact on the variables in the VAR through the matrix of impulse response M_i. A conditional forecast is then defined to be when constraints are imposed on future values of variables and/or shocks.

We construct conditional forecasts on the basis of imposing future values for some variables (or, equivalently, for future reduced-form shocks). Doan *et al* (1984) show that a unique and optimal (in the least squares sense) vector of forecast errors that satisfy the constraints on the forecasts is given by:

$$\epsilon = R'(RR')^{-1}r \tag{18}$$

where R is a $q \times k$ stacked matrix from the impulse responses $M_{h-h_n}(.,j)$, ϵ is a $k \times 1$ vector correspondingly stacked from ϵ_{t+h_n}, and r is a $q \times 1$ vector of constraints, where k is the total number of future shocks, q is the number of constraints, and $h_n = 1, \ldots, h$.[8]

Waggoner and Zha (1999) show that with conditions imposed on future variables (or reduced form shocks) the forecast distribution is invariant to orthonormal transformation of the system.[9]

Waggoner and Zha (1999) also outline a Gibbs sampling algorithm that allows parameters estimates to be conditional on the constraints in 18. Our estimation technique thus combines the block-specific estimation methodology outlined in Zha (1999) with the conditional forecasting estimation algorithm outlined in Waggoner and Zha (1999).[10]

[8] See Robertson and Tallman (1999) for an intuitive illustration of the Doan *et al* (1984) technique.

[9] Note, however, that with strong recursive blocks and lag restrictions, this result only holds for a given ordering of blocks.

[10] Briefly, forecasting requires iteration over the following steps: 1) Estimate the parameters of the structural VAR (15); 2) Forecast with the reduced-form model (16); 3) Re-estimate the structural parameters conditional on the forecast constraints (18); 4) Forecast with the newly-parameterised reduced-form model.

3 Data and model specifications

3.1 Data

All of the models are estimated using quarterly data spanning 1990Q1 to 2008Q3. The largest model we consider consists of 50 time series covering a broad range of categories, including business and consumer confidence, the housing market, the labour market, consumption and investment, production, financial markets, and the world economy. All series in the panel are seasonally adjusted using Census X12 prior to estimation. The series that are expressed in percentages (eg interest rates and unemployment rates) and those that can take negative values (eg balances of opinion and net migration) are left as levels. We transform the remainder of the series by applying natural logarithms and multiplying by 100. For most of the variables in the panel we use the random walk prior $\delta_i = 1$. However, some of the variables in the panel can be characterised as being mean-reverting. For these variables, we impose an AR prior δ_i, where δ_i is the estimated coefficient that results from regressing y_{it} on its first lag. The variables, transforms, and priors we use are displayed in appendix A.

3.2 Model specifications

We consider a range of competing VAR and BVAR specifications to assess how forecast performance changes as model size increases. All models we consider contain a subset of the 50 variables included in our largest model specification.

We consider three univariate baseline models; an AR(4) model AR, an AR model using the Schwartz-Bayesian Information Criteria to determine the lag length $AR^{SBC)}$, allowing lags to range from 1 to 4, and forecasts obtained from our priors alone $Priors$.

The BL model contains the 5 baseline variables used to determine the prior hyperparameters in each of our models; real GDP, tradable and non-tradable prices, 90-day interest rates and the trade-weighted exchange rate. We estimate three separate specifications of this model. These are a VAR with 4 lags BL, a VAR with Schwartz-Bayesian Information Criteria selected lag length BL^{SBC}, and also a VAR estimated using data-determined Bayesian priors, as in Del Negro and Schorfheide (2004) BL^{BVAR}.[11]

These models have been chosen to provide a range of baseline forecasts to compare our BVAR methodology against. However, it is also of interest how the forecast performance of the BVAR changes as the model size gets larger. To do this, we consider BVARs of 4 different sizes using the methodology laid out in section 2. The data included in each model is summarised in table 1.

The M model is similar to the medium-sized model used in Haug and Smith (2007).[12] The model has a domestic endogenous block containing the five BL variables m_1, and a foreign block m_2 containing world GDP, world CPI, and world 90-day interest rates. The domestic variables do not appear as right-hand-side variables in the foreign sector, but the foreign variables appear both in the foreign block and the domestic block.

The ML model is a variant of the large model used in Buckle *et al* (2007). This model differs from the Buckle *et al* (2007) model in three main respects. First, we express our model in levels, while Buckle *et al* (2007) specify their model in terms of deviations from trend. Second, we use slightly different data in our model. Specifically, our model splits the CPI into the tradable CPI and the non-tradable CPI; excludes climate; and includes exports of goods prices expressed in world prices instead of total export prices expressed in world prices. Third, our model imposes fewer restrictions on the variables entering each equation.[13]

In this model, the baseline domestic block m_1 is augmented with real exports, and the foreign block m_2 is augmented with goods export prices and import prices (both expressed in world prices), world equity prices, and oil prices. As in the M model, the variables in the foreign block m_2 enter both the foreign and domestic blocks m_1,m_2, while the foreign block is exogenous to the domestic block

[11] BL^{BVAR} is estimated using the Bayesian priors discussed in section 2. Following Del Negro and Schorfheide (2004), the hyperparameter λ is chosen to maximise the marginal data density using a grid search over a range of values of λ. The other hyperparameters τ and θ are set proportionately to λ, in the same proportions as used in the larger models. That is $0.1\tau = 0.01\theta = \lambda$.

[12] This model differs from the Haug and Smith (2007) model in that the CPI is split into the tradable CPI and the non-tradable CPI.

[13] For example, Buckle *et al* (2007) have four blocks of equations, and export and import prices are determined in a foreign block containing lags of export and import prices and world GDP. Our model, in contrast, determines all foreign variables endogenously within one block.

The LG^{35} model contains all of the variables in the ML model plus an extra 23 variables. The domestic block m_1 is augmented with real GDP data, housing market data, labour market data, and survey (business and consumer confidence) data. The foreign block m_2 is augmented with world 10-year interest rates. Together, the domestic and foreign blocks interact in the same way as in the ML model.

