CPB Netherlands Bureau for Economic Policy Analysis

CPB Discussion Paper | 193

The potential of a small model

Adam Elbourne
Coen Teulings

The potential of a small model

Adam Elbourne, Coen Teulings

December 22, 2011

Abstract

This discussion paper highlights potential uses of simple, small models where large traditional models are less flexible. We run a number of experiments with a small two variable VAR model of GDP growth and unemployment with both quarterly and yearly data. We compare the forecasts of these simple models with the published forecasts of the CPB and we conclude that there is not much difference. We then show how easy it is to evaluate the usefulness of a given variable for forecasting by extending the model to include world trade. Perfect knowledge of future world trade growth would help considerably but is obviously not available at the time the forecasts were made. The available world trade data doesn't improve the forecasts. Finally we also show how quick and flexible measures of the output gap can be constructed.

1 Introduction

There is a long tradition in macroeconomic forecasting and policy evaluation at the CPB, using traditional macro models. In this discussion paper, we use a very simple two-variable VAR model with GDP growth and unemployment similar to the model by Blanchard and Quah (1989) to see how a small model compares to a traditional macro model. Our ambition is limited. The paper should be seen as suggesting a research programme; it is a starting point for discussion rather than a balanced final version of an economic model. We make no claim that this model is in any sense the optimal model. We just want to show what such a simple model can bring relative to a traditional macro model and how it is often easier and more flexible to use than a traditional model.

The simple model is estimated using both quarterly and yearly data. On the one hand, using yearly data yields a problem of aggregation in time, which might bias the estimate of the dynamics of the model. On the other hand, budgetary policy has a yearly cycle, so quarterly series are not available for many policy variables. Therefore we also need to know if yearly models can provide adequate performance. For comparing the forecast quality of these models with that of the traditional macro model we run a horse race that strives to make both forecasts comparable. We run specific forecasting models for comparison with the CEP (Centraal Economisch Plan) and MEV (Macro Economische

Verkenning) forecast and discuss how they compare to the optimal race. We then show how the simple model can easily be used to evaluate the importance of specific variables in the forecasting process. We take the example of world trade. Common wisdom tells us that Dutch GDP is heavily correlated with world trade and, hence, that forecasting world trade is crucial for forecasting GDP. The validity of this view is not that obvious, however. That current GDP growth and current world trade growth are highly correlated does not make the forecast of world trade a good forecast instrument for GDP. The simple VAR model allows an easy test of this presumption, by extending the model with an equation for world trade.

Obviously, a simple two variable VAR model has only a limited potential for doing policy analysis. One issue that is often of great importance for policy analysis is the output gap, for example for evaluating the cyclically adjusted budget deficit. First, we analyse whether the VAR model can be used to calculate a proxy for the output gap. In practice measurement of the output gap is not exact, so one can wonder what constitutes a sensible operational definition of the output gap that allows for empirical testing. The interest in the output gap and the spare capacity that it is supposed to measure is in the fact that we expect that, sooner or later, output demand will recover, putting all idle capacity back to use again. Spare capacity leads to above trend output growth. We take this position and use the two variable VAR to calculate the expected above trend growth from here till infinity. We compare this measure of the output gap to alternative ways to calculate this gap. The VAR-based output gap measures are much easier to add extra information to than those measures based on a structural model. Moreover, analysing the impulse response functions of the model provides insight in the time lag for recovery of the output gap.

The remainder of this paper is structured as follows. Section 2 provides evidence that small models are not outperformed by a large traditional model and show how easily they can evaluate the value added of additional variables. Section 3 shows how a measure of the output gap can be made with a simple, easily expandable VAR model. Finally section 4 offers some concluding thoughts.

2 Forecasting

2.1 A simple model

Our first concern is to show that simple VAR models can be useful benchmarks for issues that may arise in forecasting. To be a useful benchmark for future research a model must perform reasonably well at forecasting. Here we define reasonably well as similar to the accuracy of our published forecasts. For our benchmark we use simple stationary VAR models of GDP growth and unemployment. We use VAR models because VARs are simple to estimate, are widely used in practice (Elliott and Timmermann (2008)) and a large class of models can be well approximated by a VAR model (see, for example, Fernandez-Villaverde et al. (2007)).

A VAR(p) is shown in equation (1)

$$x_t = a + A_1 x_{t-1} + ... + A_p x_{t-p} + \varepsilon_t \tag{1}$$

where x_t is a vector of endogenous variables, a is an intercept term, the A_is are matrices of regression coefficients and ε_t is an error term.

One key issue here is to choose the appropriate lag length for the VAR model (Lütkepohl (1993)), which is not always straightforward and can lead to misspecification. To get around this problem, we use the direct multi-period method instead of the iterative method (see Pesaran et al. (2011) and Marcellino et al. (2006) for recent examples and Jordà (2005) for its use for impulse responses). The traditional method of producing forecasts and impulse responses for VAR models involves simulating the estimated VAR model the requisite number of periods forward. If the wrong number of lags has been included or the estimated coefficient deviates from its true value as a result of a limited sample size, for example, the effects of this error will accumulate each extra period the model is simulated. The direct method bypasses the need to simulate the model forward by regressing the impulse response or forecast for a given period directly onto a number of lags. This results in forecasts or impulse responses that are more robust to misspecification whilst it also has the advantage of confidence intervals being much easier to compute with standard t-statistics available. The disadvantage is that the effective sample size decreases by one for each extra period forecast, thus decreasing the available degrees of freedom for estimation. For both yearly and quarterly models we specify equations directly for the yearly growth rates. For the yearly models we estimate

$$\begin{aligned} x_{t+1} &= a + A_1 x_t + \varepsilon_{t+1} \\ x_{t+2} &= a^* + A_1^* x_t + \varepsilon_{t+2}^* \end{aligned}$$

where $x_t^T \equiv [\Delta y_t, u_t]$. For the quarterly models we estimate

$$\begin{aligned} x_{t+1} &= a + A_1 x_q + A_2 x_{q-1} + \varepsilon_{t+1} \\ x_{t+2} &= a^* + A_1^* x_q + A_2^* x_{q-1} + \varepsilon_{t+1}^* \end{aligned}$$

That is, we regress the yearly variables x_{t+i} on the two most recently available quarterly counterparts x_q and x_{q-1}. By regressing yearly variables directly onto quarterly variables bypasses the need to convert a quarterly forecast into a yearly forecast. The lag lengths were chosen because they minimised forecast errors for the yearly and quarterly models.

