Finnish Economic Papers – Volume 23 – Number 1 – Spring 2010

IMPROVING UNEMPLOYMENT RATE FORECASTS USING SURVEY DATA*

PÄR ÖSTERHOLM

National Institute of Economic Research,
Box 3116, 103 62 Stockholm, Sweden;
e-mail: par.osterholm@konj.se

This paper investigates whether forecasts of the Swedish unemployment rate can be improved by using business or household survey data. We conduct a simulated out-of-sample forecast exercise in which the performance of a Bayesian VAR model with only macroeconomic variables is compared to that when the model also includes variables based on survey data. Results show that the forecasting performance at short horizons can be improved. The improvement is largest when forward-looking variables based on data from the manufacturing industry are employed. (JEL: E17, E24, E27)

1. Introduction

The aggregate unemployment rate is a variable of fundamental interest to many economic agents. For example, for monetary policy makers, it both serves as an indicator of the macroeconomy in general and carries information regarding inflationary pressure. For fiscal policy makers, the unemployment rate is linked to government expenditure as well as income due to its relationship with, for example, unemployment benefits and income taxes. There is accordingly a widespread interest in being able to generate good forecasts of the unemployment rate. The literature dealing with unemployment rate forecasting is consequently large; for a number of different applications and methodological choices, see Funke (1992), Rothman (1998), Franses *et al.* (2004), Golan and Perloff (2004), Milas and Rothman (2008) and Gustavsson and Österholm (2010).

The purpose of this paper is to investigate whether Swedish unemployment rate forecasts can be improved by using business or household survey data. The National Institute of Economic Research (Konjunkturinstitutet; henceforth NIER) conducts the *Economic Tendency Survey* in which both businesses and households are asked questions which potentially could be useful for forecasting the aggregate unemployment rate. We assess the usefulness of these data by employing them in a simulated out-of-sample forecast exercise. Specifically, we compare the

* *I am grateful to the editor Heikki Kauppi, an anonymous referee, Maria Billstam, Joakim Skalin and seminar participants at the National Institute of Economic Research for valuable comments and useful discussions.*

simulated out-of-sample forecasting performance of a Bayesian VAR model with only macroeconomic data to that when the model has had variables based on survey data added to it. This is a fairly common approach to examine Granger causality empirically which has been used in a range of applications – see, for example, Thoma and Gray (1998), Chao *et al.* (2001), Hale and Jorda (2007) and Berger and Österholm (2009) – and we believe that it is a reasonable methodological choice.[1]

The predictive power of survey data for the real economy has been investigated in a number of studies, both in- and out-of-sample; see, for example, Carroll *et al.* (1994), Ludvigson (2004), Hansson *et al.* (2005), Cotsomitis and Kwan (2006) and Kwan and Cotsomitis (2006). Often the survey data are expected to work as a leading indicator for the real variable in question. The forward-looking nature of some of the questions in the *Economic Tendency Survey* might therefore be particularly interesting. We employ five variables in this paper, each of which is based on data from a particular question in the survey. Households are asked what their expectations are regarding the development of the unemployment rate over the coming twelve months. Businesses are asked both whether the number of employees has increased, remained unchanged or decreased over the last three month period and what the outlook is for the coming three months. Since the survey data might incorporate information that is not reflected in other variables that are typically included in forecasting models, it does not seem unreasonable that forecasting performance could be improved by the inclusion of the survey data.

Our results suggest that both business and household survey data have predictive power for the Swedish unemployment rate. However, the usefulness of the data varies. The data describing the contemporaneous situation in businesses do not seem particularly valuable for forecasting. The forward-looking variables, on the other hand, all appear to have predictive power for the unemployment rate at short horizons, with the improvement being largest when data from the manufacturing industry are used.

The remainder of this paper is organised as follows. Section 2 presents the survey data in some detail. In Section 3, we present the modelling framework and Section 4 presents the results from the simulated out-of-sample forecast exercise. Finally, Section 5 concludes.

2. Data

The NIER each quarter asks representatives of Swedish firms and households about the present situation and the outlook for the near future. The information is compiled in the *Economic Tendency Survey* whose purpose is to be a quickly available source of indicators pertaining to outcome, present situation and expectations for important economic variables.

Stratified sampling of firms takes place through the business register of Statistics Sweden. More than 8000 companies are included in the survey and they are divided into four categories: manufacturing industry, construction industry, retail trade and private service sector. The questionnaires are addressed to upper management and are designed to be filled out conveniently and quickly. For example, a company can be asked whether the inflow of new orders from abroad has increased, remained unchanged or decreased. Some questions refer to the outcome for the past three months, others to expectations or plans for the coming three months. Household data are obtained through telephone interviews with a random net sample of 1500 individuals between 16 and 84 years of age. The questions asked refer both to the household's own economic situation and the aggregate economy. For each question, the responses are standardised so that the percentages of the response alternatives add up to 100. To facilitate presentation and analyses of outcomes, the concept "net figures" are employed, where a net figure is the difference between the percentage of respond-

[1] *An alternative is to rely on in-sample evidence. However, in-sample evidence is likely to lead to models that are overfitted and therefore do not forecast as well as the in-sample evidence suggests. The issue of in-sample overfitting is analysed and discussed in, for example, Hansen (1999) and Clark (2004). At the same time, it should of course be noted that simulated out-of-sample forecasts is not a flawless methodological choice either; see, for example, Inoue and Kilian (2006) for a discussion.*

ents reporting an increase and a decrease for a certain question.[2] For example, if 45 percent of respondents state that there has been an increase, 25 percent that there has been no change and 30 percent that there has been a decrease, the net figure is 45–30=15.