The LG^{50} variable model contains all of the variables in the LG^{35}, plus an additional 15 variables. The domestic block m_1 is augmented with additional real GDP data, labour market data, the current account balance, and additional money and financial market data.

4 Real-time forecasting exercise

A fundamental problem facing forecasters in real time, is that data is incrementally released throughout a quarter, so that not all series of interest are available when forecasts are to be made. In addition, data tends to be revised over time, making forecasting more difficult in real time than it may appear when looking at data ex post.

At the RBNZ, forecasts are made in preparation for the quarterly *Monetary Policy Statement (MPS)* at the end of the second month of each quarter. At this time, forecasters have available to them financial market and pricing data for the previous quarter, but only real activity data for the period two quarters previous. Table 2 summarises the data panel at each point of time.

We conduct a real-time out-of-sample forecasting exercise that allows for both the unbalanced nature of the data panel in real time, and also for revisions to data over time. To do this, we use the data that was available at the time that the initial forecasts for the *MPS* were made (the first-pass FP forecast).

In addition to containing historical data for all of the series of interest, this data set also contains forecasts for most variables in our panel.

The models we consider have relatively simple specifications for the world economy, so are unlikely to perform as well as judgemental forecasts made using a larger data set. For this reason, all of the BVAR forecasts are conditioned on the same exogenous assumptions for the world variables and oil

Table 1
Model specifications

Model	m_1	m_2	Number of Variables
BL	GDP Tradable CPI Non-tradable CPI 90-day rates Real exchange rate		5
M	BL	World GDP World CPI World 90-day interest rates	8
ML	M *Plus* Exports	M *Plus* Goods export prices Import prices Oil prices World equity prices	13
LG^{35}	ML *Plus* Migration Unemployment Wages GDP components Business surveys Consumer confidence Inflation expectations CPI House sales House prices	ML *Plus* World 10-year rates	35
LG^{50}	LG^{35} *Plus* Employment Labour force participation Current account balance Investment components Change in stocks Unskilled labour shortages Median days to sell house 5-year swap rate Monetary aggregates Real equity prices	LG^{35}	50

Table 2
Stylised data panel for different classes of variable

Time	Activity	Prices	Financial	Foreign
$t-2$	X	X	X	X
$t-1$	FP	X	X	X
t	FP	FP	FP	FP
$t+h$	O	O	O	FP

X indicates data that is available at each point of time, FP indicates forecasts that are applied from the FP forecast, and O indicates data that is missing from the panel.

prices as are used in FP.[14] Also, given informational advantages, the FP is likely to produce more accurate forecasts for the very near-term.[15] For this reason, the BVAR forecasts are conditioned on the same monitoring quarter forecasts as FP for the first two quarters of real GDP and its components, and the first quarter for pricing and financial market series.

We compare the forecasting performance of the models up to four quarters ahead over an out-of-sample period ranging from 1999Q4 to 2008Q3. At each point t in the out-of-sample evaluation period all parameters are re-estimated on all data and all conditioning assumptions. Forecast performance is evaluated using the variables contained in the BL model: real GDP, tradable CPI, non-tradable CPI, 90-day interest rates, and the trade-weighted exchange rate.

5 Results

The results of the forecasting exercise are shown in table 3. Each of the models we consider is compared against the baseline FP forecasts. This is a particularly tough benchmark, as the FP forecasts are produced as part of a rigorous forecasting process, and can be viewed as a judgementally adjusted model forecast.[16] Following Diebold and Mariano (1995), we test

[14] For world GDP, world CPI, and world interest rates the FP forecasts are largely based off *Consensus Forecasts*, while export prices, import prices and oil prices are forecast judgementally.

[15] See Matheson (2006) and Matheson (2007).

[16] These forecasts are produced using a large scale structural model, augmented with multiple smaller time-series models as well as considerable forecaster judgement.

the null hypothesis that model f and FP (denoted $f = 0$) have equal forecast accuracy on the basis of mean squared forecast error (MSFE) comparisons. Specifically, squared forecast errors are constructed over the evaluation period for each model, each variable, and each horizon:

$$\epsilon^f_{i,t+h} = (\hat{y}^f_{i,t+h} - y_{i,t+h})^2 \tag{19}$$

where $y_{i,t+h}$ is the ex-post variable at horizon h, $\hat{y}^f_{i,t+h}$ is the h-step-ahead forecast from model f, and $h = 1, \ldots, 4$. The squared forecast errors of the competing models and FP are then differenced $d_t = \epsilon^f_{i,t+h} - \epsilon^0_{i,t+h}$ to produce a sequence of squared forecast error differentials $\{d_t\}^T_{t=1}$, where $T = ((T_2 - 4) - T_1)$ and T_1 and T_2 are the first and last dates over which the out-of-sample forecasts are made, respectively. The mean difference in MSFEs is then tested by regressing the sequence of squared error differentials on a constant. A statistical difference in forecast accuracy between the competing models and the large BVAR is indicated by a constant that is statistically different from zero.[17]

[17] The variance of the coefficient estimate is adjusted for heteroskedasticity and autocorrelation using the Newey and West (1987) estimator with a truncation lag of $h-1$. The test statistic is compared to a Student's t distribution with $T-1$ degrees of freedom.

