A Bayesian VAR Analysis of Fiscal Policy in New Zealand

By Thomas Li *

This paper investigates the dynamic impacts of fiscal policy in New Zealand using Bayesian vector autoregressions (BVAR) on the latest available data. We assess both the forecasting abilities and structural results obtained from a VAR estimated on quarterly data of net tax revenue, government spending, and GDP. Our analysis demonstrates that BVARs, enhanced with appropriately selected hyperparameters, deliver more accurate out-of-sample forecasts compared to traditional VARs. Our structural VAR models reveal that fiscal decisions—tax increases and government spending—yield modest, short-lived stimulative effects on economic output. This study contributes to an understanding of the role and impact of fiscal policy in New Zealand, providing critical insights for policymakers.

I. Introduction

Over the past forty years, New Zealand has implemented a series of monetary and fiscal reforms that have led some to herald the country as a model of neoliberal economic theory (Kelsey, 2015). Prior to the 1980s, New Zealand's economy was tightly regulated with high tax rates and ever-increasing debt due to large fiscal deficits. With the 1984 election of the Labour Party, major reforms were enacted that sought to loosen governmental control of the economy in favor of deregulation and market liberalization. In particular, the Reserve Bank instituted a new regime of inflation targeting while the central government committed itself to fiscal restraint and balanced budgets.

While the early and long-term success of these free market reforms are up for debate (Kelsey, 2015), New Zealand's adoption of a (at the time) novel fiscal and monetary framework pose interesting questions for economic researchers. Already, a wealth of literature has sought to model and characterize the New Zealand economy (Claus et al., 2006; Parkyn and Vehbi, 2014; Hamer-Adams, Wong et al., 2018). We build upon this literature by assessing the dynamic impacts of fiscal policy in New Zealand using the most up-to-date data. By contributing to the understanding of the magnitudes, timing, and channels of fiscal policy on the nation's economy, we seek to aid policy makers in forming economically prudent fiscal policy decisions.

Vector autoregressions (VARs) are particularly useful instruments in this pursuit. Beyond their use as a forecasting tool, VARs allow economists to unpack structural dynamics when further interpretations are placed on residuals (this approach is known as a structural VAR). For fiscal analyses, Blanchard and Perotti (2002) (whom we refer to as BP) popularized a framework for assessing the dynamic behavior of fiscal shocks. The addition of Bayesian techniques allow us to improve model estimation and account for both forecast

* University of Pennsylvania (email: thomli@sas.upenn.edu).

and parameter uncertainty when offering results. In this paper, we seek to combine all these methodologies in analyzing the structure of fiscal policy in New Zealand.

The remainder of the paper is organized as follows. Section 2 discusses model specifications and provides an overview of the data used in our analysis. Section 3 compares the forecasting performance of a variety of VAR specifications and offers a short-term forecast of per capita GDP. Section 4 performs a structural VAR analysis uncovering the dynamic impacts of fiscal policy decisions on the economy through impulse response functions. Section 5 concludes.

II. Model and Data

A. *Model Specifications*

This paper primarily estimates the following baseline three-variable vector

$$Y_t := \begin{bmatrix} TAX_t & GOV_t & GDP_t \end{bmatrix}^\intercal$$

where TAX_t, GOV_t, GDP_t respectively denote logged real per capita values of net tax revenue, governmental spending, and gross domestic product. This baseline specification follows that of Blanchard and Perotti (2002) and allows us to trace the dynamics of fiscal policy (as encapsulated by taxes and government spending) on economic output (as measured by GDP). We also consider a larger five-variable vector given by

$$Y_t^L := \begin{bmatrix} TAX_t & GOV_t & GDP_t & INFL_t & INT_t \end{bmatrix}^\intercal$$

that adds the inflation rate $INFL_t$ and 10-year nominal interest rate INT_t to the baseline VAR. This larger specification is the same employed by Perotti (2005).

Following the literature, net tax is defined to be total tax revenues less subsidies and transfers, and government spending is final consumption expenditures on goods and services by the general government (i.e., central and local governments). The three variables in the baseline specification enter the VAR in logged real per capita terms and $INFL_t, INT_t$ in percentage points. All variables are at a quarterly frequency.

B. *Data*

We obtain data from Statistics New Zealand (the governmental statistical agency), the Treasury, and the OECD. Statistics New Zealand reports their data in real 2010 New Zealand Dollars (NZD) after adjusting for seasonal and calendar trends. Net tax revenues are calculated to be total tax revenues less transfers and subsidies, and are deflated by the GDP deflator (indexed to 100=2010). Log and per capita transformations are applied by the authors, using quarterly population data. The common date range of all variables spans 1987Q2 to 2023Q4. Further details on the data can be found in Table 1 and the logged series are plotted in Figure 1.

The plot of the baseline variables in the form they enter the VAR is provided in Figure 1. All three variables exhibit an evident upward time trend, suggesting that the time series

Table 1—Raw Data

Indicator	Source	Description
Gross Domestic Product	Statistics NZ	Seasonally adjusted, quarterly, in real NZD (2010), not annualized
Government Final Consumption Expenditure	Statistics NZ	Seasonally adjusted, quarterly, in real NZD (2010), not annualized
Total Tax Revenue	The Treasury	Monthly (aggregated to quarterly by authors), nominal, not annualized
Total Transfers & Subsidies	The Treasury	Quarterly, nominal, not annualized
Inflation Rate	Statistics NZ	Annual percentage change in CPI
10-Year Nominal Interest Rate	OECD	Not seasonally adjusted, quarterly, in percentage points, average 10-year government bond yields
GDP Deflator	Statistics NZ	Quarterly, 100=2015 (reindexed to 100=2010 by authors)
Total Population	Statistics NZ, World Bank	Quarterly from 1991Q1– (Statistics NZ), pre-1991Q1 values imputed from yearly values (World Bank)

Note: Table 1 presents the sources and additional details for the raw data on each indicator. The data file can be found in the replication materials.

are non-stationary (i.e. the mean changes over time). A formal test for stationarity using the augmented Dickey-Fuller test for unit root is presented in Appendix Table A1, which shows statistical evidence that all series except the inflation rate exhibit unit roots. We account for this nonstationarity by incorporating a deterministic time trend, as done by Blanchard and Perotti (2002) and Sims, Stock and Watson (1990). Details on allowing for deterministic time trends are presented in Section 3.