In this paper, we investigate whether the data from the *Economic Tendency Survey* can be used to improve forecasts of the Swedish unemployment rate. The benchmark model – a Bayesian VAR which we will return to in the next section – only makes use of three variables: the unemployment rate, CPI inflation and the three month treasury bill rate. VARs with these variables are standard tools in macroeconomic analysis and forecasting – see, for example, Cogley and Sargent (2001, 2005), Primiceri (2005) and Ribba (2006) – and a model based on these three variables therefore seems like a reasonable benchmark. Five questions in the survey were identified as particularly interesting to augment the benchmark model with.[3] Questions 116 and 207 are used for the manufacturing industry and 106 and 204 for the construction industry. Specifically, in questions 116 and 106, the respondents are asked whether the firm's number of employees has increased, remained unchanged or decreased during the last three months. This yields variables for the contemporaneous situation in the manufacturing and construction industries, s_t^{mc} and s_t^{cc} respectively. The expectations regarding the number of employees for the coming three months are addressed in questions 207 and 204; these forward-looking variables are denoted s_t^{mf} and s_t^{cf} for the manufacturing and construction industries respectively. For households, finally, question 7 is used. This asks how the household expects the aggregate unemployment rate to develop over the coming twelve months and the variable based on it is denoted s_t^{hf}.[4] The sample available is 1980Q1 to 2008Q4 and all data are given in Figure 1. Worth noting here is the effect that the financial turbulence had on expectations during 2008. For example, the net figure for households' expectations was -12 in 2008Q1; that is, it was more common among households to expect a falling unemployment rate than an increasing one. However, by 2008Q4 the net figure was -76, indicating that a vast majority of households was expecting the unemployment rate to increase.

3. Model

For the empirical analysis in this paper, we will rely on Bayesian VAR models. The Bayesian VAR is typically considered a good forecasting tool and has been shown to forecast well out-of-sample in a number of studies; see, for example, Doan *et al.* (1984), Litterman (1986), Adolfson *et al.* (2007) and Villani (2009).[5] Specifically, the forecasting model used in this paper is given by

$$(1) \quad G(L)x_t = \delta + \theta D_t + \eta_t,$$

where $G(L) = I - G_1 L - \ldots - G_m L^m$ is a lag polynomial of order m, x_t is an nx1 vector of economic variables, δ is an nx1 vector of intercepts and η_t is an nx1 vector of *iid* error terms fulfilling $E(\eta_t) = 0$ and $E(\eta_t \eta_t') = \Sigma$. D_t is a dummy variable that takes on the value 1 between 1980Q1 and 1992Q4. This is included to account for the different monetary policy regimes during the sam-

[2] *This is a very common way to summarise survey data. In applied econometric work, a similar approach has been used by Carabenciov et al. (2008).*

[3] *The survey contains other questions that could be of interest. Some of these, however, cannot be used since long enough time series are not available.*

[4] *Five response alternatives are available for this question: increase substantially, increase somewhat, roughly remain unchanged, decrease somewhat and decrease substantially. The net figure is still calculated as the difference between the percentage of respondents reporting an increase and a decrease – that is, it is given by (increase substantially + increase somewhat)-(decrease substantially + decrease somewhat).*

[5] *A problem when VARs are estimated with classical methods is that the number of parameters typically is large but the number of observations small. This means that parameter estimates tend to be associated with a fairly large degree of uncertainty and low precision in parameter estimates hurts forecasting performance. In the Bayesian VAR, this problem is addressed by providing prior information regarding the values of the coefficients of the model. By increasing the precision in parameter estimates forecasting performance is improved. See, for example, Sims (1980), Doan et al. (1984), Litterman (1986) and Villani (2009) for discussions regarding these issues.*