Table 3
Forecast results

			Univariate			*Multivariate*						
	h	$FP(RMSFE)$	AR	AR^{SBC}	$Priors$	BL	BL^{SBC}	BL^{BVAR}	M	ML	LG^{35}	LG^{50}
GDP	1	0.524	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	2	0.858	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	3	1.116	0.946	0.954	0.986	1.235*	1.116	0.911	0.861	0.871*	0.958	0.978
	4	1.375	0.936	0.934	0.932	1.762*	1.291	0.944	0.734	0.714*	0.850*	0.904
90 day rates	1	0.131	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	2	0.359	1.085	1.114	1.085	3.626*	1.808*	1.306*	1.294*	1.457*	1.304*	1.300*
	3	0.579	1.150	1.144	1.150	4.265*	1.850*	1.487*	1.364*	1.468*	1.331*	1.352*
	4	0.778	1.155	1.141	1.155	4.073*	1.770*	1.447*	1.266*	1.471*	1.351	1.385*
Tradables	1	0.376	1.015	1.015	1.000	1.080	1.022	1.000	1.003	1.035	1.017	1.007
	2	0.973	1.015	1.015	0.988	1.178*	0.956	1.032	0.988	1.000	0.969	0.970
	3	1.538	1.014	1.014	0.945	1.405*	0.966	1.032	0.975	0.965	0.954	0.959
	4	2.035	1.010	1.010	0.918	1.644*	0.988	1.044	0.965	0.932	0.927	0.941
Non-tradables	1	0.261	0.968	0.968	0.977	1.092	0.977	0.981	0.999	0.977	0.967	0.970
	2	0.407	0.972	0.972	0.977	1.336*	1.018	1.254	1.247*	1.063	0.960	0.941
	3	0.536	1.096*	1.097*	1.127	1.821*	1.040	1.562*	1.499*	1.180*	1.052	1.018
	4	0.617	1.236*	1.236*	1.321*	2.147*	0.980	1.733*	1.683*	1.303*	1.136	1.086
TWI	1	1.325	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	2	4.637	1.098	1.037	1.098	1.449*	1.129*	1.120*	1.045	1.064	1.032	1.040
	3	6.670	1.057	0.986	1.057	1.702*	1.112	1.135*	1.014	1.079	1.041	1.046
	4	8.570	1.007	0.924	1.007	1.830*	1.069	1.121	0.983	1.134	1.052	1.059

The numbers displayed are RMSFEs from models displayed in columns relative to the MSFEs from FP. A ratio greater (less) than one indicates a deterioration (improvement) relative to FP. $*$ denotes a significant difference in MSFEs at the 10 per cent level, according to the Diebold and Mariano (1995) test.

For most variables and horizons, the BVAR methodology performs roughly as well as the FP forecasts. In almost all cases, the BVAR performs better than the univariate and BL specification. However, there is no clear pattern to determine which specification of the BVAR performs best, with alternative specifications showing the best performance across different variables.

The BVAR methodology clearly produces better forecasting performance than univariate models or the BL specifications for GDP. In addition, the 4-step ahead forecasts for all specifications of the BVAR except LG^{50} significantly outperform FP. Over the sample we consider, the ML model produced the best forecast accuracy, and adding extra variables to this specification resulted in a deterioration in forecast accuracy.

All models struggle to match the forecasting accuracy of FP for 90-day interest rates, reflecting the informational advantage inherent in FP. In general, the BVAR models we consider do not forecast as well as univariate specifications. However, of the BVAR specifications, the best results are obtained from the M model.

The forecasting performance for tradable prices is relatively similar across all models, probably reflecting the inherent difficulty in forecasting this component. The LG^{35} model produces the lowest forecast errors of all the models, although these are not significantly different from the FP benchmark.

For non-tradable prices, the FP forecasts again provide a tough benchmark. However, there is a clear trend for larger models to perform better than smaller models. Indeed, in contrast to smaller models, the LG^{35} and LG^{50} models are not significantly different from FP at any horizon.

None of the models considered are able to outperform FP in forecasting the TWI exchange rate, and all specifications of the BVAR produce relatively similar forecasting performance.

Overall, the larger BVAR specifications tend to outperform univariate and smaller multivariate specifications. Moreover, these models broadly produce forecast accuracy that is comparable to the FP forecasts. In general, we find that the $LG^{3}5$ performs better than smaller or larger BVAR specifications.

6 Illustrating some useful tools for analysing the forecasts

In this section, we illustrate some tools for analysing our BVAR – alternative scenarios and semi-structural shock decompositions. Throughout, we employ the large BVAR with 35 variables, LG^{35}. For illustrative purposes, we examine forecasts made using the data used to compile the RBNZ's December 2006 MPS forecasts. This point in time is around the beginning of a sharp rise in world oil prices and commodity prices more generally.[18]