Further descriptive figures of these time series are presented in Figures 2–4, the features of which we now discuss alongside a few key moments in recent New Zealand economic history.

Figure 2 plots the ratio of government spending and net taxes to GDP. Prior to 2005, government spending is higher than net taxes, adding to the country's debt. Following the enactment of the Fiscal Responsibility Act in 1994—which increased accountability and standards on fiscal policy—government spending as a fraction of GDP began to decrease as

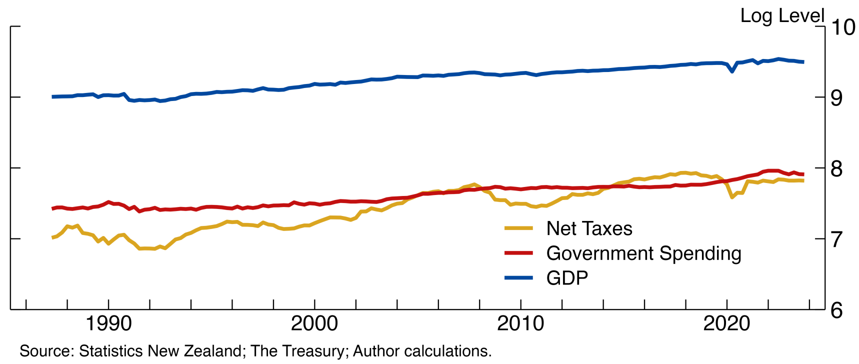

Figure 1. Variables Used in Baseline Specification

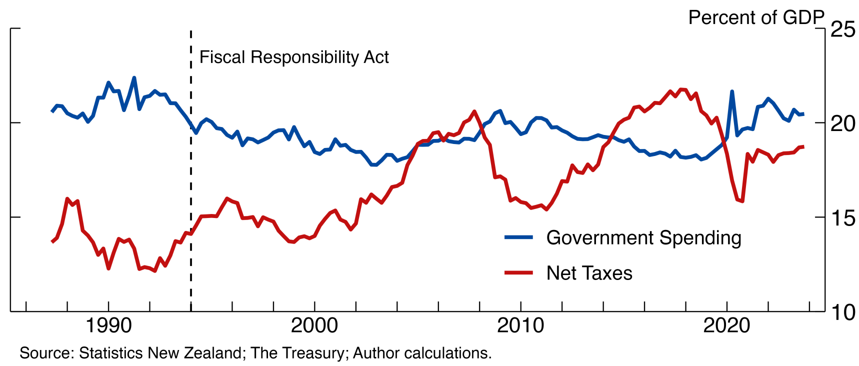

Figure 2. Ratios of Governmental Spending and Net Taxes to GDP

net taxes/GDP increased, thereby reducing the yearly budget deficit. From 2005 onwards, spending and net taxes have more or less fluctuated around the same level.

Inflation and interests rates are displayed in Figure 3. Of note are the high levels of inflation before 1992. Inflation subsided after 1990, when the Reserve Bank of New Zealand began implementing an inflation targeting monetary regime[1], and has remained roughly around 0–4% since (with an average of 2% from 1992 to 2020). Currently, New Zealand—like many economies around the world—is encountering a period of high inflationary pressures following the COVID-19 pandemic.

Real per capita GDP growth, as displayed in Figure 4, has been generally positive over the range of our sample, with a few noticeable downturns in 1991–1992 (a self-inflicted recession), 1997–1999 (Asian Financial Crisis and drought), 2008 (Global Financial Crisis), and 2020 (COVID-19). Outside these recessions, New Zealand experienced strong economic growth from 1993–1996 and 1999–2007. Over the course of the sample (1987Q2 to 2023Q4),

[1] In fact, New Zealand pioneered the inflation targeting regime that has been widely adopted in central banks around the world.

the average annualized per capita GDP growth rate is 1.35%.

Overall, our sample generally falls under stable monetary and fiscal regimes, as the Fiscal Accountability Act and inflation targeting framework began early in our sample. For our purposes, this means we have some degree of confidence that the same stochastic process underlies the entirety of our data and therefore do not need to account for additional nonstationarity beyond time trends.