Table 1 shows the estimation results for yearly models for March/CEP forecasts for both horizons. The GDP growth equation explains over 50% of the variation which translates into an in-sample standard error of 1% of GDP per year. As the horizon is extended to two years the proportion of variation explained drops to 40%, which corresponds to a standard error of 1.1%. Interestingly, if we look at the coefficients, past growth becomes less important and

Table 1: Direct estimates for yearly models: 1979-2008

	One year ahead		Two years ahead	
	Coefficient	Std. Error	Coefficient	Std. Error
Explaining ΔY				
ΔY_{-1}	0.632	0.121	0.112	0.147
U_{-1}	0.283	0.083	0.402	0.095
Constant	-0.010	0.006	-0.005	0.007
R^2	0.564		0.381	
Std. Error	0.010		0.011	
Explaining U				
ΔY_{-1}	-0.606	0.055	-0.907	0.132
U_{-1}	0.865	0.038	0.596	0.086
Constant	0.023	0.003	0.048	0.007
R^2	0.956		0.756	
Std. Error	0.004		0.010	
Cov. Errors	-0.570		-0.749	

past unemployment becomes more important as the horizon goes from one to two years. Unemployment is less volatile than GDP growth. Consequently R^2 is higher with the model explaining about 95% of unemployment variation at the one-year horizon and about 75% at two-years. Once again, going from one-year-ahead to two-years-ahead, the own lag becomes relatively less important whilst lagged GDP growth becomes more important. As expected, the errors in the two equations are negatively correlated: unexpectedly high growth typically occurs when unemployment is unexpectedly low too. Across the two horizons the correlation coefficients are similar although the magnitude of the two-year correlation is slightly larger.

Table 2 shows the estimation results for quarterly models for March/CEP forecasts for both horizons. At the one year horizon for the two lag quarterly model we can see that the model explains just under two-thirds of the variation in GDP growth; variation not explained is just under 0.9% of GDP per year. When the horizon is moved to two years ahead the standard error for the GDP equation rises to over 1%, which corresponds to only about half of the variation being explained. At both horizons the quarterly model is performing better than the yearly model. Once again, given that unemployment is less volatile than GDP growth, R^2 is higher at 0.97 and 0.84 for one- and two-years-ahead, respectively.

In-sample fit is not necessarily a good indicator of out-of-sample forecasting ability (Fildes and Makridakis (1995)). Hence, the key reason for looking at reduced form VAR models here is for making simple forecasts. As a benchmark for accuracy we use the accuracy of CPB's published forecasts, which are made in March (CEP) and September (MEV) each year. To be comparable, the VAR

Table 2: Direct estimates for quarterly models: 1979-2008

	One year ahead		Two years ahead	
	Coefficient	Std. Error	Coefficient	Std. Error
Explaining ΔY				
ΔY_{-1}	0.690	0.217	0.521	0.257
ΔY_{-2}	0.062	0.100	0.114	0.120
U_{-1}	1.104	0.281	0.395	0.330
U_{-2}	-1.027	0.273	-0.226	0.322
Constant	0.001	0.010	-0.021	0.011
R^2	0.649		0.512	
Std. Error	0.009		0.011	
Explaining U				
ΔY_{-1}	-0.256	0.188	-0.862	0.211
ΔY_{-2}	-0.113	0.087	-0.206	0.098
U_{-1}	-0.594	0.243	-1.067	0.271
U_{-2}	1.047	0.236	1.425	0.264
Constant	-0.048	0.008	-0.021	0.009
R^2	0.969		0.841	
Std. Error	0.004		0.009	
Cov. Errors	-0.661		-0.604	

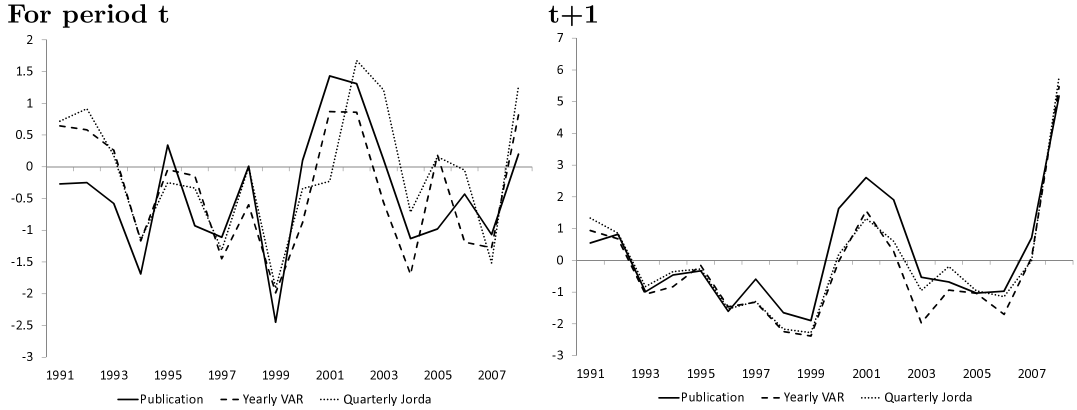
models used here are estimated on a sample up to and including the last observation that would have been available at the time of making the forecast. For example, the CEP 2003 forecast is made using data up to and including the last quarter of 2002. We made the conscious decision to stop the contest in 2008 so that the crisis of 2009 is not included. Such a large outlier would have dominated the results for all models. The accuracy of out-of-sample forecasts is shown in Table 3. The simple VAR models are, in general, similar or slightly less accurate than the published forecasts with the major exception of the MEV forecasts dated t-1. These forecasts are made for the current year with data up to and including the second quarter. The other forecasts have no observations from the year in question. It should be noted that the comparison with CPB's published forecasts is not entirely fair since the VAR models presented here have been estimated on final data whereas the published forecasts were made based upon provisional data. Nevertheless, the general picture that simple reduced form VARs can produce comparably accurate forecasts was also found by Elbourne et al. (2008) with real time data and many VAR models. Furthermore, as with most forecasting competitions, it cannot be excluded that these results are due to chance and the relatively small number of forecast years for comparison. A graphical representation of the forecast errors is shown in figures 1 and 2. It is noticeable that the individual forecasts generally move together: they all seem to be telling the same story so their similar accuracy is not surprising.

