THE UNIVERSITY OF MELBOURNE

Melbourne Institute Working Paper Series

Working Paper No. 19/09

A Latent Variable Approach to Forecasting the Unemployment Rate

Chew Lian Chua, G. C. Lim and Sarantis Tsiaplias

MELBOURNE INSTITUTE®
of Applied Economic and Social Research

A Latent Variable Approach to Forecasting the Unemployment Rate*

Chew Lian Chua, G. C. Lim and Sarantis Tsiaplias
Melbourne Institute of Applied Economic and Social Research,
The University of Melbourne

Melbourne Institute Working Paper No. 19/09

ISSN 1328-4991 (Print)

ISSN 1447-5863 (Online)

ISBN 978-0-7340-3314-7

July 2009

* This research was funded by an Australian Research Discovery Grant (DP0877424). We are grateful to John Creedy for helpful comments.

Melbourne Institute of Applied Economic and Social Research
The University of Melbourne
Victoria 3010 Australia
***Telephone* (03) 8344 2100**
***Fax* (03) 8344 2111**
***Email* melb-inst@unimelb.edu.au**
***WWW Address* http://www.melbourneinstitute.com**

Abstract

A forecasting model for unemployment is constructed that exploits the time-series properties of unemployment while satisfying the economic relationships specified by Okun's law and the Phillips curve. In deriving the model, we jointly consider the problem of obtaining estimates of the unobserved potential rate of unemployment consistent with Okun's law and Phillips curve, and associating the potential rate of unemployment to actual unemployment. The empirical example shows that the model clearly outperforms alternative forecasting procedures typically used to forecast unemployment.

Keywords: Forecasting, Unemployment, Unobserved Components.

JEL classification: C53; E27

1 Introduction

Economic forecasters are often confronted with the problem of how to design forecasting models which are based on economic theory and which also exploit the time-series properties inherent in the data. Both aspects are important: theory guides the selection of relevant variables and the specification of relationships in the forecasting model while the properties of the data guide the efficient use of information in the estimation.

The aim of this paper is to show how forecasts of the unemployment rate, one of the most important variables in economics, can be improved by applying economic theory in an econometric framework which emphasises the time series properties inherent in the data. Such an approach is, of course, not new; the novel aspect of this paper is that two of the most popular relationships in economics - Okun's Law and Phillips Curve - which relate the movements of output, prices and unemployment, are jointly expressed as a system of 'gap' equations in a state space form to produce forecasts of unemployment. The resulting forecast of unemployment is consistent with both Okun's law and the Phillips curve, as well as the time-series properties of actual unemployment.

The econometric forecasting model uses the trend-cyclical decomposition methods of Harvey et. al. (1986), Kuttner (1994) and Malley and Molana (2008). More specifically, we rely on Attfield and Silverstone (1998) who show that the unobserved components in any 'gap' equation may be identified by reference to the stochastic trend emanating from a Beveridge-Nelson decomposition, and the work of Anderson et al. (2006) who represent the Beveridge and Nelson (1981) decomposition in the single source of error state space framework. The advantage of this set-up is that it provides a parsimonious and efficient way to combine Okun's law and Phillips curve in a form that is easily estimable.

The forecasting performance of the model is evaluated against forecasts from VAR and Bayesian VAR models estimated using the same information set. The VAR and BVAR approaches are commonly used for economic forecasting and numerous studies have provided evidence regarding their forecasting capacity relative to structural univariate or AR approaches (Artis and Zhang, 1990; Ramos, 2003; see also De Gooijera and Hyndman (2006) for a general review).

The paper is organised as follows. In the next section, we specify a latent variable forecasting model using Okun's law and the Phillips curve. Section 3 details the model selection process used in this paper. The results are presented in Section 4, with Section 5 concluding the paper.

2 A latent variable forecasting model based on Okun's law and Phillips curve

This section sets up a system of gap equations that jointly capture the economic relationships postulated by Okun (1962) and Phillips (1958). Importantly, both theories signify an explanatory role for the deviation between the unemployment rate and the 'natural' rate of unemployment in the movements of output and inflation respectively.

The empirical relationship between the rate of unemployment and output suggested by Okun (1962) can be expressed as a gap equation

$$y_t - y_t^* = \beta_1(u_t - u_t^*) + \zeta_t \tag{1}$$

where y_t is the logarithm of observed real output, u_t is the observed unemployment rate, y_t^* and u_t^* correspond to potential output and the 'natural' rate of unemployment respectively, and ζ_t is an i.i.d. normal error term. Okun's coefficient, β_1, is a measure of the responsiveness of the unemployment rate to output growth. Attfield and Silverstone (1998) show that β_1 can be interpreted as a cointegration coefficient if y_t^* and u_t^* are defined in terms of long run stochastic trend values. More importantly, they show that the long run stochastic trend has a Beveridge-Nelson decomposition (Beveridge and Nelson, 1981)

Phillips (1958) asserts a similar relationship between inflation and the unemployment rate gap

$$p_t - p_t^* = \beta_2(u_t - u_t^*) + \xi_t \tag{2}$$

where p_t is the inflation rate, p_t^* represents underlying inflation, and ξ_t is an iid-normal error term. As with (1), it can be showed that the long run stochastic trend in (2) has a Beveridge and Nelson decomposition.

Pursuant to the Beveridge-Nelson decomposition of (1), (2), the set of variables y_t, p_t and u_t can be expressed in terms of their trend and cyclical components

$$y_t = y_t^* + y_t^c \tag{3}$$

$$u_t = u_t^* + u_t^c \tag{4}$$

$$p_t = p_t^* + p_t^c \tag{5}$$

where y_t^c, u_t^c and p_t^c are the (by definition, stationary) cyclical components.

From equations (3), (4) and (5), it can be shown that the gap equations (1) and (2) respectively, can be written as

$$y_t^c = \beta_1 u_t^c + \zeta_t \tag{6}$$

$$p_t^c = \beta_2 u_t^c + \xi_t \tag{7}$$

which shows that the cyclical components y_t^c and p_t^c are commonly driven by u_t^c. Replacing equation (6) into (3) and equation (7) into (5) respectively, yields

$$y_t = y_t^* + \beta_1 u_t^c + \zeta_t \tag{8}$$

$$p_t = p_t^* + \beta_2 u_t^c + \xi_t \tag{9}$$

where the error terms ζ_t and ξ_t are assumed to be contemporaneously correlated.

Anderson et al. (2006) indicate that trends and cycles can be formulated using a single source of error state space form. That is, there is perfect correlation between innovations to the trends and cycles.