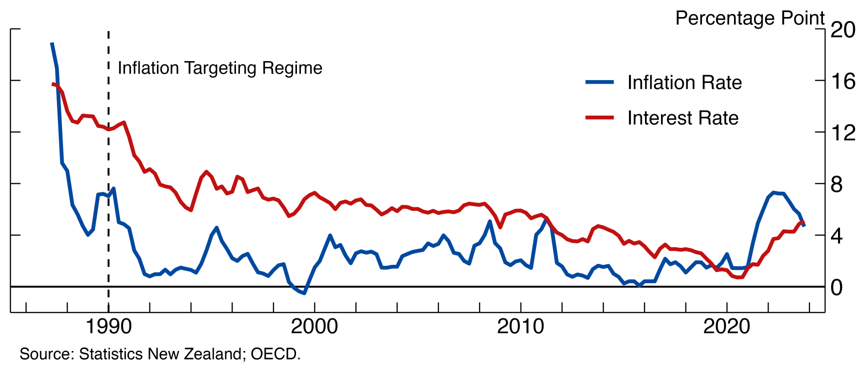

Figure 3. Annualized Inflation Rate and Nominal Interest Rate

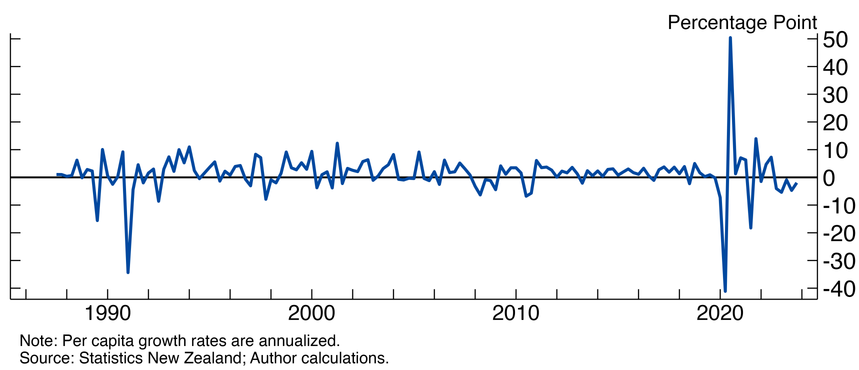

Figure 4. Annualized Real Per Capita GDP Growth Rate

III. Forecasting

In this section, we construct estimates of the two specifications via two different estimation procedures. Namely, we produce a flat-prior VAR estimated by OLS (VAR) and a Bayesian VAR using a Minnesota prior with hyperparameter selection (BVAR). For each model, we use recursive estimation to assess pseudo out-of-sample forecasting performance in a manner akin to Giannone, Lenza and Primiceri (2015) (which we hereinafter refer to as GLP). We conclude the section by providing a short-term economic outlook for the next three years.

A. Methodology

We consider and compare two estimation procedures: a flat-prior VAR estimated by OLS and a BVAR with a Minnesota prior chosen with hyperparameter selection. A flat-prior VAR with a prior proportional to $|\Sigma|^{-(n+1)/2}$ yields a posterior coefficient distribution centered at the OLS estimator (Giannone, Lenza and Primiceri, 2015)—we will henceforth refer to this model as the "VAR." The BVAR uses a Minnesota prior as first discussed by Doan, Litterman and Sims (1984) and later refined by Sims and Zha (1998).

To account for unit roots in the time series, we incorporate a deterministic trend for both estimation procedures. For OLS, a linear function of time is included in the model. For the BVAR, we employ a two-step procedure wherein we estimate the constant and time trend via OLS: $y_t = \gamma_0 + \gamma_1 t + \hat{y}_t$ and perform the Bayesian estimation on the detrended residuals $\hat{y}_t = y_t - \gamma_0 - \gamma_1 t$.

Previous literature has documented the gains from properly configuring the prior covariance matrix when performing Bayesian VAR estimation (Giannone, Lenza and Primiceri, 2015). In particular, a vector of hyperparameters $\lambda = \begin{bmatrix}\lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 & \lambda_5\end{bmatrix}$ controls the amount of shrinkage and thus the degree of informativeness in the prior. We perform hyperparameter selection of the vector λ by finding the parameters that maximize the log marginal data density (MDD), as used in Del Negro and Schorfheide (2004). To lessen the computational burden when computing the hyperparameter at each recursive iteration, we follow Schorfheide and Song (2015) by performing a one-time grid search on the initial recursive sample (1987Q2 to 2000Q1) to find suitable values of $\lambda_3, \lambda_4, \lambda_5$, which we fix at $\lambda_3 = 1$ and $\lambda_4 = \lambda_5 = 9$. In subsequent iterations, we only consider a range of values for λ_1 and λ_2. All models are estimated with four lags, as varying lag lengths do not result in noticeable changes in MDD or RMSFE (see Appendix Table A2).

To evaluate the forecasting capabilities of our models, we employ recursive estimation and forecasting starting with a estimation sample of 1987Q2 to 2000Q1 and expanding the forecast origin by one quarter at each iteration up to 2018Q4. We avoid forecasting past 2020 due to outlier fluctuations from COVID-19, which heavily bias coefficient estimates and forecast errors. For each origin T_0, we select the hyperparameters $\hat{\lambda}$ maximizing the log MDD and sample 1000 draws of model parameters (Φ, Σ) from the estimated posterior $p(\Phi, \Sigma \mid Y, \hat{\lambda})$. Using each draw of (Φ, Σ), we produce trajectories $Y_{T_0+1:T_0+4|1:T_0}$ of the next four quarters. An h-quarter-ahead point forecast is then taken to be the mean h-quarter-ahead forecast of the 1000 forecasts produced and quantiles of the forecasts are used to generate interval and density forecasts. The code to replicate these produced can be found in the supplementary zip file.

B. Point Forecast Evaluation

As a measure of model validity, we evaluate the pseudo out-of-sample forecasting performance of our models via the recursive estimation procedure described above. In particular, we provide root mean squared forecast errors (RSMFE) for each baseline variable in all the models considered (Table 2) as well as visual charts of the recursive forecasts of real GDP growth compared to the observed values (Figures 5 and 6).

Following GLP, our evaluation target is

$$z_{i,t+h} \coloneqq \frac{400}{h}(y_{i,t+h} - y_{i,t}).$$

Given that our baseline variables are expressed in log levels, $z_{i,t+h}$ approximates the annualized average growth rate over the next h quarters. The accuracy—as assessed by RMSFE—of the point forecasts of this measure are reported in Table 2 for each variable under each model. Since our first log-level forecast is of 2000Q2, taking differences means our first growth rate forecast is of 2000Q3.

Table 2—RMSFE of Point Forecasts

		Baseline		Large	
Horizon	Variable	VAR	BVAR	VAR	BVAR
One Quarter	Net Taxes	16.991	17.007	17.814	17.274
	Gov't Spending	6.307	5.041	6.393	4.891
	GDP	4.547	4.389	4.523	4.155
Four Quarters	Net Taxes	9.003	8.255	9.866	8.273
	Gov't Spending	3.138	1.844	3.010	1.841
	GDP	2.003	1.473	1.765	1.268

Note: Table 2 reports root mean squared forecast errors (RMSFE) for h-quarter-ahead annualized average growth rate of the three baseline variables, organized by VAR specification and estimation procedure. The evaluation sample used for recursive forecasting is 2000Q3 to 2019Q4.

We first compare the baseline and large specifications. Fixing the estimation procedure, the larger model generally produces smaller RMSFEs for government spending and GDP but performs worse on net taxes. This result may be attributed to the predictive power of inflation and interest rates for spending and GDP, which aligns with common economic thinking. GLP similarly found that larger models may improve performance for BVARs when aided by hyperparameter selection. Where our results differ, however, is in the improvement for the flat-prior VAR as well, as GLP suggest that the increase in estimation error can offset the gains from larger information sets. In our case, the improvement from the predictive power of inflation and interest rates may have overcome the additional estimation error.