6.1 An alternative scenario

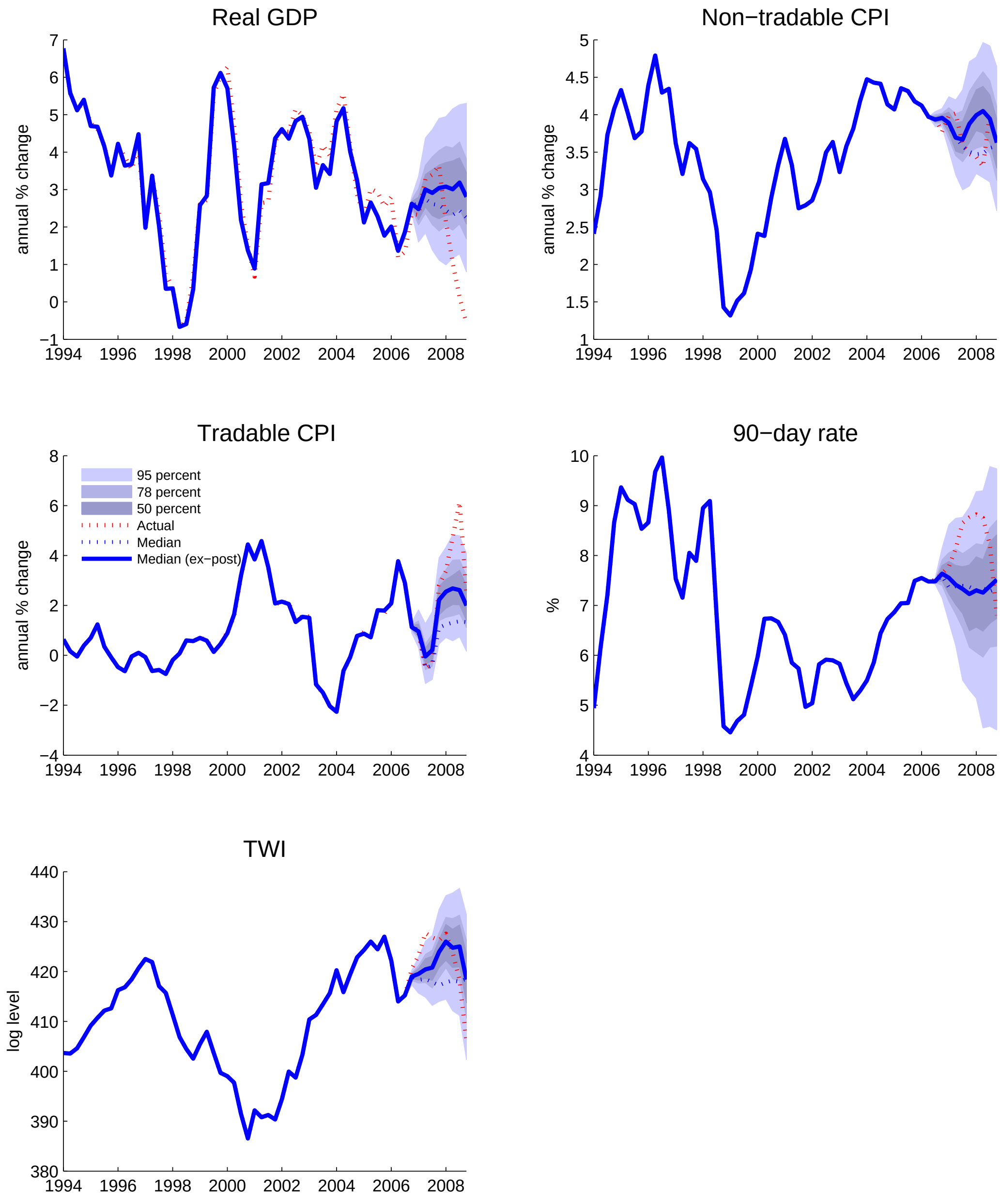
While the sharp rise in commodity prices over 2007 and the beginning of 2008 was very difficult (if not impossible) to predict in real time, we construct a counterfactual experiment in which we assume we know exactly what was going to happen to them. Implicitly, we assume the forecaster sees upside risk to commodity prices over the forecast horizon and wants to know the impact on the forecasts if that risk should eventuate. Alternative scenarios of this sort are routinely employed by central banks and other policymaking institutions. In this experiment, we use the baseline conditioning information discussed in section 4 plus conditioning information for oil prices and the world prices of New Zealand's exports and imports, assuming the forecaster knows the future paths of these variables a priori.

Figure 1 displays the forecasts for the baseline variables, along with the ex-post data that arrived after the publication of the MPS. In each panel of the figure, the dotted blue line is the ex-post data, and the red dotted line is the median forecast from our baseline model. The baseline model generally under-predicted all of the baseline variables. However, once we condition the forecasts on the future paths for commodity prices (the solid blue lines), they tend to improve. Moreover, the changes to the forecasts broadly seem to make economic sense.

The higher commodity prices boost the tradable inflation forecast and – supporting the view that the New Zealand dollar is classified as a 'commodity currency' – the TWI is now forecast to appreciate. The higher export prices outweigh the dampening effect of higher oil prices and real GDP growth rises,

[18] Between 2006Q4 and 2008Q3, oil prices rose by around 100 per cent, while the world price of New Zealand's exports rose by around 20 per cent.

Figure 1
Density forecasts, 2006Q4



putting upward pressure on non-tradable inflation further out in the forecast. Looking at the uncertainty around the forecasts, we find that ex-post data generally fall within the 95 percent interval derived from the conditional predictive density of the BVAR. The alternative scenario has surprisingly little impact on interest rates, which, according to the ex-post data, tightened substantially over the forecast period.

6.2 Semi-structural shock decompositions

VAR forecasts are generally accurate relative to structural models at shorter horizons. However, the 'black-box' nature of reduced-form VAR forecasts makes them difficult to communicate to policymakers. While structural identification can yield forecasts that have a structural interpretation, there is a wealth of different ways to identify a VAR, each with the potential to produce a conflicting economic story. Certainly, using the strongly recursive identification scheme described in section 2.3 to identify every shock in the system would require the researcher to determine whether or not each variable is impacted contemporaneously by every shock. Fortunately, an alternative, less structural approach to describing the data can be adopted in this framework.

Perhaps the most natural way to lend a structural interpretation to VAR forecasts is by way of a shock decomposition, similar to (17). The shock decomposition of the data up to time t can be written as:

$$Y_t = D + \sum_{j=1}^{t} \sum_{i=1}^{n} \sum_{k=m_{i-1}+1}^{m_i} M_{k,t-j} \epsilon_{k,j}, \qquad t = 1, \ldots, T+h \tag{20}$$

where the M_k is the matrix of impulse responses to shock k. Each variable can thus be expressed as the sum of contributions from each shock k from each of the n blocks of equations. As mentioned in sections 2.3 and 2.4, strong recursive blocks in the contemporaneous matrix yield forecasts that are invariant to the ordering of the variables within a given block of equations.[19] Moreover, as long as the ordering of the n blocks remains the same, the sums of the shock contributions from each block ($\sum_{k=m_{i-1}+1}^{m_i} M_k \epsilon_k$) are also invariant to the ordering of the variables within each block: this result also applies for forecast error variance decompositions. This allows us to remain agnostic on the ordering of variables within each block. Instead, we

[19] This is related to the recursive identification scheme discussed by Christiano *et al* (1999), who show that impulse responses to a particular shock are invariant to the ordering of the variables grouped above and below that shock.

can focus our attention on allocating variables into blocks and then ordering the blocks.