There are several methods that may be used to model the unobserved components y_t^*, u_t^*, and p_t^*. We follow the common practice of treating the latent variables as random walks with drift and an ARMA(2,1) process[1] for the stationary variable u_t^c. The model becomes

$$y_t^* = \tau_1 + y_{t-1}^* + \gamma_1 \varepsilon_{1t} \tag{10}$$

$$u_t^* = \tau_2 + u_{t-1}^* + \gamma_2 \varepsilon_{2t} \tag{11}$$

$$p_t^* = \tau_3 + p_{t-1}^* + \gamma_3 \varepsilon_{3t} \tag{12}$$

$$u_t^c = \phi_1 u_{t-1}^c + \phi_2 u_{t-2}^c + \theta \varepsilon_{2t-1} + (1-\alpha)\varepsilon_{2t}. \tag{13}$$

To allow a more flexible ARMA process, we also model a variant of u_t^c that is driven by a trigonometric process (i.e., a mixture of sine and cosine waves)

$$\begin{bmatrix} u_t^c \\ u_t^{+c} \end{bmatrix} = \rho \begin{bmatrix} \cos\lambda & \sin\lambda \\ -\sin\lambda & \cos\lambda \end{bmatrix} \begin{bmatrix} u_{t-1}^c \\ u_{t-1}^{+c} \end{bmatrix} + \begin{bmatrix} \alpha\varepsilon_{2t} \\ \alpha\varepsilon_{2t} \end{bmatrix} \tag{14}$$

where ρ is a damping factor such that $\rho \in [0,1)$ and λ is the cycle frequency spanning from 0 to π. For identification, we assume u_t^c and u_t^{+c} are driven by a common disturbance $\alpha\varepsilon_{2t}$ (see Harvey, 1989, pg. 39). u_t^{+c} is needed for the construction of u_t^c. When $\lambda = 0$, the equation u_t^{+c} becomes redundant which implies u_t^c is driven by an AR(1) process. For $|\rho| < 1$, and $0 < \lambda < \pi$, u_t^{+c} and u_t^c are stationary ARMA(2,1) processes. Unlike (13), this alternative functional form results in an ARMA process that incorporates pseudo-cyclical behaviour (Koopman et al., 2005).

Accordingly, these relationships constitute a system of structural equations that form the basis for obtaining estimates of the unemployment rate. The focus here is strictly on the unemployment rate for two reasons. First, as discussed in a later section, the measurement of an appropriate price variable is an issue we explore and second, the proposed system is only a partial explanation of the determination of the output variable.

[1] The ARMA(2,1) is chosen since it accomodates the possibility of cyclical behavior in u_t^c. Alogoskoufis and Stengos (1991) find that the US unemployment rate is appropriately modelled as an AR(2) process with ARCH errors.

2.1 A state space representation for (y_t, u_t, p_t)

In this section, we present a state space representation for (y_t, u_t, p_t). Without loss of generality, to express the state space with a common source of error in the measurement and transition equations we assume that $\zeta_t = \gamma_4 \varepsilon_{1t}$ and $\xi_t = \gamma_5 \varepsilon_{3t}$.

Together with equations (8), (9), (10), (11), (12) and (13), it can be shown that there exists a state space form

$$\begin{bmatrix} y_t \\ u_t \\ p_t \end{bmatrix} = \begin{bmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & \beta_1\phi_1 & \beta_1 \\ 0 & 1 & 0 & \phi_1 & 1 \\ 0 & 0 & 1 & \beta_2\phi_1 & \beta_2 \end{bmatrix} \begin{bmatrix} y^*_{t-1} \\ u^*_{t-1} \\ p^*_{t-1} \\ u^c_{t-1} \\ d_{t-1} \end{bmatrix} + \eta_t \tag{15}$$

$$\begin{bmatrix} y^*_t \\ u^*_t \\ p^*_t \\ u^c_t \\ d_t \end{bmatrix} = \begin{bmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \phi_1 & 1 \\ 0 & 0 & 0 & \phi_2 & 0 \end{bmatrix} \begin{bmatrix} y^*_{t-1} \\ u^*_{t-1} \\ p^*_{t-1} \\ u^c_{t-1} \\ d_{t-1} \end{bmatrix} + \begin{bmatrix} \delta_1 & \delta_2 & 0 \\ 0 & \gamma_2 & 0 \\ 0 & \delta_4 & \delta_3 \\ 0 & (1-\gamma_2) & 0 \\ 0 & \theta & 0 \end{bmatrix} \eta_t \tag{16}$$

$$\eta_t = \begin{bmatrix} (\gamma_1+\gamma_4) & \beta_1(1-\gamma_2) & 0 \\ 0 & 1 & 0 \\ 0 & \beta_2(1-\gamma_2) & (\gamma_3+\gamma_5) \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix} \tag{17}$$

where η_t is common in both the measurement and transition equations, $\delta_1 = \frac{\gamma_1}{\gamma_1+\gamma_4}$, $\delta_2 = \delta_1 \left(-\beta_1 + \beta_1\gamma_2\right)$, $\delta_3 = \frac{\gamma_3}{\gamma_3+\gamma_5}$ and $\delta_4 = \delta_3 \left(-\beta_2 + \beta_2\gamma_2\right)$.

Similarly, when equations (8), (9), (10), (11), (12) and (14) are taken together, it can be shown that the following state space form exists

$$\begin{bmatrix} y_t \\ u_t \\ w_t \end{bmatrix} = \begin{bmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & \beta_1\rho\cos\lambda & \beta_1\rho\sin\lambda \\ 0 & 1 & 0 & \rho\cos\lambda & \rho\sin\lambda \\ 0 & 0 & 1 & \beta_2\rho\cos\lambda & \beta_2\rho\sin\lambda \end{bmatrix} \begin{bmatrix} y^*_{t-1} \\ u^*_{t-1} \\ w^*_{t-1} \\ u^c_{t-1} \\ u^{+c}_{t-1} \end{bmatrix} + \eta_t \quad (18)$$

$$\begin{bmatrix} y^*_t \\ u^*_t \\ w^*_t \\ u^c_t \\ u^{+c}_t \end{bmatrix} = \begin{bmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \rho\cos\lambda & \rho\sin\lambda \\ 0 & 0 & 0 & -\rho\sin\lambda & \rho\cos\lambda \end{bmatrix} \begin{bmatrix} y^*_{t-1} \\ u^*_{t-1} \\ w^*_{t-1} \\ u^c_{t-1} \\ u^{+c}_{t-1} \end{bmatrix} + \begin{bmatrix} \delta_1 & \delta_4 & 0 \\ 0 & 1-\delta_3 & 0 \\ 0 & \delta_5 & \delta_2 \\ 0 & \delta_3 & 0 \\ 0 & \delta_3 & 0 \end{bmatrix} \eta_t \quad (19)$$