Turning to the comparison between VAR and BVAR estimation procedures, we find that the BVAR consistently performs better than the flat-prior VAR in all but one case. This improvement suggests that log MDD-based choices of hyperparameters can lead to more informative priors that induce improvements to the estimation.

Overall, these RMSFE values are in line with other work on VAR forecasting and the four-quarter-ahead forecasts are well within the standard deviations of the differenced time series data (TAX: 13.75, GOV: 5.97, GDP: 5.11). The magnitude of the forecast errors suggest, however, that the VAR forecasts are not of much use to policy makers. Take the

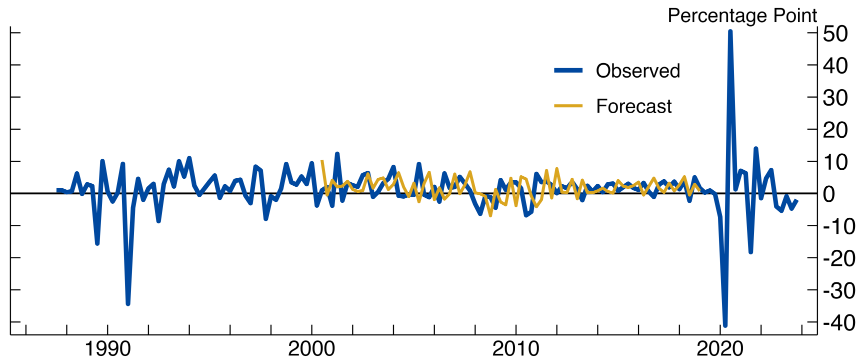

Figure 5. Forecasted vs Observed Real GDP Growth: VAR, 1-quarter-ahead

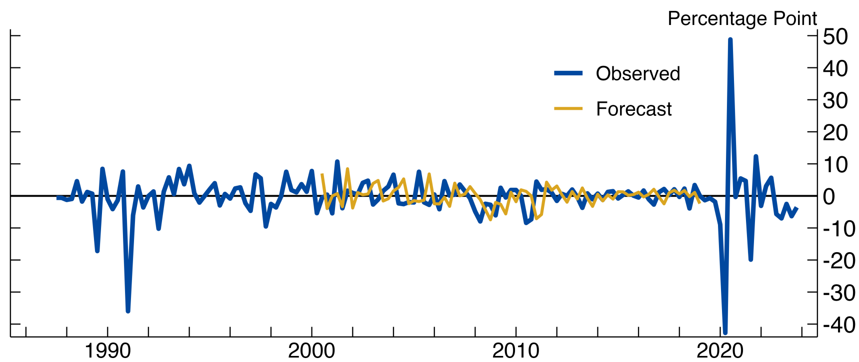

Figure 6. Forecasted vs Observed Real GDP Growth: BVAR, 1-quarter-ahead

Note: Figures 5 and 6 compare recursive 1-quarter-ahead forecasts against realized values of real per capita GDP growth from 2000Q3 to 2019Q4. Figure 5 shows the VAR forecasts and Figure 6 shows the BVAR forecasts. Both are estimated using the large specification.

baseline BVAR one-quarter-ahead GDP, for example: a RMSFE of 4.389 suggests that the forecast is off by around 4.4 percentage points on average. The four-quarter-ahead forecast is better at an average 1.473 percentage point discrepancy but is still uninformative for policymakers given that per capita GDP growth is typically around 1-2%. The forecasts for changes in net taxes are noticeably worse with errors upwards of 18 percentage points. A glance at the logged tax series in Figure 1 suggests that net taxes exhibit many fluctuations, which can be hard to predict.

Figures 5 and 6 give some insight into the nature of the forecasts. The point forecasts almost seem like a shift of the observed forecast origin forward by one quarter, with some slight attenuation toward the mean growth rate. Intuitively, point forecasts struggle to predict random fluctuations and shocks to the economy that drive GDP growth rates. For example, the models could not predict the large negative shock of the Great Financial Crisis and hence achieves a large RMSFE in 2008. As such, interval and density forecasts may provide a better picture by accounting for this uncertainty in point predictions.

C. Interval Forecast Evaluation

Interval forecasts provide a different perspective on forecasts that incorporates uncertainty in model parameters (as in the case of BVARs) and point predictions generated by the model (as in the case of both BVARs and VARs). To evaluate the interval forecast performance of the models, we provide the hit rates for 90% credible intervals generated by our models (Table 3). These credible intervals and the subsequent hit rates are generated for log-levels of the baseline variables.

Table 3—Hit Rates of Interval Forecasts

		Baseline		Large	
Horizon	Variable	VAR	BVAR	VAR	BVAR
One Quarter	Net Taxes	0.763	0.868	0.697	0.842
	Gov't Spending	0.868	0.974	0.842	0.974
	GDP	0.961	0.974	0.947	0.974
Four Quarters	Net Taxes	0.382	0.579	0.342	0.632
	Gov't Spending	0.553	0.934	0.566	0.934
	GDP	0.763	0.947	0.776	0.987

Note: Table 3 reports average hit rates for log-levels of the three baseline variables for estimated 90% credible intervals, organized by VAR specification and estimation procedure. The evaluation sample used for recursive forecasting is 2000Q2 to 2019Q4.

The BVAR interval forecasts perform better on achieving a 90% coverage frequency compared to the VAR. Most of the hit rates for the VAR interval forecasts are below the desired 90% whereas most of the hit rates for the BVAR forecasts are within a few percentage points of 90% (with the exception of four-quarter-ahead net taxes). These results are robust across model specification size and forecast horizon. Given that the BVAR takes into account model parameter uncertainty by sampling parameters $(\Phi.\Sigma)$ from a posterior parameter distribution on top of forecasting uncertainty, these results are unsurprising. Without incorporating parameter uncertainty, the VAR produces too narrow of an interval forecast, as evidenced by its deficient hit rates.