Table 3: Out-of-sample MAE for MEV and CEP forecasts: 1991-2008

		MEV		CEP	
		t-1	t	t	t+1
GDP growth					
	Published	0.68	0.97	0.80	1.34
	Yearly			0.84	1.34
	Quarterly	0.60	1.20	0.98	1.23
Unemployment					
	Published	0.24	0.47	0.38	0.72
	Yearly			0.37	0.97
	Quarterly	0.16	0.63	0.38	0.94

Figure 1: GDP growth: forecast errors based on data t-1

For period t **t+1**



Furthermore, they all show serially correlated forecast errors, which could be caused by time-varying coefficients, especially for the intercept. This issue will be discussed further below.

One final issue is that of regime changes and structural breaks. Since large macro models are typically only reestimated occasionally, structural breaks and changing relationships between key variables can influence their performance. Figure 3 shows the estimated coefficients for the yearly models for each forecast made in the competition. With the exception of the coefficient on lagged GDP growth in the GDP growth equation, the coefficients are fairly stable. Even the lagged GDP growth coefficient in the GDP growth equation only varies between 0.55 and just over 0.7. However, we suggested above that the serially correlated forecast errors could have been caused by time-varying coefficients. The change of the coefficient for lagged unemployment in the GDP growth equation increases after the series of negative forecast errors in the 1990s in figure 1. Likewise for unemployment, the decrease of the estimated intercept occurs after the series of over estimates in the 1990s in 2.

Figure 2: Unemployment: forecast errors based on data t-1

For period t **t+1**

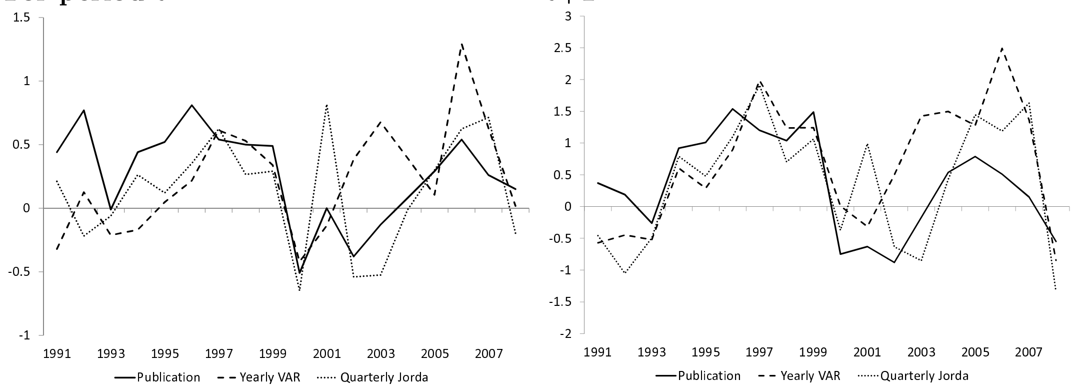
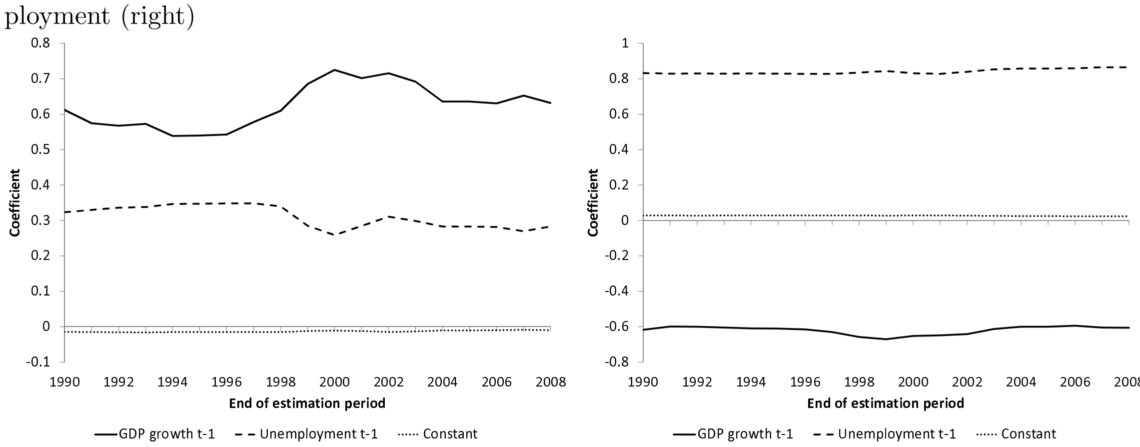


Figure 3: Stability of estimated coefficients for GDP growth (left) and unemployment (right)



2.2 The role of world trade

The Netherlands is a small open economy highly integrated in world trade, especially due to the effects of container shipping through the port of Rotterdam. As such, it would seem logical to assume that world trade growth is an important factor for explaining Dutch GDP growth. This section looks at some simple models that investigate how useful world trade growth is for forecasting Dutch GDP growth. This analysis is easy to perform in our VAR models, we simply add the extra variable and compare performance. In large traditional models it is much less straightforward because the variable in question is often an integral part of the whole model. If that variable is removed then there is no forecast for comparison.

Even if Dutch GDP is largely determined by world trade it does not follow automatically that available world trade data will help forecast GDP growth. One reason that world trade data may not be helpful is that current GDP may depend solely on current world trade, so a forecast of future GDP growth will require a forecast of world trade growth, which may not be easily forecastable. A second reason may be that the same processes that cause GDP, for example, technological developments, may also be the leading causes of world trade - hence, the key information available at a given point in time is already contained in the available GDP data.

To investigate these issues we look at quaterly VAR models for making CEP forecasts. That is, the models used in section 2.1 are extended to include world trade growth.

$$x_t^T \equiv [\Delta y_t, u_t, \Delta w_t]$$

where w_t is world trade twice reweighted. Table 4 contains the reduced form estimates when world trade growth is treated endogenously. Just under 60% of the variation of GDP growth is explained by the model. The individual coefficients on world trade growth are statistically insignificant. Moreover, their magnitudes are not particularly large suggesting that we can treat world trade growth exogenously. Table 5 contains estimates when world trade growth is assumed to be an exogenous process. Now more of the variation in GDP growth is explained, about 65%. This comes because the contemporaneous value of world trade growth is now an explanatory variable. Interestingly, the coefficient on lagged GDP growth falls by approximately the same magnitude as the new coefficient on contemporaneous world trade growth.

