Uppsala University
Department of Statistics

**MASTER THESIS**

UPPSALA
UNIVERSITET

# Out of Sample Forecast of Swedish GDP Growth by the Economic Sentiment Indicator in the Euro Area

## A Bayesian Approach

Author: Erik Spånberg
Supervisor: Anders Ågren, Professor Emeritus
Study programme: Statistics

Uppsala 2015

## Abstract

In this paper, the predictive capabilities of the Economic Sentiment Indicator (ESI), based on business and consumer surveys in the Euro area, are evaluated by out-of-sample forecasts of Swedish GDP growth. A steady state Bayesian VAR-model is applied to quarterly data from 1996 to 2014. The results show that the inclusion of the ESI improves the forecasting performance, both in the point predictive measurement Root Mean Square Errors and in the forecast density sharpness measurement Log Predictive Density Scores. These findings suggest that international confidence indicators may prove useful in forecasting macroeconomic trends for small open economies.

**Keywords:** Steady State, Bayesian VAR-model, Economic Sentiment Indicator, Small open economy, Out-of-Sample forecast, Euro area

# Contents

# List of Figures

# List of Tables

# 1 Introduction

The prime role of leading economic indicators is to anticipate economic developments. Among these the confidence indicators, obtained by consumer and business surveys, are commonly used to measure economic expectations and as such they are broadly applied to assess current economic trends, like household expenditure (e.g. Bram and Ludvigson, 1998) or even GDP growth (e.g. Mourougane and Roma, 2003). The idea that national confidence indicators can provide information about changes in national economic trends is not new, however, studies about international confidence indicators as leads for national trends are scarce.

The Economic Sentiment Indicator (ESI) is an index of consumer and business confidence indicators in the Euro area. Mourougane and Roma (2003) assess the usefulness of the indicator in forecasting real GDP growth in a selected number of larger Euro-area countries (such as Germany, France et al.), which account for 90 % of the economy in the area. In most of these countries, the indicator was found useful in this regard. Golinelli and Parigi (2004) evaluate the predictive ability of national consumer indicators upon GDP in a series of larger economies (US, UK, France among others) and find evidence that they did have such abilities under some conditions. In the United States, this is in line with the results of Howrey (2001).

There seem to be evidence that national confidence indicators provide information on national GDP and that confidence indicators in an larger economy (like the Euro area) provide information on the larger economies within that area. However, the role of international confidence indicators as leads for smaller economies is not studied in the same manner. The consumer and business economic expectations in a large economy may be able to provide information about developments in small open economies.

The purpose of this study is to evaluate if the Economic Sentiment Indicator (ESI) in Euro area Granger-cause Swedish GDP growth by comparing out-of-sample forecasting performance. For that purpose a Bayesian vector autoregressive model (hereafter abbreviated BVAR) is constructed. The model is mean-adjusted, as proposed by Villani (2009), allowing prior information about the unconditional means (steady state) of the included variables.

Granger causality is not seldom assessed by some statistical test upon parameters in a time-series econometric model or model fit comparison through a simple F-test or Wald test. Though, the original definition of causality as proposed by Granger (1969) was to provide additional predictive information. The out-of-sample approach is as such more incline with the original idea, as stated and clarified by Ashley et al. (1980). They also argue that it is in most cases more appealing for its practical implications.

The model in this study is specified by four variables, Swedish and Euro area real GDP growth, the Swedish Economic Tendency Indicator and the corresponding Euro area Economic Sentiment Indicator. The hypothesis is that by including the Euro area indicator the forecasting performance is enhanced, thus arguing for Granger causality. That information could be of interest for forecasters, policymakers and also provide a basis for further studies of international confidence indicators.

The Bayesian model allows for (and demands) the inclusion of prior distributions of the estimated parameters, leading to an altogether different set of considerations than the frequentist approach, but also altogether different possibilities Such possibilities are avoiding overfitting, prior imposed block exogenity of the Euroarea and Swedish economic variables, including prior information of steady state etc. which will be addressed in Section 3.

The paper is outlined by six sections, including the Introduction. The second section provides a description of the Economic Sentiment Indicator, followed by Section 3, describing the BVAR-model, choice of priors and the forecasting procedure. Section 4 presents the more specific study design followed by results in Section 5. Lastly, a discussion about the results and its conclusions is made in Section 6.