Recall that our large VAR only has two blocks of equations, the large endogenous domestic block and the (exogenous) foreign block. We can thus readily decompose each forecast into contributions from *all* foreign shocks and *all* domestic shocks without too much controversy regarding the ordering of the blocks (the foreign block is exogenous to the domestic block). Nonetheless, we can also take the decomposition a little further.

It is straightforward to decompose each of the foreign and domestic blocks to provide a more interpretable forecast, provided we are confident of our block recursive assumptions – the recursive ordering of the blocks.

It has become quite common to recursively identify small VARs with the following causal ordering in the contemporaneous matrix (see, for example, Zha 1999):

$$\text{Real activity} \rightarrow \text{Prices} \rightarrow \text{Financial}$$

We adopt this approach for our large BVAR by grouping the domestic variables into three groups: those variables classified as real economic activity indicators, those variables classified as price indicators, and those variables classified as financial variables. We denote this identification scheme B^* in appendix A. The results from this decomposition for our alternative scenario described in section 6.1 are displayed in figure 2, where the bars to the left (right) of the vertical lines are the real-time contributions to historical (future) paths for the variables. For each variable, the sum of the contributions equals the forecast less the deterministic components ($Y_t - D$ in equation 20).

We find that the foreign block of variables generally contribute a relatively large amount to our baseline variables, particularly over the forecast horizon.

To further improve our understanding of the drivers of the forecasts, we can also decompose the foreign block of equations in a similar way to the domestic block. Following Zha (1999), we assume that oil prices are exogenous. The causal ordering of the foreign block is:

$$\text{Oil prices} \rightarrow \text{Real activity} \rightarrow \text{Prices} \rightarrow \text{Financial}$$

We denote this identification scheme B^{**} in appendix A. The results of the decomposition of the total foreign contribution from figure 2 are displayed in figure 3.