$$\eta_t = \Sigma\varepsilon_t = \begin{bmatrix} (\gamma_1+\gamma_4) & \beta_1\alpha & 0 \\ 0 & (\gamma_2+\alpha) & 0 \\ 0 & \beta_2\alpha & (\gamma_3+\gamma_5) \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix} \quad (20)$$

where $\delta_1 = \frac{\gamma_1}{\gamma_1+\gamma_4}$, $\delta_2 = \frac{\gamma_3}{\gamma_3+\gamma_5}$, $\delta_3 = \frac{\alpha}{\alpha+\gamma_2}$, $\delta_4 = -\alpha\beta_1\frac{\gamma_1}{\alpha\gamma_1+\alpha\gamma_4+\gamma_1\gamma_2+\gamma_2\gamma_4}$, and $\delta_5 = -\alpha\beta_2\frac{\gamma_3}{\alpha\gamma_3+\alpha\gamma_5+\gamma_2\gamma_3+\gamma_2\gamma_5}$.

Maximum likelihood estimates of the model parameters are obtained using

$$\log\ L(\tau,\beta,\rho,\delta,,\lambda) \propto -\frac{T}{2}\log|\Sigma| - \frac{1}{2}\sum_{t=1}^{T}\eta_t'\Sigma\eta_t \quad (21)$$

Ord et al. (1997) showed that maximum likelihood estimation based on exponential smoothing can be employed to estimate a state space model with common errors in the measurement and transition equations. The common practice in the literature (Hyndman et al., 2002; Taylor, 2008) is to use the first few observations to estimate the initial values. As the initial values are not maximum likelihood estimates, however, optimisation is achieved by maximising a conditional likelihood function (Taylor, 2008). As the state space model comprises a vector (rather than a scalar) of common errors in the transition and measurement equations, the appropriate likelihood function is identical to that of the vector exponential smoothing framework in de Silva et al (2008). The use of the single source of error state space form avoids the need for matrix inversion to obtain the state variables (as opposed to Kalman filter based estimation). Consequently, the computational order of the model is approximately equal to that of a VAR model of similar dimension.

3 Measuring prices and model selection

An obvious impediment to implementing the model defined in Sections 2 and 3 is that, unlike output and unemployment, there are alternative measures of prices that may be used. Alternative price measures will yield different measures of potential output and, therefore, different forecasts of actual unemployment rates. Three widely adopted measures of prices are considered in this paper: wage inflation w_t, headline inflation h_t, and underlying inflation l_t. The use of wage inflation is consistent with the original formulation of the Phillips curve (Phillips, 1958) whereas later research typically uses headline or underlying inflation to measure prices in the Phillips curve context (King and Watson, 1995; Gruen et al., 1999; Hamilton, 2001).

The alternative choices available for measuring prices raise the possibility of alternative permutations of the basic forecasting model. The underlying rationale for considering permutations of the basic model is that a particular permutation may be able to better capture aggregate price changes and thus of better forecasts of the unemployment rate. Appendix A provides a list of the permutations considered which include the cases with and without Okun's Law as well as alternative combinations of two price measures. For completeness, Appendix B presents the state space forms for the permutations.

The forecasting exercise adopted in this paper generates forecasts from each permutation (or model), and a running forecast that depends on an in-sample model selection process. The in-sample selection process is implemented to minimize the potential for forecast bias (see Rapach et al. (2004) regarding the minimisation of forecast bias using variable selection procedures reliant only on in-sample data). Since alternative measures of inflation are considered in this paper, the identity of the price variable, in contrast to the output and unemployment variables, is not fixed. The likelihood function, however, contains information regarding the fit of the entire set of equations, including the inflation equation (2). To avoid choosing a particular model due to the fit associated with a given proxy for p_t and y_t, it is necessary that the model selection procedure be independent of p_t and y_t. This may be achieved by considering the marginal likelihood function of u_t. Since we assume that the model errors are normally distributed, the joint density of y_t, u_t, p_t is multivariate normal. Consequently, the computation of the marginal distribution for u_t is straightforward; the marginal distribution of u_t is univariate normal.

Akaike's information criterion (IC) (Akaike, 1974) and Schwarz's Bayesian IC (Schwarz, 1978) are commonly used to select among competing model choices.[2] Rather than incorporate the full likelihood function into the AIC and BIC, the likelihood function is replaced with the appropriate marginal likelihood in line with the preceding discussion. We denote both measures as partial AIC (PAIC) and partial BIC (PBIC) given their reliance on the partial information set.

$$PAIC = 2k - 2\ln L$$

$$PBIC = 2k\ln n - 2\ln L$$

where k is the number of parameters, n is the number of observations and L is the marginal likelihood computed using only the second element of η_t and Σ_{22}.

4 Evaluation of the forecasts

The data used in the empirical analysis are (quarterly) real GDP, the unemployment rate and the inflation rate for Australia over the period 1983:1 to 2008:2. Since the GDP and inflation measures are quarterly data, we convert the monthly unemployment rate into a quarterly figure. The three

[2]The use of AIC and BIC for single source of error models is motivated by Ord et al. (1997)

possible measures of inflation considered are: headline inflation, underlying[3] (weighted median) inflation and wage inflation based on average weekly earnings.[4] The series for GDP, unemployment, headline inflation, and average weekly earnings are obtained from the Australian Bureau of Statistics website (www.abs.gov.au) while the underlying inflation data are obtained from the Reserve Bank of Australia (RBA) website (www.rba.gov.au).

The forecasts are evaluated using a rolling forecast exercise. To do so, we first divide the sample into two periods. The first period from 1983:3 to 1995:1 is used for initial estimation while the second period, from 1995:2 to 2008:2, is the out-of-sample evaluation period. The forecast exercise proceeds as follows. Commencing from the first period, the parameters are estimated and 1 to 4 step-ahead forecasts are computed. The parameters are re-estimated with an extra period of observations and once again 1 to 4 step-ahead forecasts are computed. This process is repeated until the sample reaches 2008:1. Note that at this period, only 1 step-ahead forecasts are produced because there is only one period of observations remaining (similarly in 2007:4 only 1 and 2 step-ahead forecasts are constructed). In summary, 52 periods of observations are used to determine the forecast accuracy for 1 step ahead forecasts, 51 periods for 2 step-ahead forecasts, 50 periods for 3 step-ahead forecasts and 49 periods for 4 step-ahead forecasts. In addition, for each period we also conduct an in-sample exercise that uses the PAIC and PBIC to select the best in sample model and the selected model is used to generate 1-4 step-ahead forecasts.We use RMSE statistics to evaluate the forecast performance of the models.