The forecasting performance of the different model variations is compared in table 6. This is once again an out-of-sample forecasting competition allong the same lines as in section 2.1. The Base models are models without world trade from section 2.1. Only when the GDP forecasts are conditioned on the true future values of world trade is there any improvement in forecast accuracy. In fact, the improvements in that case are substantial with the 2 year MAE falling by 50%. So if we had a perfect forecast of world trade growth we could significantly improve our GDP growth forecasts. Unfortunately, that is not the case. We are reliant on forecasts of world trade. Table 6 shows forecast accuracy based on two different forecasts of world trade growth: one

Table 4: Yearly VAR with endogenous world trade. Standard errors in ( )

	GDP growth	Unemployment	Trade growth
ΔGDP_{-1}	0.553	-0.532	0.707
	(0.162)	(0.071)	(0.479)
$Unemployment_{-1}$	0.280	0.867	0.393
	(0.088)	(0.038)	(0.259)
$\Delta Trade_{-1}$	0.063	-0.060	-0.200
	(0.080)	(0.035)	(0.236)
Constant	-0.011	0.025	0.018
	(0.007)	(0.003)	(0.020)
R-squared	0.574	0.961	0.143
S.E. equation	0.010	0.004	0.029

Table 5: Yearly VAR with exogenous world trade. Standard errors in ( )

	GDP growth	Unemployment
ΔGDP_{-1}	0.347	-0.389
	(0.148)	(0.234)
$Unemployment_{-1}$	0.057	1.135
	(0.056)	(0.052)
$\Delta Trade$	0.186	-0.019
	(0.058)	(0.054)
$\Delta Trade_{-1}$	0.063	0.001
	(0.058)	(0.065)
R-squared	0.665	0.861
S.E. equation	0.009	0.008

Table 6: Out-of-sample MAE of world trade models: 1991-2008

	GDP growth		Unemployment		World trade growth	
	1 year	2 years	1 year	2 years	1 year	2 years
Base model						
VAR	0.84	1.34	0.37	0.97		
Direct	0.84	1.31	0.37	0.98		
World trade model						
VAR, all endogeneous	0.87	1.40	0.38	0.99	3.14	3.93
VAR exo wtg, actual	0.56	0.73	0.70	1.43	0.00	0.00
VAR exo wtg, CEP	0.80	1.36	0.67	1.43	2.06	4.09
	0.80	1.34	0.78	1.54	2.06	4.09

that is made by the estimated VAR model shown in table 4 and one where we use CPB's published world trade growth forecasts as the exogenous series for world trade in the estimated VAR model shown in table 5. Neither improves forecast accuracy significantly relative to the baseline models or the published CEP forecasts. Interestingly, knowledge of world trade does not help forecast accuracy for unemployment in these VAR models.

2.3 Conclusions

In this section a simple two-variable VAR-type model was estimated for the purpose of forecasting Dutch macro aggregates. The simple model produces estimates of in-sample correlations that we would have expected: economic growth and unemployment are negatively correlated. The models perform reasonably well at forecasting with forecast errors similar or slightly worse than CPB's published forecasts. This is further evidence along the lines started by Nelson (1972) that simple, small models can be at least as accurate as larger models and that for basic forecasting tasks they are a quicker and perfectly valid alternative to larger models. A similar result was found in Elbourne et al. (2008).

Since the basic model produces reasonable forecasts it is a useful testing ground for additional variables. Adding world trade growth to VAR models has the potential to significantly improve the accuracy of GDP growth forecasts since perfect information of future world trade growth reduces forecast errors from about 0.85% of GDP and 1.3% of GDP to about 0.6% and 0.7% at the one- and two-year-ahead horizons, respectively. However, the volume of world trade is not known perfectly in advance and must be itself be forecast. Neither of the methods used in this paper for forecasting world trade growth helps to improve the accuracy or GDP growth forecasts. For unemployment, not even perfect knowledge of future world trade growth improves the forecasts from the simple time-series models. It would be interesting to see if these results were replicable in a wider range of models, since what is true for a simple VAR model is not necessarily true for all model classes.

This sort of analysis of the added value for forecasting is much simpler and quicker to perform in a small VAR type model than in a large traditional model. In a large traditional model the role of a given variable, for example world trade, is hard-wired into the structure of the model. If you don't provide a path for future world trade you don't get a forecast for the remaining variables. In that case there is no easily available alternative model without world trade to compare the forecasts from the large model with. The same argument also holds true for any new variables that may be argued can help with forecasting. Adding these to a VAR model is much easier than adding them to a structural model. Furthermore, in addition to the ease of undertaking forecasting competitions, simple small models also allow the researcher to use techniques like Granger causality (see Granger (1969)) to evaluate the benefit of extra variables for forecasting, which are less straightforward to apply in traditional models.

3 The output gap

We can also use this simple two variable model to calculate an estimate of the output gap. This section will present two ways of doing this which relate to two different definitions of the output gap.

3.1 The Blanchard-Quah method

Blanchard and Quah (1989) proposed to decompose the reduced form errors of the two variable model into 'supply' shocks and 'demand' shocks. They did this by the assumption that only supply shocks have a long-run effect on the level of GDP. For illustration purposes we will ignore the intercept term in the VAR and we will also restrict ourselves to one lag for simplicity; a constant and longer lags also work with this approach, although the maths is less tidy. This leaves us with the following reduced form

$$x_t = A_1 x_{t-1} + \varepsilon_t. \tag{2}$$

We can apply any transformation B to the errors of this model that we like

$$\varepsilon_t = B e_t \tag{3}$$

where

$$E\left[\varepsilon_t \varepsilon_t'\right] = E\left[B e_t e_t' B'\right]. \tag{4}$$

Typically we assume that the underlying structural errors are uncorrelated so that $E\left[e_t e_t'\right] = I$, then each shock can be analysed in isolation. Now suppose we select some transformation B that allows us to intepret the errors in some meaningful way, then we have a structural model:

$$x_t = A_1 x_{t-1} + B e_t. \tag{5}$$

Because the reduced form covariance matrix is symmetric, equation (3) only imposes 3 restrictions on B. The Blanchard-Quah method involves imposing an extra restriction on B through the long-run properties of the model. Specifically, only one of the shocks is allowed to have a long-run effect on the level of GDP. So what is the long-run effect of a shock on the level of GDP? To see this, we need to substitute out x_{t-1} to give

$$x_t = A_1 \left(A_1 x_{t-2} + Be_{t-1}\right) + Be_t$$

and simplifying

$$x_t = {A_1}^2 x_{t-2} + Be_t + A_1 Be_{t-1}.$$

Then substituting out x_{t-2} gives

$$x_t = {A_1}^3 x_{t-3} + Be_t + A_1 Be_{t-1} + {A_1}^2 Be_{t-2}.$$

If the VAR is stationary, the coefficent on lagged x, ${A_1}^k$, tends to zero as the horizon increases meaning that x_t is solely a function of past shocks. Moreover, the effect of a shock on x_t diminishes geometrically as the horizon increases. Since x_t contains the growth rate of GDP, we need to add these effects together to get the long-run effect on the level of GDP. Since the effects diminish geometrically we can use the matrix equivalent of the sum-to-infinity formula to define the long-run effect matrix, C:

$$C = \left(I - A_1\right)^{-1} B. \tag{6}$$

In the bivariate system considered here imposing one restriction on C is enough to identify the structural matrix B. The restriction imposed is that one of the shocks, which will be labeled a 'demand' shock has zero long-run effect on the level of GDP.