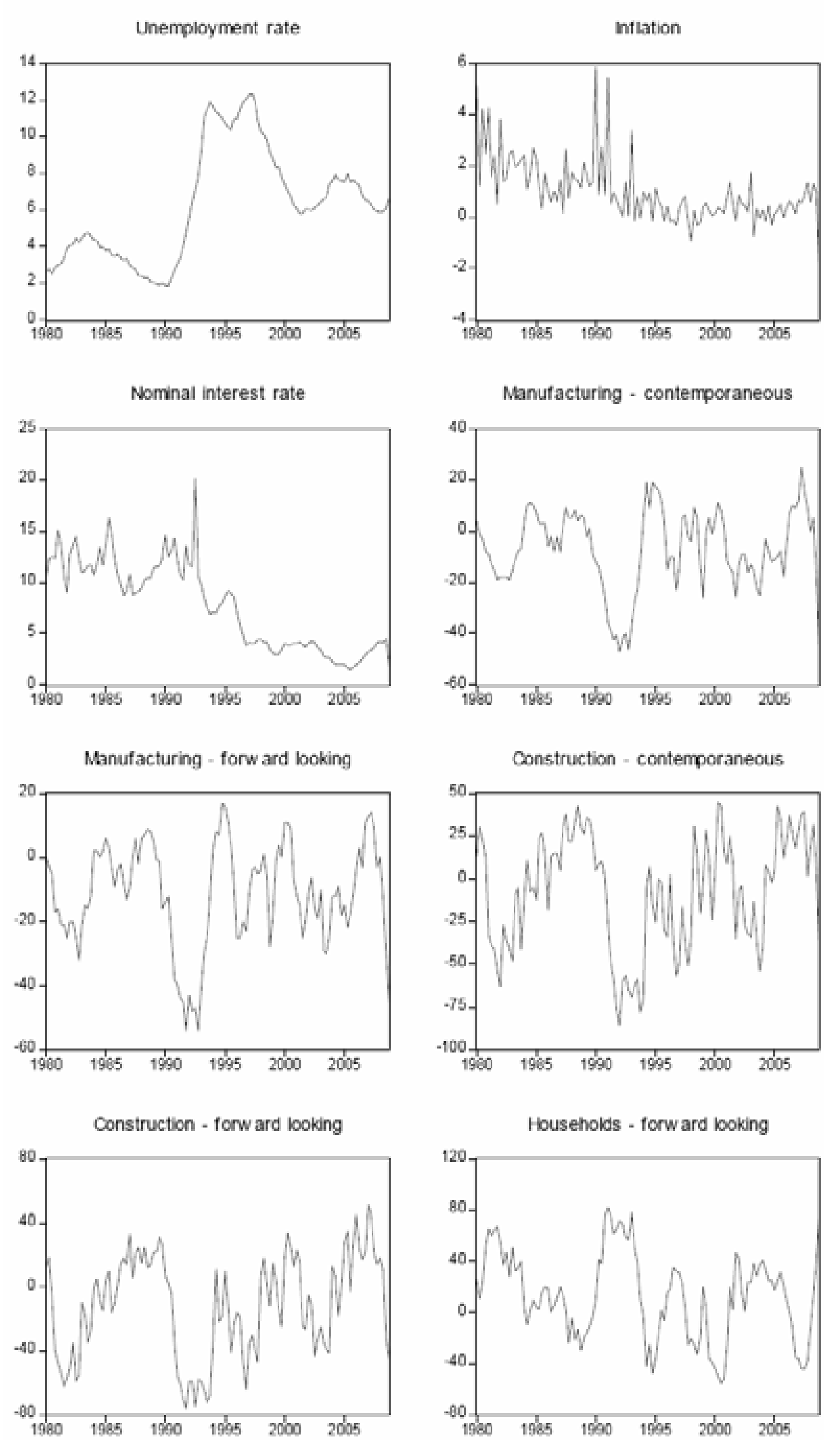

Figure 1. Data
Notes: Unemployment rate, inflation and nominal interest rate are given in percent. Remaining series are net figures.

ple. While the shift to inflation targeting did not necessarily affect the time-series properties of the unemployment rate or the dynamic relationship between variables, it is typically assumed that it generated a level shift for both inflation and the nominal interest; see, for example, Adolfson *et al.* (2007) and Österholm (2008).

In all models, the lag length is set to $m=4$. Regarding the dynamics of the model, the prior on $vec(G)$, where $G = (G_1 \dots G_4)$, is given by $vec(G) \sim N_{4n^2}(\lambda_G, \Omega_G)$ and is of Minnesota style. The Minnesota prior takes its starting point in the observation that a univariate random walk (potentially with drift) is a reasonably good forecasting model for the level of many macroeconomic variables. Variables in levels (first differences) accordingly have a prior mean on the coefficient on the first own lag of one (zero); all other coefficients in G_i have a prior mean of zero. The tightness around the mean is adjusted through a number of hyperparameters. In line with the univariate random walk benchmark, coefficients on long lags have smaller variances than those on short lags and the variances on cross-lag coefficients are smaller than those on own-lag coefficients. This is achieved by setting the overall tightness to 0.2, cross-equation tightness to 0.5 and choosing a lag decay parameter of 1.[6] While Bayesian econometricians typically prefer to use informative priors, it is not always easy to have an opinion *ex ante*. This is true for the intercepts and coefficients on the dummy variable in equation (1). Following the standard in the literature, diffuse priors for $\Psi = (\delta \;\; \theta)$ are accordingly used and are given by $vec(\Psi) \sim N_{2n}(\lambda_\Psi, \Omega_\Psi)$. Finally, the prior for the covariance matrix is a mainstream diffuse prior, $p(\Sigma) \propto |\Sigma|^{-(n+1)/2}$. The numerical evaluation of the posterior distribution is conducted using the Gibbs sampler – see, for example, Tierny (1994) for a technical discussion – with the number of draws set to 10000. Regarding the forecasts from the models, these are generated in a straightforward manner. At each point in time and for every draw from the posterior distribution, a sequence of shocks is drawn and used to generate future data. In this manner we generate 10000 paths of the unemployment rate; the median forecast from this predictive density is used as point estimate.[7]