Table 1 presents the models' forecasting accuracy. In the table, we report the performance of each permutation listed in Appendix A. We also provide composite forecasts computed using the mean of the set of forecasts provided by the alternative permutations, as well as forecasts associated with the in-sample model selection procedure specified in Section 2.2. We benchmark the accuracy of the forecasts using three basic forecasting models: a Bayesian

[3]Unlike the headline inflation figure, the weighted median measure of underlying inflation decreases the influence of the most volatile price changes, thereby producing a smoother measure of consumer prices.

[4]The start date of this series on the ABS website is 1983:1.

VAR, a standard VAR, and a random walk.

Table 1: RMSE statistics for the unemployment rate

	Unemployment rate					Forecast rank				
	1	2	3	4	GM	1	2	3	4	GM
M_1	0.1517	0.1398	0.1321	0.1440	0.1417	2	17	6	7	6
M_2	0.1646	0.1358	0.1362	0.1462	0.1453	20	10	10	8	9
M_3	0.1617	0.1342	0.1420	0.1609	0.1493	15	6	13	20	17
M_4	0.1608	0.1299	0.1312	0.1431	0.1407	14	3	5	6	3
M_5	0.1545	0.1282	0.1360	0.1546	0.1429	4	1	9	13	8
M_6	0.1631	0.1412	0.1382	0.1479	0.1473	18	19	11	10	12
M_7	0.1597	0.1364	0.1431	0.1590	0.1492	13	12	16	17	16
M_8	0.1752	0.1456	0.1349	0.1468	0.1499	27	22	8	9	18
M_9	0.1581	0.1399	0.1464	0.1609	0.1511	10	18	21	20	20
M_{10}	0.1719	0.1445	0.1469	0.1609	0.1557	24	21	22	20	22
M_{11}	0.1744	0.1536	0.1541	0.1764	0.1643	26	26	25	25	25
M_{12}	0.1573	0.1361	0.1426	0.1524	0.1469	7	11	15	11	11
M_{13}	0.1622	0.1358	0.1308	0.1386	0.1414	16	10	4	2	5
M_{14}	0.1596	0.1288	0.1268	0.1378	0.1377	12	2	1	1	1
M_{15}	0.1630	0.1323	0.1297	0.1403	0.1408	17	5	3	4	4
M_{16}	0.1579	0.1386	0.1454	0.1617	0.1506	9	16	20	22	19
M_{17}	0.1589	0.1371	0.1422	0.1557	0.1482	11	14	14	14	14
M_{18}	0.1562	0.1372	0.1395	0.1537	0.1464	6	15	12	12	10
M_{19}	0.1684	0.1460	0.1516	0.1655	0.1576	22	23	24	24	23
M_{20}	0.1650	0.1343	0.1326	0.1420	0.1429	21	7	7	5	8
M_{21}	0.1691	0.1433	0.1435	0.1612	0.1539	23	20	18	21	21
M_{22}	0.1724	0.1485	0.1505	0.1632	0.1583	25	24	23	23	24
CF_u	0.1577	0.1307	0.1295	0.1394	0.1389	8	4	2	3	2
$PAIC_u$	0.1548	0.1369	0.1445	0.1586	0.1484	5	13	19	16	15
$PBIC_u$	0.1545	0.1356	0.1433	0.1577	0.1475	4	8	17	15	13
$BVAR$	0.1237	0.1502	0.1959	0.2553	0.1746	1	25	26	26	26
VAR	0.1644	0.1976	0.2695	0.3554	0.2362	19	27	27	27	27
RW	1.0000	1.0000	1.0000	1.0000	1.0000	28	28	28	28	28

The VAR and BVAR models are constructed using the five variables at our disposal: unemployment rate, real GDP growth, headline inflation, underlying inflation, and wage inflation. These are all variables for which economic theory specifies important co-relationships. For the BVAR, we follow LeSage (1990) in adopting a Minnesota (or Litterman) prior for the intercept and slope coefficients, with the parameters estimated using Theil's mixed estimation approach (see, also, Litterman, 1986). Pursuant to the Minnesota prior, we assume an a priori random walk forecasting framework. We apply 'tight' hyperparameters for the prior, thereby reducing the weight attached to higher lags of the dependent variables.[5] Lag lengths are chosen using the AIC.

Of the 22 permutations considered, the permutation featuring unemployment, GDP, underlying inflation and a trigonemetric process for the cyclical unemployment component u_t^c (permutation 14 in Appendix A) provides the best RMSE forecasts for 3 and 4 quarters ahead and the 2nd smallest RMSE for 2-quarter ahead forecasts. Permutation 14 also outperforms the combined forecast CF, generated using all 22 permutations, and the forecasts based on PAIC and PBIC in-sample selection at every forecast horizon. This is expected given that the PAIC and PBIC approaches tend to select permutations 12, 16, 17, and 18 which are not among the better forecasting models (see Table 2). The improved forecasts obtained using permutation 14 suggest that the inclusion of the smoother underlying inflation variable to capture the price movements in Phillips equation yields a better forecast for unemployment than either the relatively noisier headline inflation figure or wage inflation.

The unemployment forecasts generated by all but one permutation of the combined Okun's and Phillips models produce a smaller RMSE than the VAR model for any forecast horizon. This suggests that the theoretical advantages of the approach considered here, relative to an 'uninformative' or unrestricted model such as the VAR, result in significant forecast improvements with

[5]For equation i, we adopt an overall tightness parameter of 0.1, a harmonic lag decay of 1, a weight of 0.1 for lags of variable i, and symmetric weights of 0.1 for lags of other variables.

relatively little additional computational requirements.

Numerous studies have provided evidence that the Bayesian variant of the VAR model yields forecasting improvements relative to the standard VAR (LeSage, 1990; Shoesmith, 1992; Joutz et al., 1995; Amisano and Serati, 1999). Consequently, the BVAR provides a useful benchmark for assessing the forecasting qualities of the combined Okun and Phillips model. In this respect, the BVAR model produces the best single period-ahead unemployment forecast. This improvement, however, is not maintained for the 2, 3 and 4 step ahead forecasts, with the BVAR producing higher RMSE than all but one of the permutations considered.

The forecasts generated using the in-sample PAIC and PBIC based selection procedures also produce smaller average RMS errors than the VAR or BVAR approaches, suggesting that the forecasting improvement does not stem from any selection bias. Importantly, although the RMSE statistics for the BVAR at 4-steps ahead are approximately twice the magnitude of the 1-step ahead RMSE, the RMSE statistics for the PAIC and PBIC based forecasts are of a similar magnitude at all four forecast steps. Consequently, unlike the approach forwarded here, the quality of the forecasts generated by a BVAR imbued with theoretically appropriate output, unemployment and price data appear to be limited to a single period. In contrast, the combined Okun's law and Phillips curve approach appears to produce accurate forecasts for all four quarters. These results suggest that the combined approach specified here produces incremental forecasting benefits, especially where the forecast period of interest is greater than the next-period ahead.