We further verify that the hit sequence for the h-step-ahead forecast has no autocorrelation beyond order $h-1$. Taken together, the approximately 90% hit rate and lack of serial correlation in the hit sequence suggests that the BVARs have produced well-calibrated interval forecasts that are not dynamically misspecified.

D. Short-term Economic Outlook

Based on our analysis of the forecasting performance of the models considered, the BVAR outperforms the VAR when it comes to both point and density forecasts. The addition of inflation and interest rates also appears to add some additional predictive power to our forecasts of GDP growth. As such, we conclude the forecasting section of the paper by providing a density forecast of New Zealand's per capita GDP over the next three years using the large BVAR fitted to the data from 1987Q2 to 2023Q4.

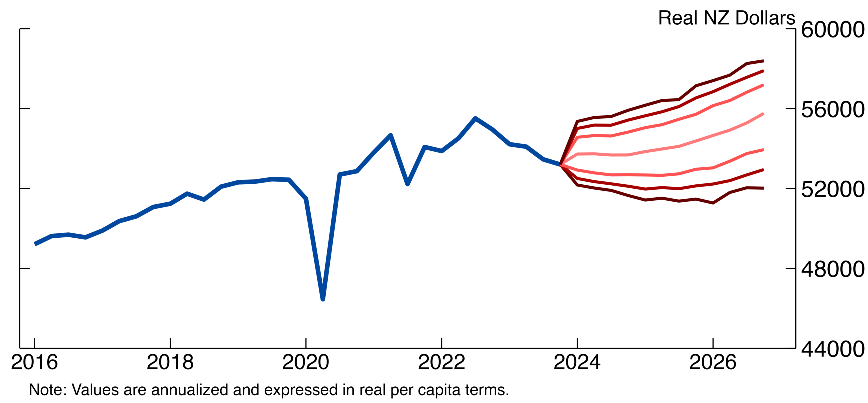

Figure 7. 12-quarter-ahead Density Forecast of Real GDP Per Capita

Note: Figure 7 displays a density forecast of real GDP per capita over the next 12 quarters, as estimated by the large BVAR specification over the data from 1987Q2 to 2023Q4. The red shadings respectively denote from darkest (outer) to lightest (inner), 90%, 80%, 60%, and median forecasts.

The median forecast presumes economic growth over the next three years at a rate similar to the pre-pandemic trend (we project 1.57% growth per year, compared to the historical average of 1.47%). There are some risks of continued economic contraction over the next few years, as evidenced by the lower bands of the density forecast, but also upside risk of strong economic growth. By 2026Q4, the worst-case scenario suggested by the lower 90% interval forecast is a -2.25% decline in real GDP per capita from 2023Q4 (an average -0.75% per year). High inflation may be weighing down on the growth projections. The best-case scenario, on the other hard, is a 9.30% rise in real GDP per capita (an average 3.10% per year). Thus, with 90% probability, real GDP per capita in 2026Q4 will be between NZ$52018 and NZ$58391.

IV. Structural Analysis

We now turn to a structural VAR (SVAR) model of New Zealand fiscal policy. In particular, we seek to assess the dynamic effects of unexpected shocks in government spending and taxes on economic output. We begin with a discussion of our identification assumptions, then provide impulse response functions (IRFs), and conclude the section with an analysis of fiscal policy dynamics in New Zealand.

A. Identification

For simplicity and following the example of BP, we consider the baseline VAR specification containing real per capita net tax revenue, government spending, and real GDP (ordered as such). We conduct our structural analysis by estimating a structural BVAR on the full pre-COVID sample (1987Q2 to 2019Q4) with a Minnesota prior and hyperparameter selection. With the estimated posterior, we generate 1000 draws of model parameters (Φ, Σ) and compute an IRF trajectory for each draw. From these 1000 trajectories, we derive a mean IRF forecast as well as a 90% credible band by taking the 5th and 95th quantiles.

We employ a recursive identification scheme given by $Y_t = X\Phi + \Sigma_{tr}\varepsilon_t$, or in its expanded form,

$$\begin{bmatrix} TAX_t \\ GOV_t \\ GDP_t \end{bmatrix} = X\Phi + \begin{bmatrix} \sigma_{11}^{tr} & 0 & 0 \\ \sigma_{21}^{tr} & \sigma_{22}^{tr} & 0 \\ \sigma_{31}^{tr} & \sigma_{32}^{tr} & \sigma_{33}^{tr} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{TAX} \\ \varepsilon_t^{GOV} \\ \varepsilon_t^{GDP} \end{bmatrix}, \qquad \varepsilon \sim \mathcal{N}(0, I)$$

where X stacks the lags, Φ is the matrix of coefficients, and Σ_{tr} is the lower triangular Cholesky decomposition of the covariance matrix Σ of the forecast residuals u_t (i.e., $\Sigma = \mathbb{E}[u_t u_t^\intercal]$).

The recursive scheme states that GDP output is contemporaneously affected by all shocks, a plausible economic assumption. Given lags in data reporting and decision making, tax and government spending decisions are assumed to be unaffected by contemporaneous output shocks.

Our identification also assumes that taxes only respond contemporaneously to tax shocks whereas government spending responds contemporaneously to both tax and spending shocks. Thus, the potentially unfounded assumption that taxes move first is baked into the identification. This choice of ordering is relatively arbitrary, but we note that BP likewise choose to assume that tax decisions come before government spending decisions. Regardless, we present the alternative ordering in the Appendix as a robustness check and indeed find no difference in the results when assuming that spending decisions come before taxes.