Before we turn our attention to making a measure of the output gap it will be informative to look at the dynamic properties of this system to see if the identified shocks can plausibly be labelled 'supply' and 'demand'. Figure 4 shows impulse response functions for both 'supply' and 'demand' shocks for the quarterly model. The average effect of a negative 'supply' shock is to lower GDP in the long-run. In the mean time, unemployment falls. At no horizon for either of the responses to the 'supply' shock are they statistically significantly different from zero, however. This is not surprising given the ongoing debate in the literature regarding the sign of the repsonse of employment to technology shocks (see, for example, Christiano et al. (2004)). In contrast, the responses to the 'demand' shock are statistically different from zero. After a negative 'demand' shock unemployment rises and GDP falls. The estimates presented here suggest that a 'demand' shock that had a peak impact on GDP of a 1% reduction would, on average, raise unemployment by about 0.4 percentage points. For unemployment the baseline is back within the confidence bounds after about 5-6 years, which is slightly slower than GDP which takes only 4-5 years. Figure 5 show the same responses but from a model estimated on yearly data. In general, the pattern is similar to the quarterly model. The effects of 'supply' shocks are not statistically significant but show the same pattern whereby shocks that would lower long-run GDP, lower unemployment in the short-run. Again, the ‘demand’ shock responses are also comparable: a 1% peak increase in GDP would raise unemployment by about 0.5 percentage points and GDP again crosses the baseline slightly before unemployment. Therefore we can conclude that the model gives plausible responses to demand and supply shocks for the Netherlands and, hence, the shocks themselves are plausible.

One way to generate an historical measure of the output gap is to define trend GDP as the estimated Blanchard-Quah model with demand shocks set to zero. The difference between the stochastic trend caused by the supply shocks and the real world world observed values of GDP is the component caused by the transitory demand shocks. That is, take the estimated structural residuals

$$e_t = \left[e_t^{supply}, e_t^{demand}\right]'$$

and create a counterfactual series where there are no demand shocks:

$$e_t^* = \left[e_t^{supply}, 0\right]'.$$

Then use this set of shocks to simulate what would have happened with no demand shocks:

$$x_t^* = A_1 x_{t-1}^* + B e_t^*. \tag{7}$$

Then the output gap is simply the realisations minus the counterfactual series:

$$Gap_t = x_t - x_t^*. \tag{8}$$

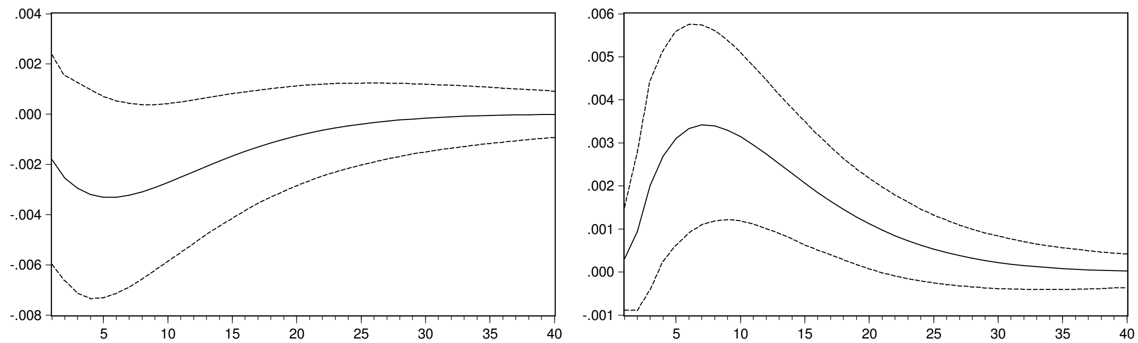
Claus et al. (2000) detail how a similar model, with an additional capacity utilisation variable, has been one of the methods used by the Reserve Bank of New Zealand to infer the level of potential output. Once again, it should be stressed that this method relies on the assumption of constant coefficients, which makes it less than ideally suited to estimating the output gap today. Provisional conclusions from Teulings and Zubanov (2010) tell us that banking crisis are additive (no interaction with the output gap at the start of banking crisis) and that they yield a large immediate loss (6%), with slow and little recovery afterwards. This suggests at the very least the correct time series model that includes the crisis should contain a structural break in the level of GDP. This can be handled fairly easily in a small VAR model with a simple dummy variable.

3.2 The multivariate Beveridge-Nelson method

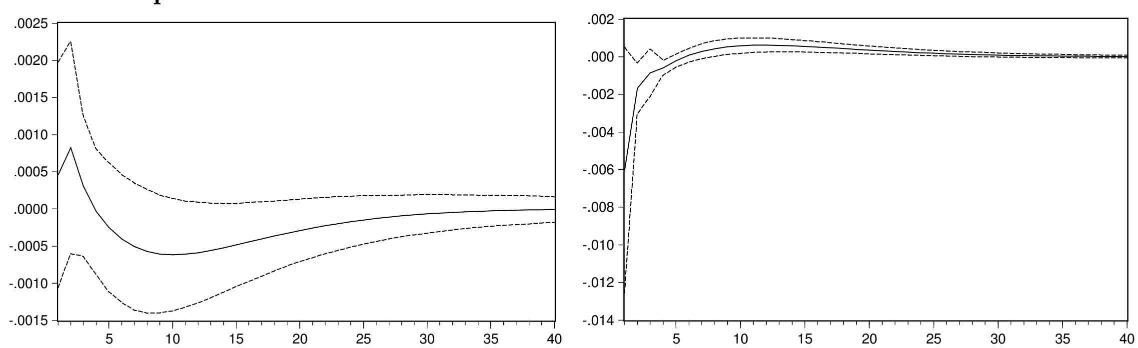
An output gap like measure can also be made directly from the reduced form estimates. Theory tells us that output gaps close over time, so one potential measure of the output gap is the expected excess future growth above trend. That faster growth comes from closing the output gap. Consider equation (2),