The RMSE statistics are presented relative to the RMSE for the random walk (RW) benchmark. Values greater (lesser) than unity indicate a higher (lower) RMSE than the RW benchmark. CF is generated using the mean of the 22 permutations considered. The $PAIC$ and $PBIC$ models are determined using the approach specified in Section 2.2. The VAR and BVAR models are vector autogression and Bayesian vector autoregression models respectively. GM is the geometric mean of the RMSE for 1 to 4 steps ahead.

Table 2. Models selected from PAIC and PBIC for best in-sample fit for u_t for period between 1995:2 to 2008:03

	AIC	BIC		AIC	BIC
M_1	0	0	M_{12}	22	0
M_2	0	0	M_{13}	0	0
M_3	0	0	M_{14}	0	0
M_4	0	0	M_{15}	0	0
M_5	0	0	M_{16}	7	10
M_6	1	1	M_{17}	14	26
M_7	0	0	M_{18}	8	15
M_8	0	0	M_{19}	0	0
M_9	0	0	M_{20}	0	0
M_{10}	0	0	M_{21}	0	0
M_{11}	0	0	M_{22}	0	0

5 Conclusion

This paper showed that the unemployment rate could be forecasted from an economic model of GDP growth, the unemployment rate and the inflation rate which was consistent with two popular economic relationships - Okun's Law and Phillips Curve. More importantly the relationships can be parsimoniously and efficiently specified as a system of gap equations which can then be expressed as a Beveridge decompositional single source state-space model. This formulation is advantageous because it produces an estimate of potential unemployment consistent with both theories that is used to forecast the unemployment rate. Moreover, the single source of error state space formulation also reduces the computational requirements of the approach, as the system is equivalent to a VAR model of similar dimension.

Overall, the forecasts generated by combining the theoretical insights of Okun and Phillips appear to improve on those generated using standard atheoretical forecasting models such as the VAR and BVAR models. Nearly all the permutations of the combined approach improve on the unemployment forecasts provided by the BVAR, which performs poorly for all but one-step ahead predictions. Importantly the RMSE of the forecasts generated using the combined approach, in contrast to those generated using the VAR and BVAR models, are similar for forecasts at 1, 2, 3 and 4 -quarters ahead suggesting that the combination of Okun's law and Phillips curve advocated here is useful for generating both short and medium term unemployment forecasts typically used in macroeconomic analysis.

6 Appendix A: Models

Table A1. Combination of unemployment rates, GDP, headline inflation, underlying inflation and wage inflation with u_t^c specfied as either an ARMA process or a trigonometric process.

Model	GDP	UR	HI	UI	WI	$ARMA$	
1	×	×				×	
2	×	×	×			×	
3	×	×		×		×	
4	×	×			×	×	
5		×	×			×	
6		×		×		×	
7		×			×	×	
8	×	×	×	×	×	×	
9	×	×	×	×		×	
10	×	×		×	×	×	
11	×	×	×		×	×	
12	×	×					×
13	×	×	×				×
14	×	×		×			×
15	×	×			×		×
16		×	×				×
17		×		×			×
18		×			×		×
19	×	×	×	×	×		×
20	×	×	×	×			×
21	×	×		×	×		×
22	×	×	×		×		×

7 Appendix B

Let h_t, l_t and w_t represent headline, underlying and wage inflation respectively and let h_t^*, l_t^* and w_t^* be the corresponding unobserved trend components. Examples of the state space form for the 2,4 and 5 equation systems are provided below. For each of the systems, u_t^c is expressed as an ARMA(2,2) process and, alternatively, as a trigonometric process.

- **2-equation system (y_t,u_t) with ARMA(2,2) process in u_t^c**

$$\begin{bmatrix} y_t \\ u_t \end{bmatrix} = \begin{bmatrix} \tau_1 \\ \tau_2 \end{bmatrix} + \begin{bmatrix} 1 & 0 & \beta_1\phi_1 & \beta_1 \\ 0 & 1 & \phi_1 & 1 \end{bmatrix} \begin{bmatrix} y_{t-1}^* \\ u_{t-1}^* \\ u_{t-1}^c \\ d_{t-1} \end{bmatrix} + \eta_t \tag{22}$$

$$\begin{bmatrix} y_t^* \\ u_t^* \\ u_t^c \\ d_t \end{bmatrix} = \begin{bmatrix} \tau_1 \\ \tau_2 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \phi_1 & 1 \\ 0 & 0 & \phi_2 & 0 \end{bmatrix} \begin{bmatrix} y_{t-1}^* \\ u_{t-1}^* \\ u_{t-1}^c \\ d_{t-1} \end{bmatrix} + \begin{bmatrix} \delta_1 & \delta_2 \\ 0 & \gamma_2 \\ 0 & 1-\gamma_2 \\ 0 & \theta \end{bmatrix} \eta_t$$

where $\eta_t = \begin{bmatrix} (\gamma_1+\gamma_4) & \beta_1(1-\gamma_2) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$, $\delta_1 = \frac{\gamma_1}{\gamma_1+\gamma_4}$ and $\delta_2 = \delta_1 (\beta_1\gamma_2 - \beta_1)$.

- **2-equation system (y_t,u_t) with trigonometric process in u_t^c**

$$\begin{bmatrix} y_t \\ u_t \end{bmatrix} = \begin{bmatrix} \tau_1 \\ \tau_2 \end{bmatrix} + \begin{bmatrix} 1 & 0 & \beta_1\rho\cos\lambda & \beta_1\rho\sin\lambda \\ 0 & 1 & \rho\cos\lambda & \rho\sin\lambda \end{bmatrix} \begin{bmatrix} y_{t-1}^* \\ u_{t-1}^* \\ u_{t-1}^c \\ u_{t-1}^{+c} \end{bmatrix} + \eta_t \tag{23}$$

$$\begin{bmatrix} y_t^* \\ u_t^* \\ u_t^c \\ u_t^{+c} \end{bmatrix} = \begin{bmatrix} \tau_1 \\ \tau_2 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \rho\cos\lambda & \rho\sin\lambda \\ 0 & 0 & -\rho\sin\lambda & \rho\cos\lambda \end{bmatrix} \begin{bmatrix} y_{t-1}^* \\ u_{t-1}^* \\ u_{t-1}^c \\ u_{t-1}^{+c} \end{bmatrix} + \begin{bmatrix} \delta_1 & \delta_3 \\ 0 & 1-\delta_2 \\ 0 & \delta_2 \\ 0 & \delta_2 \end{bmatrix} \eta_t$$

where $\delta_1 = \frac{\gamma_1}{\gamma_1+\gamma_4}$, $\delta_2 = \frac{\alpha}{\alpha+\gamma_2}$, $\delta_3 = -\alpha\beta_1\frac{\gamma_1}{\alpha\gamma_1+\alpha\gamma_4+\gamma_1\gamma_2+\gamma_2\gamma_4}$ and

$$\eta_t = \begin{bmatrix} (\gamma_1+\gamma_4) & \beta_1\alpha \\ 0 & (\gamma_2+\alpha) \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