As indicated above, the benchmark model is a trivariate Bayesian VAR with the unemployment rate (u_t), CPI inflation (Δp_t) and the three month treasury bill rate (i_t). We accordingly set $x_t = (u_t \;\; \Delta p_t \;\; i_t)'$ in the benchmark model. When survey data are employed, the model is extended with variables based on survey data. For example, when the usefulness of the household survey data is investigated, we set $x_t = (u_t \;\; \Delta p_t \;\; i_t \;\; s_t^{hf})'$.

4. Results

The simulated out-of-sample forecasts are generated the following way: We initially estimate the models using data from 1980Q1 to 1999Q4 and then use the estimated models to generate forecasts of the unemployment rate four quarters ahead. We then extend that sample one period, re-estimate the models and generate new forecasts four quarters ahead and so on.[8] The last evaluation is conducted on a model estimated from 1980Q1 to 2008Q3 and forecasted one period ahead.

We evaluate the forecasting performance of the competing models by comparing the root mean square forecast errors (RMSFEs) at the different forecasting horizons. For a convenient presentation of the results, we show the *relative RMSFEs*. The relative RMSFE at horizon h is here defined as

$$(2) \quad RR_h = RMSFE_{S,h} / RMSFE_{NS,h}$$

where $RMSFE_{S,h}$ and $RMSFE_{NS,h}$ are the RMSFEs of the model with survey data and without survey data respectively. A relative RMSFE smaller than one accordingly means that the model with

[6] *These are values recommended by Doan (1992) which are commonly used in empirical work; see, for example, Villani (2009). See Litterman (1986) for a discussion regarding the hyperparameters.*

[7] *Since the RMSFE is the evaluation criterion used in this paper, it can be argued that the mean should be used rather than the median. Since the predictive densities are very close to symmetric though, the median provides a convenient approximation to the mean.*

[8] *That is, it is a recursive simulated out-of-sample forecast exercise.*

survey data outperforms the model without survey data. We choose to focus on (relative) RMSFEs and use them as criterion for assessing performance. In line with the philosophy expressed in Armstrong (2007), we do not aim to assess whether the difference in forecasting performance is significant.

The usefulness of the five survey data variables is initially investigated by comparing the RMSFEs of the five four-variate BVARs to that of the benchmark trivariate BVAR. The relative RMSFEs are given in Table 1 and RMSFEs of all models are given in Table A1 in the Appendix. As can be seen, augmenting the model with s_t^{cc} or s_t^{mc} tends to push the relative RMSFE above unity; only at the one-quarter horizon when s_t^{mc} is used is it smaller than one. One explanation for this finding is that even though the BVAR reduces the problem of over-parameterisation, it is not immune to it. As variables with only a moderate additional informational content are added, this does not outweigh the cost in terms of lost precision in estimation.

Turning to the forward-looking variables instead, the relative RMSFE is smaller than one at all horizons for both s_t^{cf} and s_t^{mf}. For s_t^{hf}, it is smaller than unity only at the one- and two-quarter horizons. Improvements using s_t^{cf} and s_t^{hf} are small though and indicate a reduction in the RMSFE of only a few percent at most. Relying on s_t^{mf}, on the other hand, the improvement in forecast accuracy is larger – at the one- and two-quarter horizons, the RMSFE is reduced by more than ten percent.

The results so far indicate that the three forward-looking variables individually have predictive power for the unemployment rate. An empirical strategy that might pay off is of course to include more than one of these variables at a time in the system. The forecasting performance of four additional models is accordingly investigated, reflecting the three combinations of two forward-looking variables at a time, and the model with all three variables included at the same time. Results from this exercise show that the relative RMSFE from these four models all are below one at the one- to three-quarter horizons. But while all four models constitute improvements over the benchmark trivariate BVAR, the forecasting performance is not as good as for the four-variate model where $x_t = (u_t \ \ \Delta p_t \ \ i_t \ \ s_t^{mf})'$.[9]