Figure 4: Quarterly impulse responses to supply (left) and demand (right) shocks

Unemployment response



Growth response



GDP level response

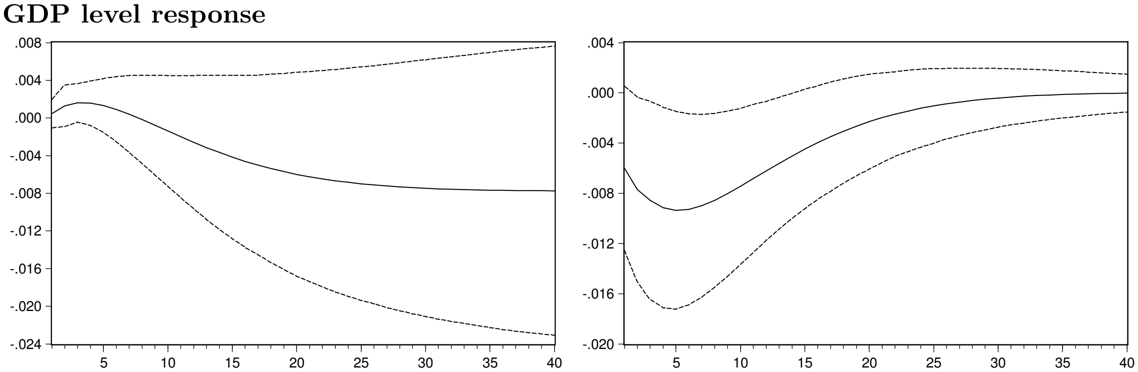
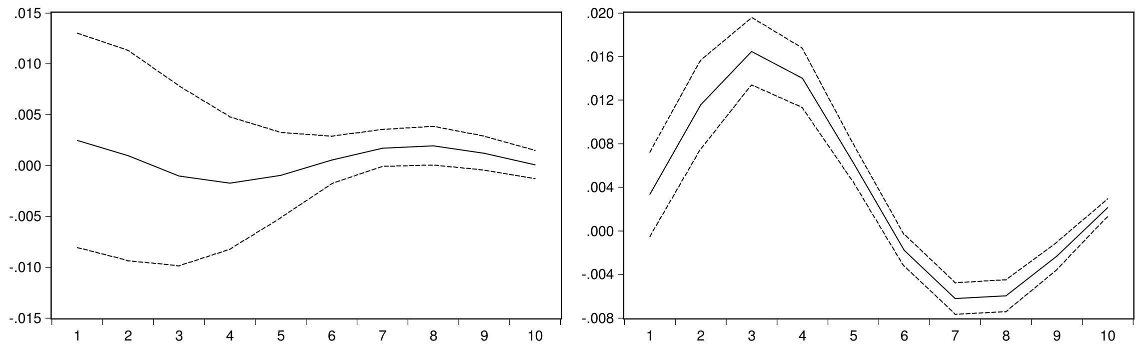
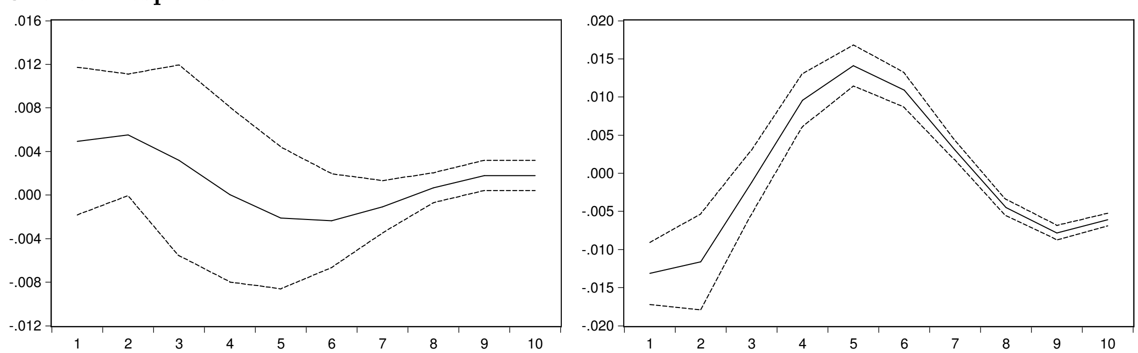


Figure 5: Yearly impulse responses to supply (left) and demand (right) shocks

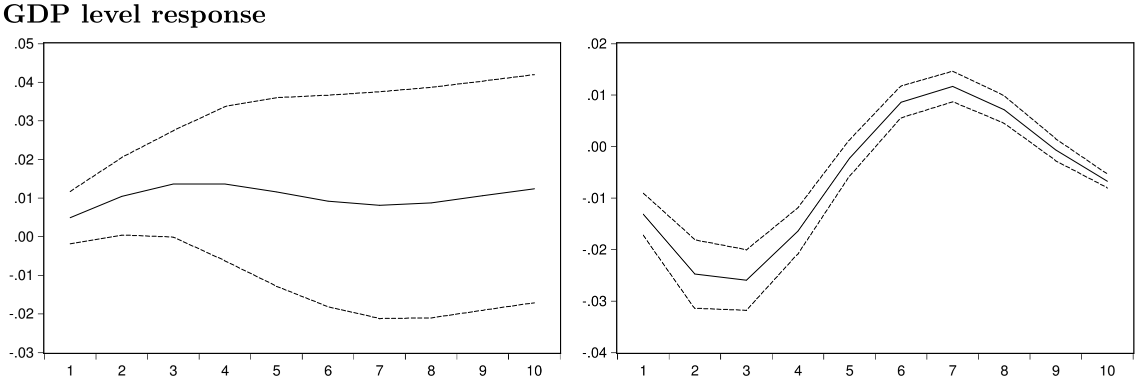
Unemployment response



Growth response



GDP level response



but now including an intercept. Define x to be the steady state of x_t. The expression for x reads

$$x = (I - A)^{-1} a.$$

Define z_t to be the deviation of x_t from its steady state value

$$z_t \equiv x_t - x.$$

Since the first element of x_t is the growth of GDP, z_t is the deviation of the growth rate from its long term average. Define z_t^* as the value of z_t conditional on the information available at time $t = 0$:

$$z_t^* \equiv \mathrm{E}\left[z_t | z_0\right].$$

Since $z_t^* = A z_{t-1}^*$, an expression for z_t can be found by solving the differential equation.