- **4-equation system (y_t,u_t, w_t, h_t) with ARMA(2,2) process in u_t^c**

$$\begin{bmatrix} y_t \\ u_t \\ w_t \\ h_t \end{bmatrix} = \begin{bmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \\ \tau_4 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 & \beta_1\phi_1 & \beta_1 \\ 0 & 1 & 0 & 0 & \phi_1 & 1 \\ 0 & 0 & 1 & 0 & \beta_2\phi_1 & \beta_2 \\ 0 & 0 & 0 & 1 & \beta_3\phi_1 & \beta 3 \end{bmatrix} \begin{bmatrix} y_{t-1}^* \\ u_{t-1}^* \\ w_{t-1}^* \\ h_{t-1}^* \\ u_{t-1}^c \\ d_{t-1} \end{bmatrix} + \eta_t \qquad (24)$$

$$\begin{bmatrix} y_t^* \\ u_t^* \\ w_t^* \\ h_t^* \\ u_t^c \\ d_t \end{bmatrix} = \begin{bmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \\ \tau_4 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \phi_1 & 1 \\ 0 & 0 & 0 & 0 & \phi_2 & 0 \end{bmatrix} \begin{bmatrix} y_{t-1}^* \\ u_{t-1}^* \\ w_{t-1}^* \\ h_{t-1}^* \\ u_{t-1}^c \\ d_{t-1} \end{bmatrix}$$
$$+ \begin{bmatrix} \delta_1 & \delta_2 & 0 & 0 \\ 0 & \gamma_2 & 0 & 0 \\ 0 & \delta_3 & \delta_4 & 0 \\ 0 & \delta_5 & 0 & \delta_6 \\ 0 & 1-\gamma_2 & 0 & 0 \\ 0 & \theta & 0 & 0 \end{bmatrix} \eta_t$$

where $\eta_t = \begin{bmatrix} (\gamma_1+\gamma_6) & \beta_1(1-\gamma_2) & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \beta_2(1-\gamma_2) & (\gamma_3+\gamma_7) & 0 \\ 0 & \beta_3(1-\gamma_2) & 0 & (\gamma_4+\gamma_8) \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \\ \varepsilon_{4t} \end{bmatrix}$,
$\delta_1 = \frac{\gamma_1}{\gamma_1+\gamma_6}$, $\delta_2 = \frac{\gamma_1}{\gamma_1+\gamma_6}(-\beta_1+\beta_1\gamma_2)$, $\delta_3 = \frac{\gamma_3}{\gamma_3+\gamma_7}(-\beta_2+\beta_2\gamma_2)$, $\delta_4 = \frac{\gamma_3}{\gamma_3+\gamma_7}$, $\delta_5 = \frac{\gamma_4}{\gamma_4+\gamma_8}(-\beta_3+\beta_3\gamma_2)$ and $\delta_6 = \frac{\gamma_4}{\gamma_4+\gamma_8}$.

- **4-equation system (y_t,u_t, w_t, h_t) with trigonometric process in** u_t^c

$$\begin{bmatrix} y_t \\ u_t \\ w_t \\ h_t \end{bmatrix} = \begin{bmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \\ \tau_4 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 & \beta_1\rho\cos\lambda & \beta_1\rho\sin\lambda \\ 0 & 1 & 0 & 0 & \rho\cos\lambda & \rho\sin\lambda \\ 0 & 0 & 1 & 0 & \beta_2\rho\cos\lambda & \beta_2\rho\sin\lambda \\ 0 & 0 & 0 & 1 & \beta_3\rho\cos\lambda & \beta_3\rho\sin\lambda \end{bmatrix} \begin{bmatrix} y_{t-1}^* \\ u_{t-1}^* \\ w_{t-1}^* \\ h_{t-1}^* \\ u_{t-1}^c \\ u_{t-1}^{+c} \end{bmatrix} + \eta_t \tag{25}$$

$$\begin{bmatrix} y_t^* \\ u_t^* \\ w_t^* \\ h_t^* \\ u_t^c \\ u_t^{+c} \end{bmatrix} = \begin{bmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \\ \tau_4 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \rho\cos\lambda & \rho\sin\lambda \\ 0 & 0 & 0 & 0 & -\rho\sin\lambda & \rho\cos\lambda \end{bmatrix} \begin{bmatrix} y_{t-1}^* \\ u_{t-1}^* \\ w_{t-1}^* \\ h_{t-1}^* \\ u_{t-1}^c \\ u_{t-1}^{+c} \end{bmatrix} + \begin{bmatrix} \delta_1 & \delta_2 & 0 & 0 \\ 0 & 1-\delta_7 & 0 & 0 \\ 0 & \delta_3 & \delta_4 & 0 \\ 0 & \delta_5 & 0 & \delta_6 \\ 0 & \delta_7 & 0 & 0 \\ 0 & \delta_8 & 0 & 0 \end{bmatrix} \eta_t$$

where $\eta_t = \begin{bmatrix} (\gamma_1+\gamma_6) & \beta_1\alpha & 0 & 0 \\ 0 & (\gamma_2+\alpha) & 0 & 0 \\ 0 & \beta_2\alpha & (\gamma_3+\gamma_7) & 0 \\ 0 & \beta_3\alpha & 0 & (\gamma_4+\gamma_8) \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \\ \varepsilon_{4t} \end{bmatrix}$, $\delta_1 = \frac{\gamma_1}{\gamma_1+\gamma_6}$, $\delta_2 = -\alpha\beta_1\frac{\gamma_1}{\alpha\gamma_1+\alpha\gamma_6+\gamma_1\gamma_2+\gamma_2\gamma_6}$, $\delta_3 = -\alpha\beta_2\frac{\gamma_3}{\alpha\gamma_3+\alpha\gamma_7+\gamma_2\gamma_3+\gamma_2\gamma_7}$, $\delta_4 = \frac{\gamma_3}{\gamma_3+\gamma_7}$, $\delta_5 = -\alpha\beta_3\frac{\gamma_4}{\alpha\gamma_4+\alpha\gamma_8+\gamma_2\gamma_4+\gamma_2\gamma_8}$, $\delta_6 = \frac{\gamma_4}{\gamma_4+\gamma_8}$ and $\delta_7 = \frac{\alpha}{\alpha+\gamma_2}$.