Table 1. Relative RMSFEs for estimated Bayesian VAR models

	Horizon in quarters			
	1	**2**	**3**	**4**
$x_t = (u_t \ \Delta p_t \ i_t \ s_t^{cc})'$	1.01	1.07	1.06	1.06
$x_t = (u_t \ \Delta p_t \ i_t \ s_t^{mc})'$	0.96	1.02	1.05	1.13
$x_t = (u_t \ \Delta p_t \ i_t \ s_t^{cf})'$	0.96	0.97	0.97	0.99
$x_t = (u_t \ \Delta p_t \ i_t \ s_t^{mf})'$	0.88	0.84	0.91	0.94
$x_t = (u_t \ \Delta p_t \ i_t \ s_t^{hf})'$	0.98	0.97	1.00	1.03
$x_t = (u_t \ \Delta p_t \ i_t \ s_t^{cf} \ s_t^{mf})'$	0.89	0.85	0.92	0.95
$x_t = (u_t \ \Delta p_t \ i_t \ s_t^{cf} \ s_t^{hf})'$	0.96	0.94	0.97	1.01
$x_t = (u_t \ \Delta p_t \ i_t \ s_t^{mf} \ s_t^{hf})'$	0.91	0.88	0.95	1.00
$x_t = (u_t \ \Delta p_t \ i_t \ s_t^{cf} \ s_t^{mf} \ s_t^{hf})'$	0.92	0.89	0.96	1.00

Note: u_t is the unemployment rate, Δp_t is CPI inflation and i_t the three month treasury bill rate. s_t^{cc} and s_t^{mc} describe the contemporaneous situation regarding the number of employees in the construction and manufacturing industries respectively; s_t^{cf} and s_t^{mf} describe the corresponding expectations. s_t^{hf} describes households' expectations concerning the aggregate unemployment rate. The relative RMSFE is calculated using the trivariate model with only macroeconomic variables as the benchmark model. A relative RMSFE smaller than one indicates that the model augmented with survey data performs better than the benchmark model.

[9] *Using combinations of forward-looking variables and s_t^{mc} was not a particularly successful strategy either in terms of reducing the RMSFE. In no case was forecasting performance better than when four-variate model including s_t^{mf} as the only survey data variable was used. Results are not reported but are available upon request.*

It can finally be noted that the variables based on survey data also all appear to have reasonable properties in the system. Illustrative impulse response functions for two cases are shown in Figures A1 and A2 in the Appendix. Orthogonal shocks, ε_t, are identified using the reduced-form shocks according to $\varepsilon_t = P^{-1}\eta_t$ where P is obtained from the standard Cholesky decomposition of Σ as $\Sigma = PP'$ and the variables are ordered as in x_t. Most importantly, a shock to a business survey variable – that is, a positive shock to employment or employment plans – decreases future unemployment; see Figure A1. In a similar fashion, a shock to the household survey variable – that is, an expectation about a higher unemployment rate in the future – raises future unemployment; see Figure A2. This analysis is of course not structural by any means. However, the survey data series could be useful also in empirical analysis using structural VARs. The importance of the information set for correct identification of structural shocks is undisputable. For example, the so called price puzzle – that is, a rise in the price level in response to a contractionary monetary policy shock – has been shown to be alleviated or disappear when appropriate forward-looking variables are included in the model; see, for example, Sims (1992), Christiano *et al.* (1999) and Hanson (2004).[10] It seems likely that the survey data employed in this paper can be used for similar purposes.

5. Conclusions

In this paper, it has been investigated whether Swedish unemployment rate forecasts can be improved by using business and household survey data. Results show that several of the survey data variables employed have predictive power for the Swedish unemployment rate at short horizons. Though improvements are not dramatic, it is apparent that data from the *Economic Tendency Survey* in a fruitful way can be incorporated in macroeconomic forecasting models. That our results support the use of forward-looking variables is not surprising; it seems reasonable that these variables contain more information that is not already reflected in the other variables included in the forecasting model.

Our findings are obviously model dependent but nevertheless suggest that survey data can be useful when forecasting the real economy. That the information from the manufacturing industry – rather than the construction industry – has the largest value added in a macroeconomic model is also interesting to note. Discussions regarding which sector of the economy carries the largest informational content from a forecasting point of view are common among forecasters and policymakers. Being able to focus on the most relevant indicator(s) for a certain variable is useful, regardless of whether forecasts are generated by econometric models or judgemental methods.

[10] *Similar issues also arise in analysis of fiscal policy; see, for example, Yang (2007).*

References

Adolfson, M., M.K. Andersson, J. Lindé, M. Villani, and A. Vredin (2007). "Modern Forecasting Models in Action: Improving Macro Economic Analyses at Central Banks." *International Journal of Central Banking* 3, 111–144.

Armstrong, J. S. (2007). "Significance Tests Harm Progress in Forecasting." *International Journal of Forecasting* 23, 321–327.

Berger, H., and P. Österholm (2009). "Does Money Still Matter for U.S. Output?" *Economics Letters* 102, 143–146.

Carabenciov, I., I. Ermolaev, C. Freedman, M. Juillard, O. Kamenik, D. Korshunov, D. Laxton, and J. Laxton (2008). A Small Quaterly Multi-Country Projection Model with Financial-Real Linkages and Oil Prices. IMF Working Paper 08/280, International Monetary Fund.

Chao, J., C. Valantina, and N. Swanson (2001). "An Out-of-Sample Test for Granger Causality." *Macroeconomic Dynamics* 5, 598–620.