The subsequent analysis in this section will use the recursive identification scheme, but we also considered[2] an alternative identification scheme proposed by BP that is popular in the literature for fiscal SVARs. Indeed, the BP identification scheme has been used in other fiscal SVAR analyses of New Zealand (Claus et al., 2006; Parkyn and Vehbi, 2014; Hamer-Adams, Wong et al., 2018). This identification better isolates the effect of *discretionary* fiscal policy shocks on the economy by accounting for endogenous variations that occur from automatic stabilizers responding to cyclical economic conditions. To do so, they use the following identification assumption:

$$Au_t := \begin{bmatrix} 1 & 0 & -a_1 \\ 0 & 1 & -b_1 \\ -c_1 & -c_2 & 1 \end{bmatrix} \begin{bmatrix} u_t^{TAX} \\ u_t^{GOV} \\ u_t^{GDP} \end{bmatrix} = \begin{bmatrix} 1 & a_2 & 0 \\ b_2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_t^{TAX} \\ \varepsilon_t^{GOV} \\ \varepsilon_t^{GDP} \end{bmatrix} =: B\varepsilon_t$$

[2]Unfortunately, the weather was too nice so I gave up trying to implement the BP identification scheme. It boils down to solving the system of equation $\Sigma = A^{-1}B\Sigma_\varepsilon B^\intercal(A^{-1})^\intercal$, which has 6 knowns and 6 unknowns after imposing the restrictions discussed in Blanchard and Perotti (2002).

where

$$\varepsilon \sim \mathcal{N}\left(\begin{bmatrix}0\\0\\0\end{bmatrix}, \begin{bmatrix}\sigma^2_{TAX} & 0 & 0\\ 0 & \sigma^2_{GOV} & 0\\ 0 & 0 & \sigma^2_{GDP}\end{bmatrix}\right) =: \mathcal{N}(0, \Sigma_\varepsilon)$$

This scheme assumes that unexpected taxes are due to contemporaneous reactions to unexpected GDP (a_1), responses to government spending shocks (a_2), and responses to tax shocks (ε_t^{TAX}). Similarly, unexpected government spending movements are due to contemporaneous reactions to unexpected GDP (b_1), responses to tax shocks (b_2), and responses to spending shocks (ε_t^{GOV}). Unexpected movements in GDP are assumed to be due to contemporaneous responses to unexpected taxes (c_1) and unexpected spending (c_2), and output shocks (ε_t^{GDP}). There are no contemporaneous reactions to output shocks and GDP reacts to tax and spending shocks only as they transmit through u_t^{TAX} and u_t^{GOV} (i.e., when the shocks move actual taxes and spending and are not offset by automatic stabilizers).

The BP identification scheme better handles the endogeneity concerns that arise from automatic stabilizer and would likely produce better structural results. A rigorous paper should choose the BP scheme over recursive identification, but for our purposes, the recursive scheme should suffice.

B. Impulse Response Functions

The impulse response functions are presented in Figure 8. These IRFs trace out the dynamic effect of a positive one-standard-deviation shock on taxes (column 1), government spending (column 2), and output (column 3) over a horizon of 40 quarters or 10 years. The y-axis can be interpreted as percentage point deviation from the steady state.

We first focus on the effect of a shock to taxes. In the context of fiscal policy, this shock can be thought of as a surprise increase in taxes collected by the government. The tax shock is persistent and remains significant for around 3 years. The shock may also increase government spending by half a percentage point in the subsequent years, but the results are statistically insignificant. The initial impact on output is positive (around a 0.5 percentage point increase) and dies out after 2–4 years. This is a peculiar result: one might expect an increase in taxes to stifle economic activity by reducing consumption spending or investment. Previous fiscal VARs of New Zealand report mixed results on this tax puzzle: Dungey and Fry (2009) likewise find positive short-term impacts on output from tax shocks whereas Parkyn and Vehbi (2014) and Claus et al. (2006) find the opposite. Some authors have suggests a rise in productivity following a tax shock, but it is also possible that our model is misspecified by not accounting for endogeneity concerns with automatic stabilizers.

Moving onto the government spending shock (e.g., a fiscal spending package), we see the shock persist for up to 3–4 years before leveling out. Curiously, the increase in government spending is coupled with a decrease in tax revenues, although a null effect is within the 90% credible bands. The contemporaneous effect on output is slightly positive (a quarter of a percentage point) and turns somewhat negative over the following years, potentially due to the higher real interest rates that follow an increase in spending. This response of output to a spending shock aligns with the results of similar studies (Parkyn and Vehbi,