$$z_t^* = A^t z_0 = B \Lambda^t d,$$

where $A = B \Lambda B^{-1}$ with Λ being the diagonal matrix of eigenvalues and B a matrix of eigenvectors. d is a vector of initial conditions at $t = 0$, which can be solved from

$$d = B^{-1} z_0,$$

The state of the economy at $t = 0$ can be fully characterised by the two state variables of the system, the vector z_0. For the purpose of this paper, the output gap is defined as the deviation of current GDP from its long run growth path, or equivalently, to the cumulative excess growth above steady state of GDP that is to be expected from now until infinity. The output gap can therefore be defined as the sum of expected future excess growth z_t^*.

$$f \equiv \sum_{t=1}^{\infty} z_t^* = B \left[\Lambda + \Lambda^2 + \Lambda^3 ...\right] B^{-1} z_0 = B \Lambda \left[I - \Lambda\right]^{-1} B^{-1} z_0. \tag{9}$$

The output gap is the first element of f. The first row of $B \Lambda \left[I - \Lambda\right]^{-1} B^{-1}$ are the coefficients relating the state variables z_0 to the output gap.

This characterisation works only if all eigenvalues are real. If the eigenvalues are complex, the solution is somewhat different. The eigenvalues satisfy $\lambda^+ = a + bi$ and $\lambda^- = a - bi$. Let v be the eigenvector associated with λ^-. Define the matrices:

$$\Lambda \equiv \begin{bmatrix} a & -b \\ b & a \end{bmatrix}, B \equiv \left[\mathrm{Re}\,(v)\,, \mathrm{Im}\,(v)\right].$$

Then, equation (9) applies. This is the Multivariate Beveridge-Nelson decomposition for the cyclical component (Beveridge and Nelson (1981) introduced this technique for univariate series). Specifically, positive values (higher than typical expected growth) imply that the current value is below the stochastic trend. It must be stressed that this stationary component is not the same as the standard definition of an output gap. The standard definition is the deviation of current output from current potential whereas this measure gives the deviation of current output from future potential. One key advantage of this approach is that it only uses past information at each point in time to make each estimate of the output gap - there is no requirement for future values to give an estimate of the trend.

With an additional lag, a similar methodology can be worked out. In that case, the state of the economy is summarised by four instead of two variables, the growth of GDP and the unemployment rate of last period and of the period before. Where only one lag is sufficient to characterise the annual data, a model for the quarterly data requires two lags. Hence, the output gap in that model is a function of four state variables.

3.3 Results

Figure 6 plots various output gap measures next to each other. Those indicated by 'BQ_Q' and 'BQ_Y' are the output gaps made by the quarterly and yearly Blanchard-Quah VARs introduced here. By definition the output gap in the Blanchard-Quah method is zero before the first observation used for estimation. From 1988 to 2001 the BQ output gaps are from VAR estimated on the data from 1988 to 2001. From 2002 onwards the output gaps are psuedo real-time estimates of the output gap. For example, the quarterly output gap for the 2nd quarter of 2005 use data up to and including 2005. From 2002 onwards we also show the forecast based Beveridge-Nelson output gap described in section 3.2 above. For comparison we also display three other measure of the outputgap based on the latest available data. The other measures shown are the European Commission estimate, 'EC', the CPB's own measure, 'CPB', and finally from an HP filter, 'HP'. All of the measures display similar features although there are some large discrepancies as to the magnitudes. For example, all of the methods pick up a peak around 2008, but the quarterly Blanchard-Quah measure put GDP about only 1% above trend, whereas the EC, the CPB method and the HP filter all put the gap closer to 3%. The psuedo real-time VAR based measures were all slow to pick up on the trough around 2003.

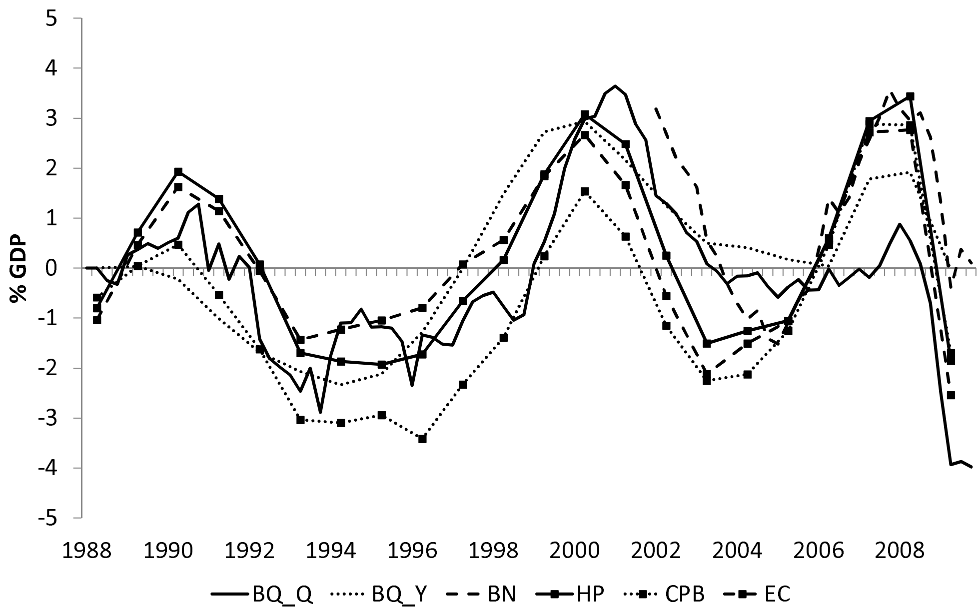
For the Beveridge-Nelson decomposition we provide the relation between the state variable z_t (the deviation from their long run value of the growth of GDP and of unemployment) and the output gap. We have:

$$\begin{aligned} Gap_t^{annual} &= -0.386 g_{t-1} + 1.278 u_{t-1}, \\ Gap_t^{quartely} &= -0.784 g_{t-1} + 4.098 u_{t-1} - 0.404 g_{t-2} - 1.404 u_{t-2}. \end{aligned}$$

The output gap turns out to be a negative function of last period's growth rate and a positive function of unemployment. A high growth rate during last period predicts lower growth in the future. A high unemployment rate shows that part of labour supply is not employed currently, which opens the possibility of faster GDP growth by reemploying idle labour supply in the near future. One percentage point unemployment is expected to lead more than one percent excess future growth of GDP. This captures both labour hoarding and the discouraged worker effect. Both effects imply that the measured unemployment rate understates the actual degree of underemployment, since part of the formally employed workers are in fact not doing much productive work (labour hoarding) and since part of the workers who are actually willing to take a job do not make the effort to report as unemployed. In the quarterly model, these effects are even more pronounced. The less pronounced pattern in the annual data suggests that part of the actual dynamics is swept under the carpet due to an aggregation-in-time bias.