Christiano, L.J., M. Eichenbaum, and C. Evans (1999). "Monetary Policy Shocks: What Have We Learned and to What End?" In *Handbook of Macroeconomics*, vol. 1A. Eds. M. Woodford, and J.B. Taylor. Elsevier, Amsterdam.

Clark, T.E. (2004). "Can Out-of-Sample Forecast Comparisons Help Prevent Overfitting?" *Journal of Forecasting* 23, 115–139.

Cogley, T., and T.J. Sargent (2001). "Evolving Post-World War II U.S. Inflation Dynamics." *NBER Macroeconomics Annual* 16, 331–373.

Cogley, T., and T.J. Sargent (2005). "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII US." *Review of Economic Dynamics* 8, 262–302.

Cotsomitis, J.A., and A.C.C. Kwan (2006). "Can Consumer Confidence Forecast Household Spending? Evidence from the European Commission Business and Consumer Surveys." *Southern Economic Journal* 72, 597–610.

Diebold, F.X., and R.S. Marino (1995). "Comparing Predictive Accuracy." *Journal of Business and Economic Statistics* 13, 253–263.

Doan, T.A. (1992). *RATS Users Manual*, Version 4.

Doan, T.A., R.B. Litterman, and C.A. Sims (1984). "Forecasting and Conditional Projection Using Realistic Prior Distributions." *Econometric Reviews* 3, 1–100.

Franses, P.H., R. Paap, and B. Vroomen (2004). "Forecasting Unemployment Using an Autoregression with Censored Latent Effects Parameters." *International Journal of Forecasting* 20, 255–271.

Funke, M. (1992). "Time-Series Forecasting of the German Unemployment Rate." *Journal of Forecasting* 11, 111–125.

Golan, A., and J.M. Perloff (2004). "Superior Forecasts of the U.S. Unemployment Rate Using a Nonparametric Method." *Review of Economics and Statistics* 86, 433–438.

Gustavsson, M., and P. Österholm (2010). "The Presence of Unemployment Hysteresis in the OECD: What Can We Learn from Out-of-Sample Forecasts?" *Empirical Economics* 38, 779–792.

Hale, G., and Ó. Jordà (2007). "Do Monetary Aggregates Help Forecast Inflation?" *Federal Reserve Bank of San Fransisco Economic Letter* 2007–10, 1–3.

Hansen, B.E. (1999). "Discussion of 'Data mining reconsidered'." *Econometrics Journal* 2, 192–201.

Hanson, M.S. (2004). "The 'Price Puzzle' Reconsidered." *Journal of Monetary Economics* 51, 1385– 1413.

Hansson, J., P. Jansson and M. Löf (2005). "Business Survey Data: Do They Help in Forecasting GDP Growth?" *International Journal of Forecasting* 21, 377–389.

Inoue, A., and L. Kilian (2006). "On the Selection of Forecasting Models." *Journal of Econometrics* 130, 273–306.

Kwan, A.C.C., and J.A. Cotsomitis (2006). "The Usefulness of Consumer Confidence in Forecasting Household Spending in Canada: A National and Regional Analysis." *Economic Inquiry* 44, 185–197.

Litterman, R.B. (1986). "Forecasting with Bayesian Vector Autoregressions – Five Years of Experience." *Journal of Business and Economic Statistics* 5, 25–38.

Ludvigson, S. (2004). "Consumer Confidence and Consumer Spending." *Journal of Economic Perspectives* 18, 29–50.

Milas, C., and P. Rothman (2008). "Out-of-Sample Forecasting of Unemployment Rates with Pooled STVECM Forecasts." *International Journal of Forecasting* 24, 101–121.

Österholm, P. (2008). "Can Forecasting Performance Be Improved by Considering the Steady State? An Application to Swedish Inflation and Interest Rate." *Journal of Forecasting* 27, 41–51.

Primiceri, G. (2005). "Time Varying Structural Vector Autoregressions and Monetary Policy." *Review of Economic Studies*, 72, 821–852.

Ribba, A. (2006). "The Joint Dynamics of Inflation, Unemployment and Interest Rate in the United States Since 1980." *Empirical Economics* 31, 497–511.

Rothman, P. (1998). "Forecasting Asymmetric Unemployment Rates." *Review of Economics and Statistics* 80, 164–168.

Sims, C.A. (1980). "Macroeconomics and Reality." *Econometrica* 48, 1–48.

Sims, C.A. (1992). "Interpreting the Macroeconomic Time Series Facts: The Effects of Monetary Policy." *European Economic Review* 36, 975–1000.

Thoma, M.A., and J.A. Gray (1998). "Financial Market Variables Do Not Predict Real Activity." *Economic Inquiry* 36, 522–539.

Tierny, L. (1994). "Markov Chains for Exploring Posterior Distributions." *Annals of Statistics* 22, 1701–1762.

Villani, M. (2009). "Steady State Priors for Vector Autrogregressions." *Journal of Applied Econometrics* 24, 630–650.

Yang, S.-C.S. (2007). "Tentative Evidence of Tax Foresight." *Economics Letters* 96, 30–37.