Figure 2
Shock decompositions, 2006Q4

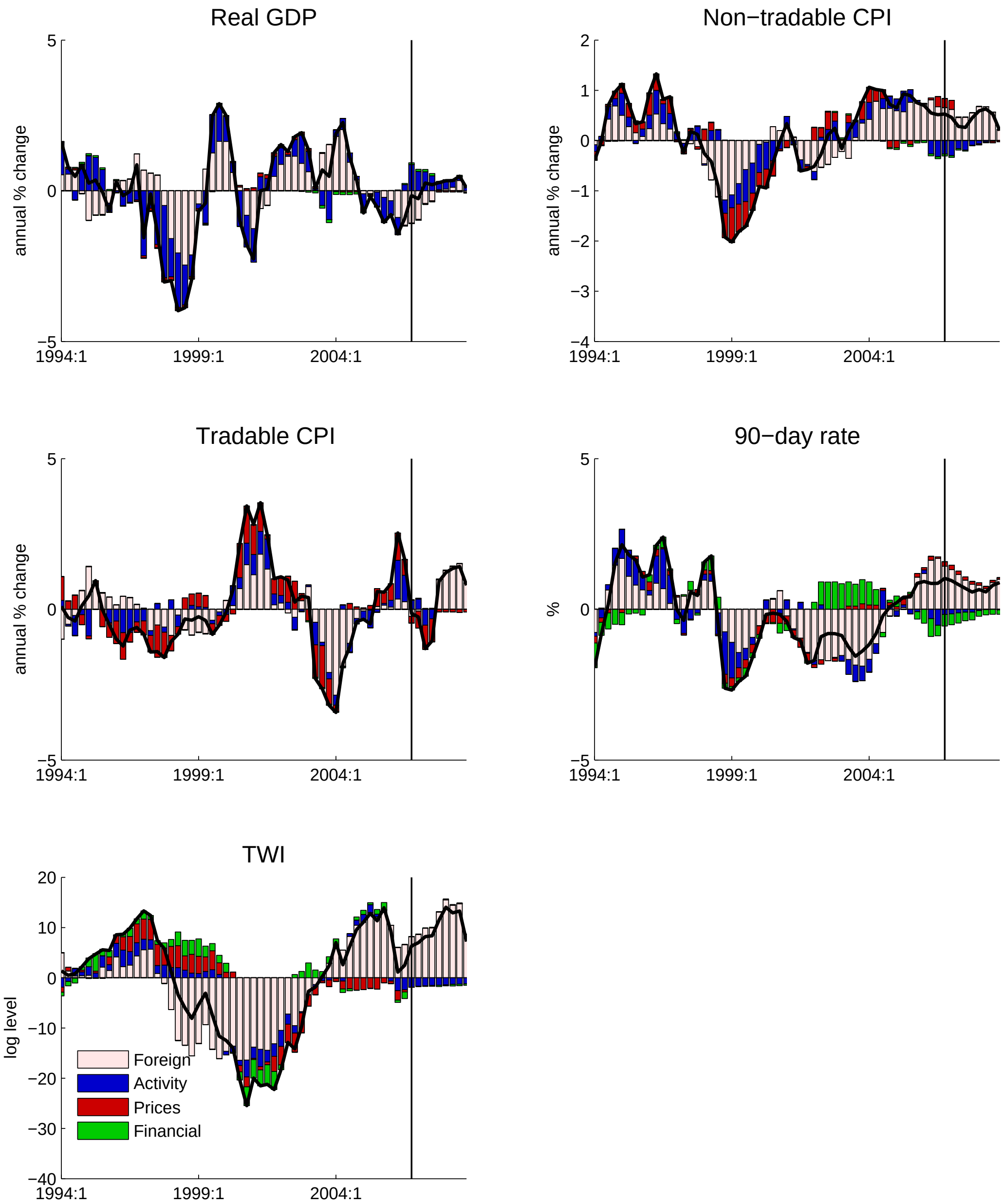


Figure 3
Decompositions of the foreign contribution, 2006Q4

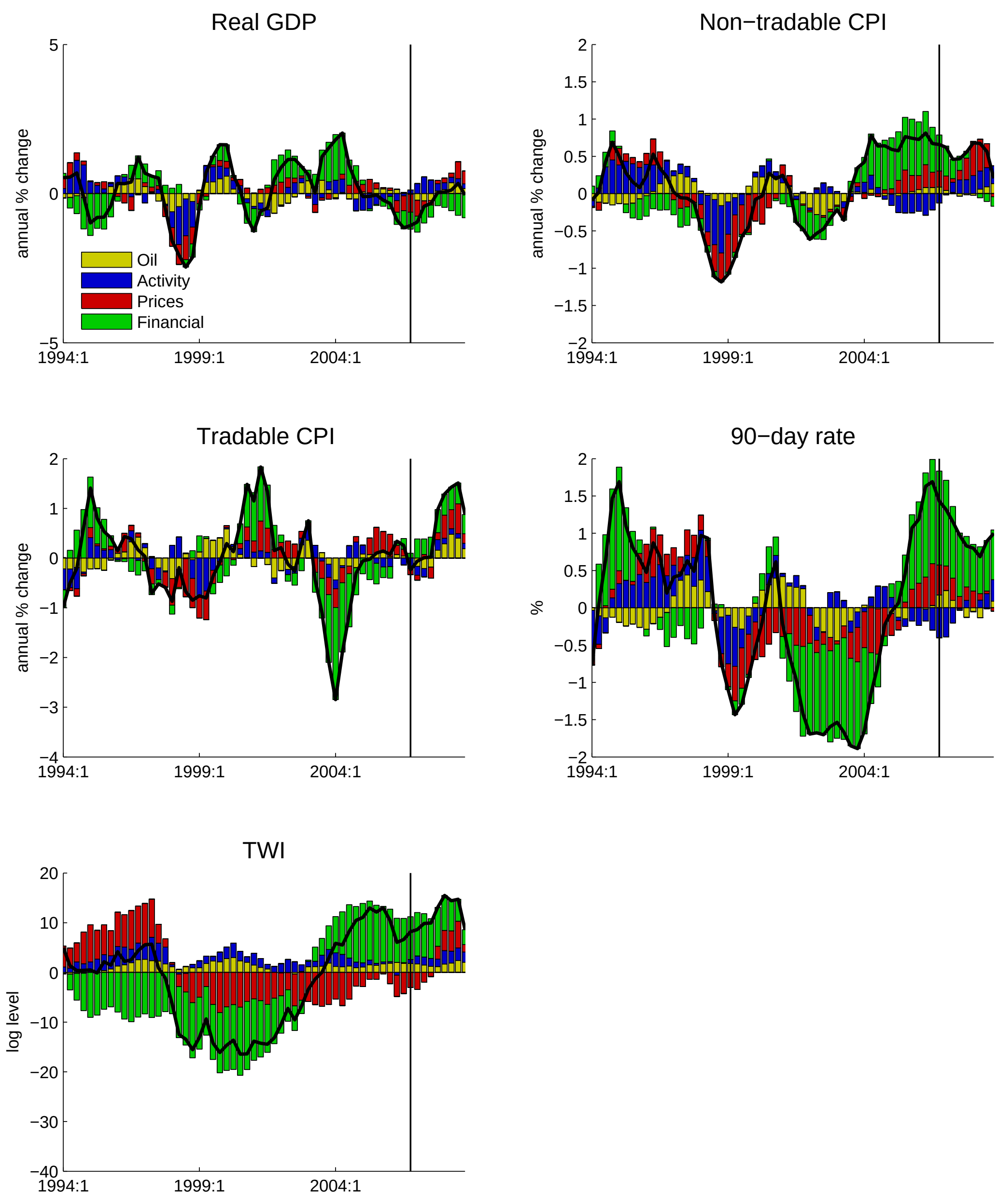


Table 4
Forecast error variance decomposition, 2006Q4

%		Foreign				Domestic		
	h	Oil	Activity	Prices	Financial	Activity	Prices	Financial
GDP	1	6.8	8.1	1.3	4.2	79.7	0.0	0.0
	2	5.1	12.2	4.1	10.0	68.6	0.0	0.0
	4	3.5	7.7	5.6	17.1	65.9	0.1	0.1
	8	2.3	5.7	5.1	21.2	65.4	0.2	0.2
Non-tradable CPI	1	0.7	7.8	12.3	1.3	42.1	35.8	0.0
	2	2.7	22.0	13.1	2.4	30.8	29.0	0.0
	4	6.1	27.4	14.7	4.8	24.4	22.4	0.0
	8	6.6	26.4	17.1	6.4	23.9	19.5	0.0
Tradable CPI	1	5.1	3.7	7.7	1.9	41.2	40.4	0.0
	2	5.5	1.6	12.6	10.7	33.2	36.3	0.0
	4	7.4	2.1	13.1	10.7	31.2	35.5	0.0
	8	9.4	1.6	10.4	13.1	30.8	34.7	0.0
90-day rate	1	0.7	4.4	3.0	7.4	50.1	4.3	30.1
	2	1.1	3.4	4.7	18.9	42.0	4.2	25.8
	4	3.3	6.0	4.1	23.7	37.7	3.7	21.5
	8	3.8	5.6	4.1	21.9	39.6	3.7	21.3
TWI	1	0.4	9.2	28.9	8.0	11.2	16.8	25.4
	2	0.8	7.0	26.1	23.5	9.1	14.6	18.9
	4	0.4	4.0	27.0	33.3	8.7	12.2	14.3
	8	0.9	2.7	24.4	41.3	8.4	10.5	11.8

We find that foreign financial shocks are relatively large contributors to the overall impact that the foreign block has on our TWI and the 90-day rate forecasts. Together, the foreign financial shocks act to tighten domestic monetary conditions over the forecast horizon, acting to reduce domestic GDP growth. Foreign price shocks, on the other hand, generally act to boost the tradable and non-tradable inflation forecasts.