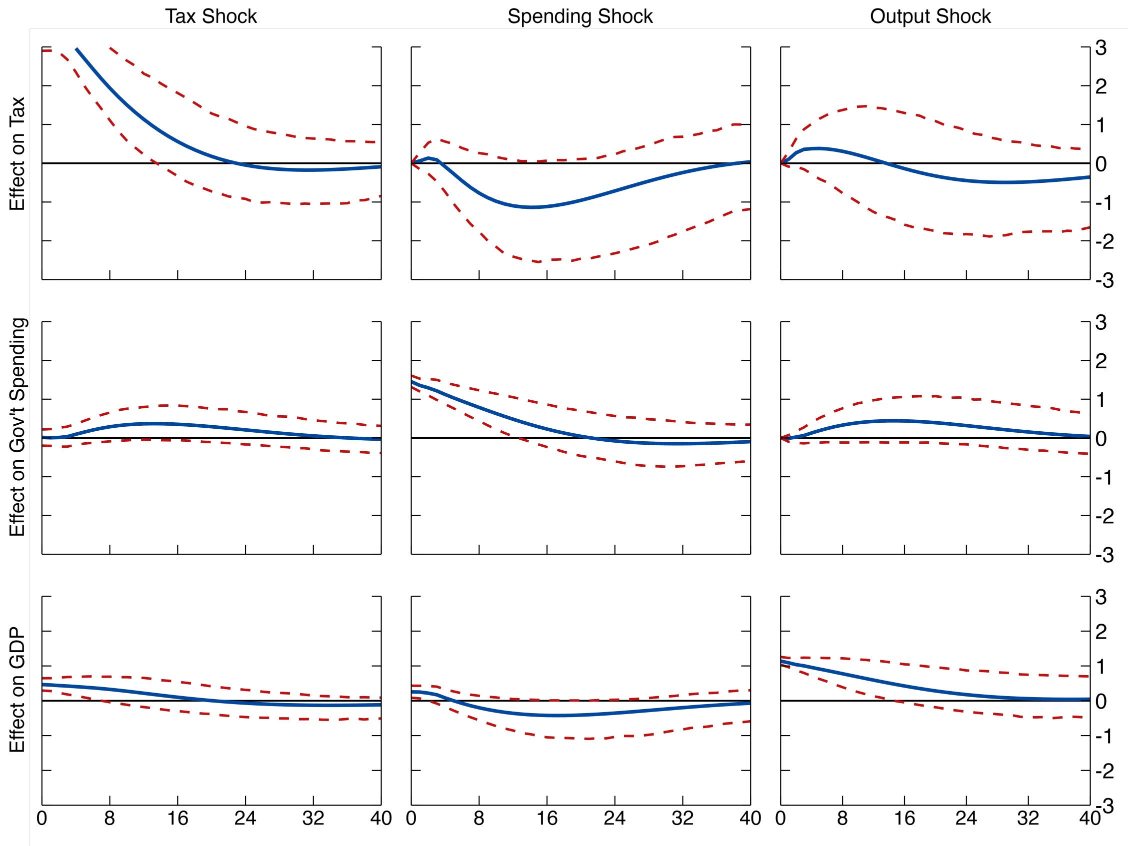

FIGURE 8. IMPULSE RESPONSE FUNCTIONS

Note: Figure 8 presents impulse response functions for one-standard-deviation shocks in net taxes, government spending, and output over a horizon of 40 quarters. The y-axis can be interpreted as percentage point changes relative to a baseline. The solid blue lines are mean IRFs and the dashed red bands provide a 90% credible interval around the mean IRF. The IRFs are estimated by the baseline BVAR with the following order: Taxes, Spending, GDP.

2014; Dungey and Fry, 2009; Claus et al., 2006).

A robustness check switching the recursive ordering such that spending is ordered before taxes is presented in Appendix A1). The results are essentially identical.

Figure 9 traces the cumulative multiplier effect of a dollar spent by the government. Note the large credible bands past 8 quarters: the long-term multiplier estimate is quite unstable, so we limit our analysis to the first 8 quarters. The multiplier is initially positive around 1, suggesting that a dollar of government spending in turn yields a dollar of economic output. The multiplier then appears to decline towards zero and possibly goes negative, indicating that the initial dollar has no (or a negative) long-term effect on economic output. These results are questionable, and given the large uncertainty, we do not read too much into it.

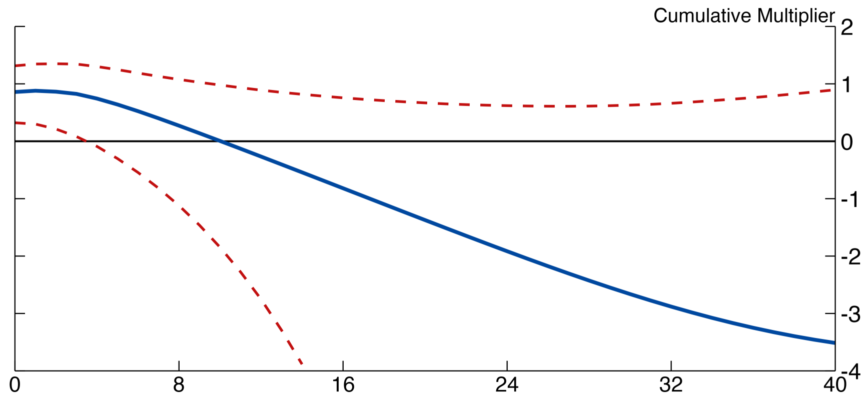

Figure 9. Cumulative Fiscal Spending Multiplier

Note: Figure 9 presents the cumulative multiplier on a dollar spent by the government over a horizon of 10 years. The solid blue line is the mean and the dashed red bands are a 90% credible interval. We calculate the NZD equivalents of the IRF shock on government spending by multiplying the real 2023Q4 value of spending and GDP by their respective IRF trajectory (which are expressed in terms of percent change). The cumulative multiplier effect is derived by dividing the cumulative NZD GDP effect by the cumulative NZD spending effect.

V. Conclusion

This paper estimated reduced-form and structural fiscal VARs to model the dynamic effects of fiscal policy decisions in New Zealand. In our forecast performance assessments, we find that Bayesian models with hyperparameter selection achieve the lowest forecast errors and best density forecasts, especially when inflation and interest rates are added in. Using the large BVAR, we forecast economic growth over the next three year, with a mean annual growth of 1.57% (falling between -0.75% and 3.10% with 90% probability).

Our SVAR analyses suggest that both tax and spending shocks provide a modest boost to economic output in the short term (respectively 0.46 and 0.25 percentage point increases contemporaneously) before zeroing out in the long-term. We also find that government spending has a one-to-one contemporaneous impact on output, wherein a dollar spent by the government raises GDP temporarily by around a dollar before the cumulative effect declines in the medium-term. Our results are relatively robust to the choice of lags and recursive ordering, and credible bands indicates some—but not large—degree of uncertainty in our main results.

These fiscal analyses provide valuable insight into the economic effects of New Zealand's fiscal decisions. There is, of course, room for further research testing alternative structural identification schemes and incorporating more variables to gain a fuller picture of the complexities of fiscal policy.

REFERENCES

Blanchard, Olivier, and Roberto Perotti. 2002. "An empirical characterization of the dynamic effects of changes in government spending and taxes on output." *the Quarterly Journal of economics*, 117(4): 1329–1368.

Claus, Iris, Aaron Gill, Boram Lee, and Nathan McLellan. 2006. "An empirical investigation of fiscal policy in New Zealand." New Zealand Treasury Working Paper.

Del Negro, Marco, and Frank Schorfheide. 2004. "Priors from general equilibrium models for VARs." *International Economic Review*, 45(2): 643–673.

Doan, Thomas, Robert Litterman, and Christopher Sims. 1984. "Forecasting and conditional projection using realistic prior distributions." *Econometric reviews*, 3(1): 1–100.

Dungey, Mardi, and Renée Fry. 2009. "The identification of fiscal and monetary policy in a structural VAR." *Economic Modelling*, 26(6): 1147–1160.

Giannone, Domenico, Michele Lenza, and Giorgio E Primiceri. 2015. "Prior selection for vector autoregressions." *Review of Economics and Statistics*, 97(2): 436–451.