Appendix

Table A1. RMSFEs for estimated Bayesian VAR models

	Horizon in quarters			
	1	**2**	**3**	**4**
$x_t = (u_t\ \Delta p_t\ i_t)'$	0.225	0.343	0.467	0.552
$x_t = (u_t\ \Delta p_t\ i_t\ s_t^{cc})'$	0.228	0.367	0.493	0.586
$x_t = (u_t\ \Delta p_t\ i_t\ s_t^{mc})'$	0.216	0.350	0.493	0.623
$x_t = (u_t\ \Delta p_t\ i_t\ s_t^{cf})'$	0.216	0.331	0.451	0.545
$x_t = (u_t\ \Delta p_t\ i_t\ s_t^{mf})'$	0.198	0.287	0.427	0.518
$x_t = (u_t\ \Delta p_t\ i_t\ s_t^{hf})'$	0.220	0.331	0.466	0.571
$x_t = (u_t\ \Delta p_t\ i_t\ s_t^{cf}\ s_t^{mf})'$	0.200	0.293	0.430	0.524
$x_t = (u_t\ \Delta p_t\ i_t\ s_t^{cf}\ s_t^{hf})'$	0.217	0.324	0.453	0.559
$x_t = (u_t\ \Delta p_t\ i_t\ s_t^{mf}\ s_t^{hf})'$	0.217	0.324	0.453	0.559
$x_t = (u_t\ \Delta p_t\ i_t\ s_t^{cf}\ s_t^{mf}\ s_t^{hf})'$	0.206	0.304	0.449	0.554

Note: u_t is the unemployment rate, Δp_t is CPI inflation and i_t the three month treasury bill rate. s_t^{cc} and s_t^{mc} describe the contemporaneous situation regarding the number of employees in the construction and manufacturing industries respectively; s_t^{cf} and s_t^{mf} describe the corresponding expectations. s_t^{hf} describes households' expectations concerning the aggregate unemployment rate.