More generally, we can also examine forecast error variance decompositions to find the *typical* contributions to the variance of the forecasts from each of our blocks of equations. We display such a decomposition in table 4. The variance decomposition results show that domestic activity shocks are relatively large contributors to most of the baseline forecasts. Domestic price shocks have a relatively large impact on the tradable and non-tradable CPI forecast variance, as well as on the TWI forecast variance. In fact, the contributions to the TWI forecast variance are intuitively appealing, with the largest contributions coming from domestic and foreign financial and price shocks. The contributions to the tradable forecast variance also have some appeal, with foreign price shocks (including oil) accounting for more than 20 percent of the forecast variance at longer horizons.

7 Conclusion

We examined the forecasting performance of BVARs of different sizes using an unbalanced panel of data. We found that our BVAR methodology outperformed univariate and VAR benchmarks in a real-time out-of-sample forecasting exercise, and produced comparable forecast accuracy to the judgementally adjusted model forecasts produced by the RBNZ. Our results on the optimal number of variables to include in the BVAR were inconclusive, but a 35 variable model tended to forecast better than larger or smaller models.

We demonstrated techniques for imposing judgement on the forecasts, and also for analysing shock decompositions without making strong identifying assumptions. These techniques allow a semi-structural interpretation of the forecasts to be made, and greatly aid in communicating the economic story underlying the forecasts to policy makers.

References

Banbura, M, D Gianonne, and L Reichlin (2008), "Large Bayesian VARs," *Journal of Applied Econometrics*, forthcoming.

Bloor, C and T D Matheson (2008), "Analysing shock transmission in a data-rich environment: A large BVAR for New Zealand," *Reserve Bank of New Zealand, Discussion paper*, 2008/09.

Buckle, R A, K Kim, H Kirkham, N McLellan, and J Sharma (2007), "A structural VAR business cycle model of a volatile small open economy," *Economic Modelling*, 24, 990–1017.

Christiano, L J, M Eichenbaum, and C Evans (1999), "Monetary policy shocks: What have we learned and to what end?" in *Handbook of Macroeconomics*, eds G S Taylor and M Woodford, vol 1, chap 2, 65–148, Elsevier.

Croushore, D (2006), "Forecasting with real-time macreconomic data," in *Handbook of economic Forecasting*, eds C W G Graham Elliott and A Timmerman.

Cushman, D O and T Zha (1997), "Identifying monetary policy in a small open economy under flexible exchange rates," *Journal of Monetary Economics*, 39, 433–48.

De Mol, C, D Gianonne, and L Reichlin (2008), "Forecasting using a large number of predictors: Is Bayesian regression a valid alternative to Principal Components?" *Journal of Econometrics*, 146(2), 318–328.

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

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 projections using realistic prior distributions," *Econometric Reviews*, 3, 1–100.

Gianonne, D, L Reichlin, and D Small (2008), "Nowcasting GDP and inflation: The real-time informational content of macroeconomic data releases," *Journal of Monetary Economics*, 55(4), 655–676.

Haug, A A and C Smith (2007), "Local linear impulse responses for a small open economy," *Reserve Bank of New Zealand, Discussion paper*, 2007/09.

Kadiyala, K R and S Karlsson (1997), "Numerical methods for estimation and inference in Bayesian VAR-models," .

Litterman, R (1986), "Forecasting with Bayesian vector autoregressions - five years of experience," *Journal of Business and Economic Sta-*

tistics, 4, 25–38.

Matheson, T D (2006), "Factor model forecasts for New Zealand," *International Journal of Central Banking*, June, 169–237.

Matheson, T D (2007), "An analysis of the informational content of New Zealand data releases: the importance of business opinion surveys," *Reserve Bank of New Zealand*, *Discussion paper*, 2007/13.

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

Robertson, J C and E W Tallman (1999), "Vector autoregressions: forecasting and reality," *Federal Reserve Bank of Atlanta*, *Economic Review*, 4–18.

Sims, C A (1993), "A nine-variable probabilistic macroeconomic forecasting model," in *Business Cycles, Indicators, and Forecasting*, eds J H Stock and M W Watson, University of Chicago Press, Chicago.

Sims, C A and T Zha (1998), "Bayesian methods for dynamic multivariate analysis," *International Economic Review*, 39(4), 949–968.

Waggoner, D F and T Zha (1999), "Conditional forecasts in dynamic multivariate models," *The Review of Economics and Statistics*, 81(4), 639–651.

Waggoner, D F and T Zha (2003), "A Gibbs sampler for structural vector autoregressions," *Journal of Economic Dynamics and Control*, 28, 349–366.

Zha, T (1999), "Block recursion and structural vector autoregressions," *Journal of Econometrics*, 90, 291–316.

Appendices

A Data

Key

$[T, P, B, B^*, B^{**}]$		[Transform, Prior, Block, Block, Block]
where:		
T	0 1	level log level (multiplied by 100)
P	1 $\delta(< 1)$	Random Walk Mean reverting
B	$[m_1, m_2]$	[Domestic, Foreign] Blocks used for forecasting
B^*	$[m_1, m_2, m_3, m_4]$	[Financial, Prices, Activity, Foreign] Blocks used for figure 2
B^{**}	$[(B^* - m_4), m_4, m_5, m_6, m_7]$	[(B^* - Foreign),Financial, Oil, Prices, Activity] Blocks used for foreign sector in figure 3
Models	$[1, 2, 3, 4, 5]$	[BL, M, ML, LG^{35}, LG^{50}] x indicates variable included in model