- **5-equation system (y_t,u_t, w_t, h_t, l_t) with ARMA(2,2) process in** u_t^c

$$\begin{bmatrix} y_t \\ u_t \\ w_t \\ h_t \\ l_t \end{bmatrix} = \begin{bmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \\ \tau_4 \\ \tau_5 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \beta_1\phi_1 & \beta_1 \\ 0 & 1 & 0 & 0 & 0 & \phi_1 & 1 \\ 0 & 0 & 1 & 0 & 0 & \beta_2\phi_1 & \beta_2 \\ 0 & 0 & 0 & 1 & 0 & \beta_3\phi_1 & \beta_3 \\ 0 & 0 & 0 & 0 & 1 & \beta_4\phi_1 & \beta_4 \end{bmatrix} \begin{bmatrix} y_{t-1}^* \\ u_{t-1}^* \\ w_{t-1}^* \\ h_t^* \\ l_t^* \\ u_{t-1}^c \\ d_{t-1} \end{bmatrix} + \eta_t \quad (26)$$

$$\begin{bmatrix} y_t^* \\ u_t^* \\ w_t^* \\ h_t^* \\ l_t^* \\ u_t^c \\ d_t \end{bmatrix} = \begin{bmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \\ \tau_4 \\ \tau_5 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \phi_1 & 1 \\ 0 & 0 & 0 & 0 & 0 & \phi_2 & 0 \end{bmatrix} \begin{bmatrix} y_{t-1}^* \\ u_{t-1}^* \\ w_{t-1}^* \\ h_{t-1}^* \\ l_{t-1}^* \\ u_{t-1}^c \\ d_{t-1} \end{bmatrix}$$
$$+ \begin{bmatrix} \delta_1 & \delta_2 & 0 & 0 & 0 \\ 0 & \gamma_2 & 0 & 0 & 0 \\ 0 & \delta_3 & \delta_4 & 0 & 0 \\ 0 & \delta_5 & 0 & \delta_6 & 0 \\ 0 & \delta_7 & 0 & 0 & \delta_8 \\ 0 & 1-\gamma_2 & 0 & 0 & 0 \\ 0 & \theta & 0 & 0 & 0 \end{bmatrix} \eta_t$$

where $\delta_1 = \frac{\gamma_1}{\gamma_1+\gamma_6}$, $\delta_2 = \frac{1}{\gamma_1+\gamma_6}\left(\beta_1\gamma_1\gamma_2 - \beta_1\gamma_1\right)$, $\delta_3 = \frac{1}{\gamma_3+\gamma_7}\left(\beta_2\gamma_2\gamma_3 - \beta_2\gamma_3\right)$, $\delta_4 = \frac{\gamma_3}{\gamma_3+\gamma_7}$, $\delta_5 = \frac{1}{\gamma_4+\gamma_8}\left(\beta_3\gamma_2\gamma_4 - \beta_3\gamma_4\right)$, $\delta_6 = \frac{\gamma_4}{\gamma_4+\gamma_8}$,
$\delta_7 = \frac{1}{\gamma_5+\gamma_9}\left(\gamma_2\beta_4\gamma_5 - \beta_4\gamma_5\right)$, $\delta_8 = \frac{\gamma_5}{\gamma_5+\gamma_9}$ and

$$\eta_t = \begin{bmatrix} (\gamma_1+\gamma_6) & \beta_1(1-\gamma_2) & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & \beta_2(1-\gamma_2) & (\gamma_3+\gamma_7) & 0 & 0 \\ 0 & \beta_3(1-\gamma_2) & 0 & (\gamma_4+\gamma_8) & 0 \\ 0 & \beta_4(1-\gamma_2) & 0 & 0 & (\gamma_5+\gamma_9) \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \\ \varepsilon_{4t} \\ \varepsilon_{5t} \end{bmatrix}$$

- **5-equation system (y_t,u_t, w_t, h_t, l_t) with trigonometric process in u_t^c**

$$\begin{bmatrix} y_t \\ u_t \\ w_t \\ h_t \\ l_t \end{bmatrix} = \begin{bmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \\ \tau_4 \\ \tau_3 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \beta_1\rho\cos\lambda & \beta_1\rho\sin\lambda \\ 0 & 1 & 0 & 0 & 0 & \rho\cos\lambda & \rho\sin\lambda \\ 0 & 0 & 1 & 0 & 0 & \beta_2\rho\cos\lambda & \beta_2\rho\sin\lambda \\ 0 & 0 & 0 & 1 & 0 & \beta_3\rho\cos\lambda & \beta_3\rho\sin\lambda \\ 0 & 0 & 0 & 0 & 1 & \beta_4\rho\cos\lambda & \beta_4\rho\sin\lambda \end{bmatrix} \begin{bmatrix} y_{t-1}^* \\ u_{t-1}^* \\ w_{t-1}^* \\ h_{t-1}^* \\ l_{t-1}^* \\ u_{t-1}^c \\ u_{t-1}^{+c} \end{bmatrix} + \eta_t$$

$$
\begin{bmatrix} y_t^* \\ u_t^* \\ w_t^* \\ h_t^* \\ l_t^* \\ u_t^c \\ u_t^{+c} \end{bmatrix} = \begin{bmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \\ \tau_4 \\ \tau_5 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \rho\cos\lambda & \rho\sin\lambda \\ 0 & 0 & 0 & 0 & 0 & -\rho\sin\lambda & \rho\cos\lambda \end{bmatrix} \begin{bmatrix} y_{t-1}^* \\ u_{t-1}^* \\ w_{t-1}^* \\ h_{t-1}^* \\ l_{t-1}^* \\ u_{t-1}^c \\ u_{t-1}^{+c} \end{bmatrix} +
$$

$$
\begin{bmatrix} \delta_1 & \delta_2 & 0 & 0 & 0 \\ 0 & 1-\delta_9 & 0 & 0 & 0 \\ 0 & \delta_3 & \delta_4 & 0 & 0 \\ 0 & \delta_5 & 0 & \delta_6 & 0 \\ 0 & \delta_7 & 0 & 0 & \delta_8 \\ 0 & \delta_9 & 0 & 0 & 0 \\ 0 & \delta_9 & 0 & 0 & 0 \end{bmatrix} \eta_t
$$