Hamer-Adams, Anna, Martin G Wong, et al. 2018. "Quantifying fiscal multipliers in New Zealand: The evidence from SVAR models." Reserve Bank of New Zealand.

Kelsey, Jane. 2015. *The New Zealand experiment: A world model for structural adjustment?* Bridget Williams Books.

Parkyn, Oscar, and Tugrul Vehbi. 2014. "The effects of fiscal policy in New Zealand: Evidence from a VAR model with debt constraints." *Economic Record*, 90(290): 345–364.

Perotti, Roberto. 2005. "Estimating the effects of fiscal policy in OECD countries." *Available at SSRN 717561.*

Schorfheide, Frank, and Dongho Song. 2015. "Real-time forecasting with a mixed-frequency VAR." *Journal of Business & Economic Statistics*, 33(3): 366–380.

Sims, Christopher A, and Tao Zha. 1998. "Bayesian methods for dynamic multivariate models." *International Economic Review*, 949–968.

Sims, Christopher A, James H Stock, and Mark W Watson. 1990. "Inference in linear time series models with some unit roots." *Econometrica: Journal of the Econometric Society*, 113–144.

Appendix

Table A1—Augmented Dickey-Fuller Unit Root Tests

Variable	ADF Test Statistic	p-value
Net Taxes	-2.951	0.180
Government Spending	-2.307	0.449
GDP	-2.249	0.473
Inflation Rate	-3.657	0.031
Interest Rates	-1.937	0.603

Table A2—RMSFE and Log MDD of Alternative Lag Specifications

Horizon	Variable	2 Lags	4 Lags	6 Lags	8 Lags
One Quarter	Net Taxes	17.558	17.007	16.878	17.077
	Gov't Spending	5.135	5.041	5.108	5.076
	GDP	4.416	4.389	4.354	4.348
Four Quarters	Net Taxes	8.500	8.255	8.117	8.111
	Gov't Spending	1.900	1.844	1.891	1.919
	GDP	1.545	1.473	1.516	1.507
Log MDD		856.09	856.90	856.47	856.45

Note: Table 2 reports root mean squared forecast errors (RMSFE) for h-quarter annualized average growth rate. The baseline specification (Taxes, Spending, GDP) and BVAR estimation with hyperparameter selection on λ is used for all models. The evaluation sample used for recursive forecasting is 2000Q1 to 2019Q4. Log MDD is calculated on the sample from 1989Q2 to 2019Q4.

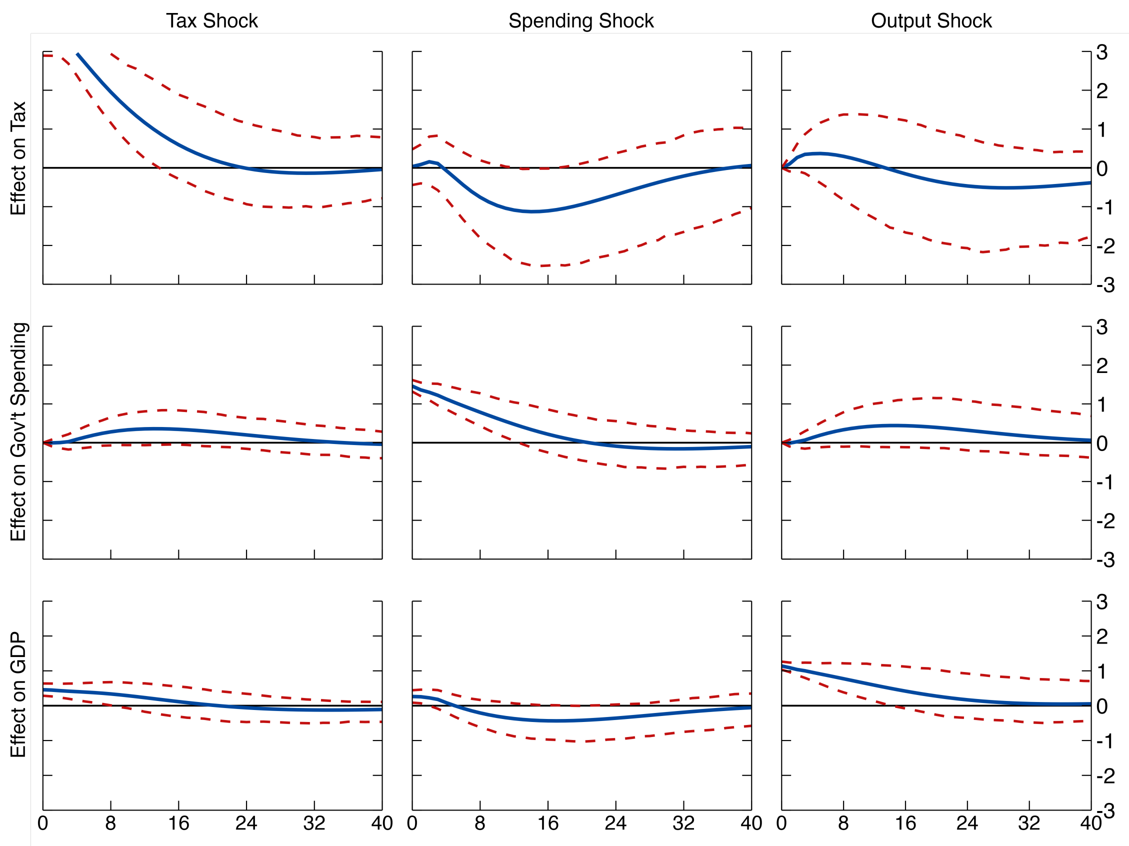

FIGURE A1. IMPULSE RESPONSE FUNCTIONS

Note: Figure A1 presents impulse response functions for one-standard-deviation shocks in net taxes, government spending, and output over a horizon of 40 quarters. The y-axis can be interpreted as percentage point changes relative to a baseline. The solid blue lines are mean IRFs and the dashed red bands provide a 90% credible interval around the mean IRF. The IRFs are estimated by the baseline BVAR with the following order: Spending, Taxes, GDP. The results are nearly identical to those achieved with the original ordering of Taxes, Spending, GDP.