Identifier	Description	$[T, P, B, B^{*}, B^{**}]$	Models $[1, 2, 3, 4, 5]$
ASECT	Real equity price index (deflated with CPI)	[1, 1, 1, 1, 1]	[o, o, o, o, x]
MM3	Money aggregate – M3	[1, 1, 1, 1, 1]	[o, o, o, o, x]
MM1	Money aggregate – M1	[1, 1, 1, 1, 1]	[o, o, o, o, x]
RTWI	Trade-weighted exchange rate	[1, δ, 1, 1, 1]	[x, x, x, x, x]
R5YS	5-year swap rate	[0, δ, 1, 1, 1]	[x, x, x, x, x]
R90DAY	90 day bank bill rate	[0, δ, 1, 2, 2]	[o, o, o, x, x]
PQHPI	House price index (QVNZ)	[1, 1, 1, 2, 2]	[o, o, o, o, x]
EBEACN	Expected costs next quarter (QSBO)	[0, δ, 1, 2, 2]	[o, o, o, x, x]
EBEASPN	Expected selling price next quarter (QSBO)	[0, δ, 1, 2, 2]	[o, o, o, x, x]
ERCPI3	2 year ahead inflation expectation (RBNZ)	[0, δ, 1, 2, 2]	[o, o, o, x, x]
PCPIS	Headline CPI	[1, 1, 1, 2, 2]	[o, o, o, x, x]
PNT	Non-tradable CPI	[1, 1, 1, 2, 2]	[x, x, x, x, x]
PTR	Tradable CPI	[1, 1, 1, 2, 2]	[x, x, x, x, x]
LLISTOX	Labour cost index	[1, 1, 1, 2, 2]	[o, o, o, x, x]
AHDAYSAL	Median days to sell house (REINZ)	[0, 1, 1, 3, 3]	[o, o, o, o, x]
AHSALED	House sales (REINZ)	[1, 1, 1, 3, 3]	[o, o, o, x, x]
EBECU	Capacity Utilisation (QSBO)	[0, δ, 1, 3, 3]	[o, o, o, x, x]
EBEFLS	Difficulty finding skilled labour (QSBO)	[0, δ, 1, 3, 3]	[o, o, o, x, x]
EBEFLU	Difficulty finding unskilled labour (QSBO)	[0, δ, 1, 3, 3]	[o, o, o, o, x]
EWMC	Consumer confidence (WP-McDermott-Miller)	[0, δ, 1, 3, 3]	[o, o, o, x, x]
EBEDTAN	Domestic trading activity next quarter (QSBO)	[0, δ, 1, 3, 3]	[o, o, o, x, x]
EBEPRFN	Profit expectations next quarter (QSBO)	[0, δ, 1, 3, 3]	[o, o, o, x, x]
EBEGBO	Business confidence (QSBO)	[0, δ, 1, 3, 3]	[o, o, o, x, x]
NGDPPZ	Real GDP – Production	[1, 1, 1, 3, 3]	[x, x, x, x, x]
NCPZ	Real GDP – Consumption (private)	[1, 1, 1, 3, 3]	[o, o, o, x, x]
NVIZ	Real GDP – Change in stocks	[0, δ, 1, 3, 3]	[o, o, o, o, x]
NIPZ	Real GDP – Private investment	[1, 1, 1, 3, 3]	[o, o, o, x, x]
NIPDZ	Real GDP – Private investment (dwellings)	[1, 1, 1, 3, 3]	[o, o, o, x, x]
NITIAZ	Real GDP – Total investment (intangible assets)	[1, 1, 1, 3, 3]	[o, o, o, o, x]
NIMNRZ	Real GDP – Market investment (non-residential)	[1, 1, 1, 3, 3]	[o, o, o, o, x]

– continued from previous page

Identifier	Description	$[T, P, B, B^*, B^{**}]$	Models $[1, 2, 3, 4, 5]$
NIMTEZ	Real GDP – Market investment (transport)	[1, 1, 1, 3, 3]	[o, o, o, o, x]
NIMPZ	Real GDP – Market investment (plant and mach)	[1, 1, 1, 3, 3]	[o, o, o, o, x]
NMZ	Real GDP – Imports	[1, 1, 1, 3, 3]	[o, o, o, x, x]
NXZ	Real GDP – Exports	[1, 1, 1, 3, 3]	[o, o, x, x, x]
NIGZ	Real GDP – Investment (government)	[1, 1, 1, 3, 3]	[o, o, o, o, x]
NCGZ	Real GDP – Consumption (government)	[1, 1, 1, 3, 3]	[o, o, o, x, x]
TBC	Trade balance	[0, 1, 1, 3, 3]	[o, o, o, o, x]
LHURZ	Unemployment rate (HLFS)	[0, 1, 1, 3, 3]	[o, o, o, x, x]
LHPR	Participation rate (HLFS)	[0, 1, 1, 3, 3]	[o, o, o, o, x]
LHEMP	Employed (HLFS)	[1, 1, 1, 3, 3]	[o, o, o, o, x]
LMIGDZ	Migration (long-term departures)	[1, 1, 1, 3, 3]	[o, o, o, x, x]
LMIGAZ	Migration (long-term arrivals)	[1, 1, 1, 3, 3]	[o, o, o, x, x]
IEQWLDM	Real world equity price (deflated with world CPI)	[1, 1, 2, 4, 4]	[o, o, x, x, x]
RNLROW	World 10 year bond rate (80-20 split, US and AU)	[0, 1, 2, 4, 4]	[o, o, o, x, x]
RNROW	World 90 day rate (80-20 split, US and AU)	[0, 1, 2, 4, 4]	[o, x, x, x, x]
IOILP	World oil price (Dubai: US dollars)	[1, 1, 2, 4, 5]	[o, o, x, x, x]
TITOTPWZ	World price of NZ exports	[1, 1, 2, 4, 6]	[o, o, x, x, x]
TETOTPWZ	World price of NZ imports	[1, 1, 2, 4, 6]	[o, o, x, x, x]
IWCPI	World CPI (5 country weighted average)	[1, 1, 2, 4, 6]	[o, x, x, x, x]
IWGDPZ	World GDP (5 country weighted average)	[1, 1, 2, 4, 7]	[o, x, x, x, x]