The fact that the sum of the coefficients for g_{t-1} and g_{t-2} in the quaterly model is about four times the coefficient of g_{t-1} in the annual model fits the expectation, since the annual growth rate is on average equal to four time the quarterly growth rate.

Figure 6: Output gaps



3.4 Conclusions

This section has detailed two ways of making output gap statistics from a simple VAR model and has compared them to existing measures. The usefulness of these output gap measures depends on their intended use. For policy analysis the difference in definition between the two measures may also be important. If we want to know what the cyclical component of the budget deficit is then we probably want to look at an output gap relative to current potential. If we want to know if the current deficit will cause sustainability problems we may wish to look at the gap relative to future potential. If the output gaps are supposed to be used in forecasting then these output gap calculations are not very useful. This is especially so for the forward-looking multivariate Beveridge-Nelson output gap since it is clear that this measure tells us nothing that the forecast itself from the model would tell us. In fact, this output gap measure merely takes the information contained in a period-by-period forecast and summarises those individual pieces of information at each forecast horizon by one statistic.

It would also be relatively easy to add additional variables to help distinguish between permanent and transitory components of GDP. For example, Claus et al. (2000) add a measure of capacity utilisation to the simple Blanchard-Quah model. Incorporating extra information into the multivariate Beveridge-Nelson approach is a simple case of adding an extra variable to the forecasting VAR. In a large traditional model adding a variable would typically require adding some structure and some story to accomodate the extra variable in the model, which is not a straightforward nor uncontroversial task.

4 Conclusions

This paper has argued that simple, small models can provide quick, easy and useful analyses that would typically be undertaken using large models. The simple two- and three-variable VAR type models presented here have provided quick and easy evaluations of the importance of the available world trade data for forecasting and some possibilities for measuring the output gap quickly. Neither issue is as quick or easy to perform with large traditional models.

References

Beveridge, S. and C. Nelson, 1981, A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the business cycle, *Journal of Monetary economics*, vol. 7, no. 2, pp. 151–174.

Blanchard, O. and D. Quah, 1989, The dynamic effects of aggregate demand and supply disturbances, *The American Economic Review*, vol. 79, no. 4, pp. 655–673.

Christiano, L., M. Eichenbaum and R. Vigfusson, 2004, The response of hours to a technology shock: Evidence based on direct measures of technology, *Journal of the European Economic Association*, vol. 2, no. 2-3, pp. 381–395.

Claus, I., P. Conway and A. Scott, 2000, *The output gap: measurement, comparisons and assessment*, Reserve Bank of New Zealand.

Elbourne, A., H. Kranendonk, R. Luginbuhl, B. Smid and M. Vromans, 2008, Evaluating CPB's published gdp growth forecasts, CPB Documents 172, CPB Netherlands Bureau for Economic Policy Analysis.

Elliott, G. and A. Timmermann, 2008, Economic forecasting, *Journal of Economic Literature*, vol. 46, no. 1, pp. 3–56.

Fernandez-Villaverde, J., J. Rubio-Ramirez, T. Sargent and M. Watson, 2007, ABCs (and Ds) of understanding VARs, *The American Economic Review*, vol. 97, no. 3, pp. 1021–1026.

Fildes, R. and S. Makridakis, 1995, The impact of empirical accuracy studies on time series analysis and forecasting, *International Statistical Review/Revue Internationale de Statistique*, vol. 63, no. 3, pp. 289–308.

Granger, C., 1969, Investigating causal relations by econometric models and cross-spectral methods, *Econometrica: Journal of the Econometric Society*, pp. 424–438.

Jordà, O., 2005, Estimation and inference of impulse responses by local projections, *American Economic Review*, vol. 95, no. 1, pp. 161–182.

Lütkepohl, H., 1993, Introduction to multiple time series analysis.

Marcellino, M., J.H. Stock and M.W. Watson, 2006, A comparison of direct and iterated multistep ar methods for forecasting macroeconomic time series, *Journal of Econometrics*, vol. 135, no. 1-2, pp. 499–526.

Nelson, C., 1972, The prediction performance of the frb-mit-penn model of the us economy, *The American Economic Review*, vol. 62, no. 5, pp. 902–917.

Pesaran, M.H., A. Pick and A. Timmermann, 2011, Variable selection, estimation and inference for multi-period forecasting problems, *Journal of Econometrics*, vol. 164, no. 1, pp. 173–187.

Teulings, C.N. and N. Zubanov, 2010, Is economic recovery a myth? robust estimation of impulse responses, CESifo Working Paper Series 3027, CESifo Group Munich.

Fildes, R. and S. Makridakis, 1995, The impact of empirical accuracy studies on time series analysis and forecasting, *International Statistical Review/Revue Internationale de Statistique*, vol. 63, no. 3, pp. 289–308.

Granger, C., 1969, Investigating causal relations by econometric models and cross-spectral methods, *Econometrica: Journal of the Econometric Society*, pp. 424–438.

Jordà, O., 2005, Estimation and inference of impulse responses by local projections, *American Economic Review*, vol. 95, no. 1, pp. 161–182.

Lütkepohl, H., 1993, Introduction to multiple time series analysis.

Marcellino, M., J.H. Stock and M.W. Watson, 2006, A comparison of direct and iterated multistep ar methods for forecasting macroeconomic time series, *Journal of Econometrics*, vol. 135, no. 1-2, pp. 499–526.

Nelson, C., 1972, The prediction performance of the frb-mit-penn model of the us economy, *The American Economic Review*, vol. 62, no. 5, pp. 902–917.

Pesaran, M.H., A. Pick and A. Timmermann, 2011, Variable selection, estimation and inference for multi-period forecasting problems, *Journal of Econometrics*, vol. 164, no. 1, pp. 173–187.

Teulings, C.N. and N. Zubanov, 2010, Is economic recovery a myth? robust estimation of impulse responses, CESifo Working Paper Series 3027, CESifo Group Munich.

Publisher:

CPB Netherlands Bureau for Economic Policy Analysis
P.O. Box 80510 | 2508 GM The Hague
T (070) 3383 380

November 2011 | ISBN 978-90-5833-530-2