where $\delta_1 = \frac{\gamma_1}{\gamma_1+\gamma_6}$, $\delta_2 = -\alpha\beta_1\frac{\gamma_1}{\alpha\gamma_1+\alpha\gamma_6+\gamma_1\gamma_2+\gamma_2\gamma_6}$, $\delta_3 = -\alpha\beta_2\frac{\gamma_3}{\alpha\gamma_3+\alpha\gamma_7+\gamma_2\gamma_3+\gamma_2\gamma_7}$, $\delta_4 = \frac{\gamma_3}{\gamma_3+\gamma_7}$, $\delta_5 = -\alpha\beta_3\frac{\gamma_4}{\alpha\gamma_4+\alpha\gamma_8+\gamma_2\gamma_4+\gamma_2\gamma_8}$, $\delta_6 = \frac{\gamma_4}{\gamma_4+\gamma_8}$, $\delta_7 = -\alpha\beta_4\frac{\gamma_5}{\alpha\gamma_5+\alpha\gamma_9+\gamma_2\gamma_5+\gamma_2\gamma_9}$, $\delta_8 = \frac{\gamma_5}{\gamma_5+\gamma_9}$, $\delta_9 = \frac{\alpha}{\alpha+\gamma_2}$, and

$$
\eta_t = \begin{bmatrix} \gamma_1+\gamma_6 & \beta_1\alpha & 0 & 0 & 0 \\ 0 & \gamma_2+\alpha & 0 & 0 & 0 \\ 0 & \beta_2\alpha & \gamma_3+\gamma_7 & 0 & 0 \\ 0 & \beta_3\alpha & 0 & \gamma_4+\gamma_8 & 0 \\ 0 & \beta_4\alpha & 0 & 0 & \gamma_5+\gamma_9 \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \\ \varepsilon_{4t} \\ \varepsilon_{5t} \end{bmatrix}.
$$

References

[1] Alogoskoufis, G.S. & Stengos, T. (1991). Testing for Nonlinear Dynamics in Historical Unemployment Series. *European University Institute Working Paper*, vol. ECO 91/38.

[2] Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19, 716–723.

[3] Amisano, G., & Serati, M. (1999). Forecasting cointegrated series with BVAR models. *Journal of Forecasting*, 18, 463-476.

[4] Anderson, H.M., Low C.N. & Snyder, R.D. (2006). Single source of error state space approach to the Beveridge Nelson decomposition. *Economics Letters,* 91, 104-109.

[5] Artis, M.J. & Zhang, W. (1990). BVAR forecasts for the G-7. *International Journal of Forecasting*, 6, 349-362.

[6] Attfield C. L.F. & Silverstone B. (1998). Okun's law, cointegration and gap variables. *Journal of Macroeconomics*, 20, 625-637.

[7] Beveridge, S. & Nelson, C.R. (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*, 7, 151-174.

[8] de Silva, A., Hyndman, R.J. & Snyder, R.D. (2008). Vector Exponential Smoothing. In: R.J. Hyndman et al, Forecasting with exponential smoothing: the state space approach, Berlin: Springer-Verlag, pp. 287-299.

[9] De Gooijera, J.G. & Hyndman, R.J. (2006). 25 years of time series forecasting. *International Journal of Forecasting*, 22, 443-473.

[10] Gruen, D., Pagan, A. & Thompson, C. (1999). The Phillips curve in Australia. *Journal of Monetary Economics*, 44, 223-258.

[11] Hamilton J.D. (2001), A parametric approach to flexible nonlinear inference. *Econometrica*, 69, 537–573.

[12] Harvey, A.C (1989). *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press.

[13] Harvey, A.C., Henry, B., Peters, S. & Wren-Lewis, S. (1986). Stochastic trends in dynamic regression models: an application to the employment-output equation. *The Economic Journal*, 96, 975 986.

[14] Hyndman, R. J., Koehle A. B., Ralph D. S. & Grose S.(2002). A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of Forecasting*, 18, 439-454.

[15] Joutz, F.L., Maddala, G.S. & Trost, R.P. (1995). An integrated Bayesian Vector Autoregression and error correction model for forecasting electricity, consumption, and prices. *Journal of Forecasting*, 14, 287-310.

[16] Koopman, S. J., Lee, K. M. & Wong S. Y. (2005) Trend-cycle decomposition models with smooth-transition parameters: evidence from US economic time series. In D. van Dijk, C. Milas and P.A. Rothman (eds), *Nonlinear Time Series Analysis of Business Cycles*, Elsevier.

[17] LeSage, J.P. (1990). A comparison of the forecasting ability of ECM and VAR models. *Review of Economics and Statistics*, 72, 664-671.

[18] Litterman, R.B. (1986). Forecasting with Bayesian vector autoregressions — five years of experience. *Journal of Business and Economic Statistics*, 4, 25-38.

[19] King, R.G. & Watson, M.W. (1995). The Postwar U.S. Phillips Curve: A revisionist econometric history. *Carnegie-Rochester Conference Series on Public Policy*, 41, 157-219.

[20] Kuttner, K.N. (1994). Estimating Potential Output as a Latent Variable. *Journal of Business and Economic Statistics*, 12, 361-368.

[21] Malley, J. & Molana H. (2008). Output, Unemployment and Okun's Law: Some Evidence from the G7. *Economics Letters,* 101, 113-115.

[22] Okun, A. M. (1962). Potential GDP: Is Measurement and Significance. *American Statisitical Association*, *Proceedings of the Business and Economics Statistics Section,* 98-104.

[23] Ord, J. K., Koehler, A. B. & Snyder, R. D. (1997). Estimation and Prediction for a Class of Dynamic Nonlinear Statistical Models. *Journal of the American Statistical Association*, 92, 1621-1629.

[24] Phillips, A. W. (1958). The Relationship between Unemployment and the Rate of Change of Money Wages in the United Kingdom 1861-1957. *Economica*, 25, 283–299.

[25] Ramos, F.F.R. (2003). Forecasts of market shares from VAR and BVAR models: a comparison of their accuracy. *International Journal of Forecasting*, 19, 95-110.

[26] Rapach, D.E., Wohar, M.E. & Rangvid, J. (2004). Macro variables and international stock return predictability. *International Journal of Forecasting*, 21, 137-166.

[27] Shoesmith, G.L. (1992). Cointegration, error correction, and improved regional VAR forecasting. *Journal of Forecasting*, 11, 91-109.

[28] Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6, 461-464.

[29] Taylor, J.W. (2008). Exponentially Weighted Information Criteria for Selecting Among Forecasting Models. *International Journal of Forecasting*, 24, 513-524.