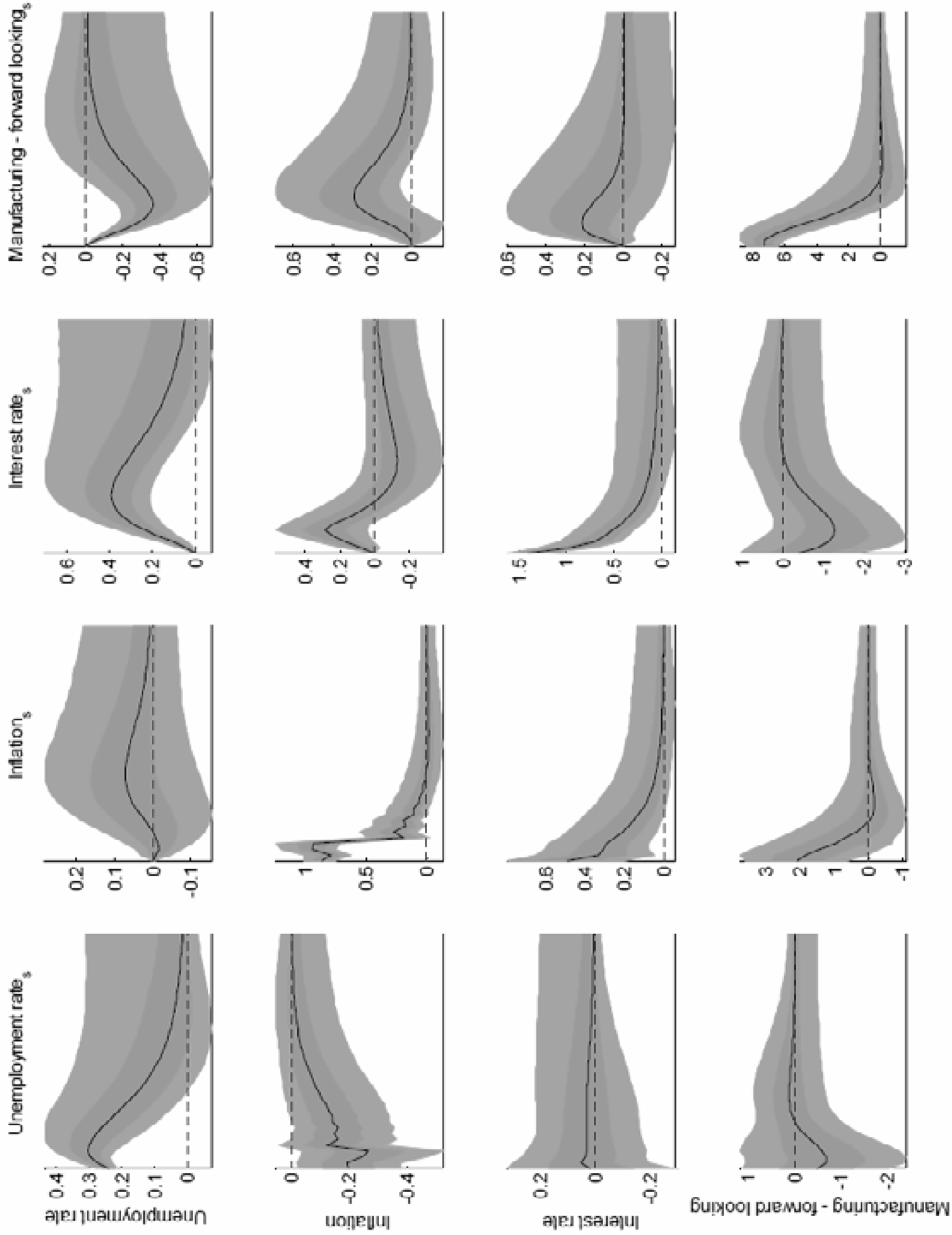

Figure A1. Impulse response functions from four-variate model using a forward-looking variable based on data from the manufacturing industry

Note: Black line is the median. Coloured bands are 68 and 95 percent confidence bands. Maximum horizon is 40 quarters.

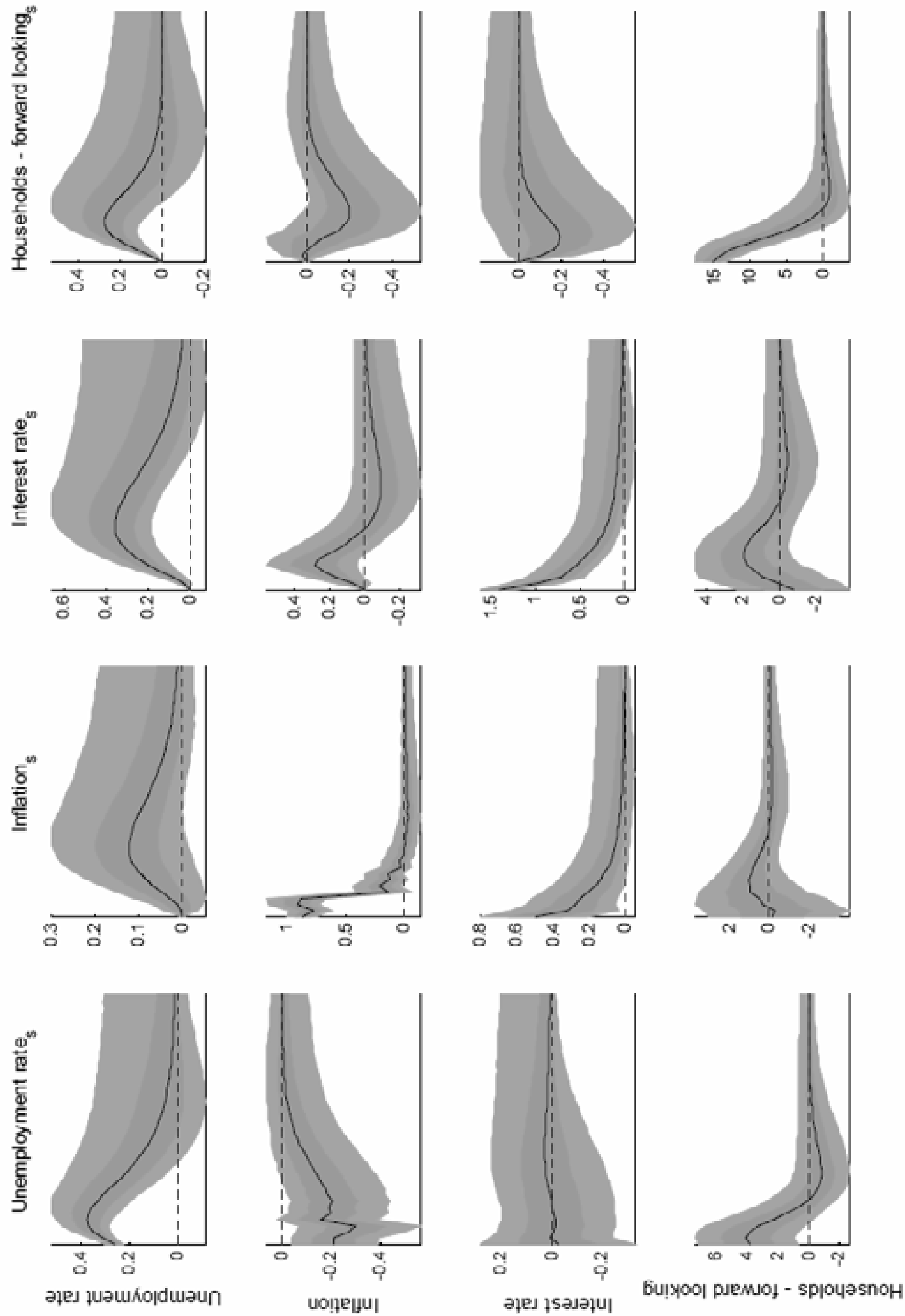

Figure A2. Impulse response functions from four-variate model using a forward-looking variable based on data from households

Note: Black line is the median. Coloured bands are 68 and 95 percent confidence bands. Maximum horizon is 40 quarters.