# 2 The Economic Sentiment Indicator

The Economic Sentiment Indicator (ESI) is a composite indicator of business and consumer surveys in the member states of the European Union. The European Commision has formed the "Joint Harmonised EU Programme of Business and Consumer Surveys", which is managed by the Directorate-General for Economic and Financial Affairs (DG ECFIN) (see Directorate-General for Economic and Financial Affairs, 2014). The purpose of this programme is to harmonise surveys in the European Union, allowing for comparisons between member states as well as forming aggregate measurements for the EU and the Euro area. National institutes are in charge of conducting the surveys, with the approval of the European Commision, which partly covers the survey costs. Two basic principles for the national institutes are stated; "use the same harmonised questionnaires" and "conduct national surveys, and transmission of the results, according to a common timetable" (Directorate-General for Economic and Financial Affairs, 2014, p. 5). Harmonisation does not mean exact uniformity in this case.

Monthly surveys in five economic sectors is used to assemble the ESI: industry, services, construction, retail trade and from consumers. The survey-sectors are individually summarised to confidence indicators by individual components, balancing questions about order book and economic expectations, which then in turn are summarized by the ESI.

The ESI is formed through 15 standardized individual components from the confidence indicators with explicit weights allocated to the five sectors, leading to the composite design described in Table 2.1.

| Survey Sector | Weight Share |
|---|---|
| Industry | 40 % |
| Services | 30 % |
| Consumers | 20 % |
| Construction | 5 % |
| Retail Trade | 5 % |

Table 2.1: Weight Share of the Survey Sectors

In this study, the Economic Sentiment Indicator in Euro area is assessed. The Euro area is regarded as all the EU member states with the single currency, today consisting of a total of 19 countries.

The replies of the questionnaires are aggregated on the euro-area level through a weighted average of the national replies. These weights are chosen as the euro-area member share of a reference series, e.g. gross value added at constant prices for Construction or private final consumption expenditure for Consumers.They are smoothed by a 2-year moving average and are updated once a year. Details of the reference series, aggregation procedure and the harmonised questionnaires are stated in the User guide by the DG ECFIN (see Directorate-General for Economic and Financial Affairs, 2014).

The Economic Tendency Indicator (ETI) in Sweden corresponds to the ESI, with some smaller differences regarding the seasonal adjustment procedure and the consumer indicator definition. (see National Institute of Economic Research).

# 3 Bayesian Vector Autoregression

The mean-adjusted Bayesian Vector Autoregressive model proposed by Villani (2009) can be described in the general form by Equation 3.1

$$\mathbf{\Pi}(L)(\mathbf{x}_t - \boldsymbol{\mu}) = \boldsymbol{\eta}_t \tag{3.1}$$

where $\mathbf{\Pi}(L) = \mathbf{I} - \mathbf{\Pi}_1 L - \cdots - \mathbf{\Pi}_m L^m$ is a lag polynomial of order $m$, $\mathbf{x}_t$ is a $n \times 1$ vector of stationary variables, $\boldsymbol{\mu}$ is $n \times 1$ vector of the unconditional means of the included variables and a $n \times 1$ vector $\boldsymbol{\eta}_t \sim N(\mathbf{0}, \mathbf{\Sigma})$ with *iid* elements.

The model form has been used earlier in studies for similar purposes, when determining Granger causality of a variable on real economic variables. An example is Stockhammar and Österholm (2015) who are studying the effects of an US economic policy uncertainty index upon Swedish GDP growth. This paper has a closely related methodology to their study.

The standard VAR-model was proposed by Sims (1980), arguing for its great flexibility, but with the cost of an apparent risk of over-parametrisation. Sims suggested that imposing prior information of parameters could deal with the issue. This followed by a progress of developing the Bayesian approach of imposing prior distributions for VAR-models (e.g. Doan et al., 1984; Litterman, 1986).

The mean-adjustment is a departure from the conventional BVAR-model.Villani (2009) proposed this adjustment as a way of model the unconditional mean (or as he calls it the "steady state") of the variables explicitly and as such include prior information of it. The parameters in the conventional model seldom have intuitive choices of prior belief, (e.g. imposing a prior to the coefficient for the third lag of variable 1 upon variable 2). On the other hand, economists often claim to have quite strong prior opinions of the steady state of macroeconomic variables such as GDP growth. Villani simply claims that it is easier to have good prior beliefs of the steady state than of the dynamics. The addition of such beliefs could arguably improve forecasting performance and take into account long-term economic trends as the forecasts tend to converge to these states. In his study, the forecasting of inflation, GDP growth and three-months interest rates in Sweden is improved by this adjustment. Its usage has grown to be more common in economic forecasting (e.g Clark, 2009; Berger and Österholm, 2011) and in academic teaching literature (e.g Clements and Hendry, 2011).

The BVAR will be identified recursively using standard Cholesky decomposition of the covariance matrix (see Sims, 1980, for details)

The estimation and forecasting procedures in this study are are all made in MATLAB®[1], with codes produced by and obtained from Mattias Villani (for handbook, see Villani, 2007).

## 3.1 Prior and Posterior distributions

In a Bayesian approach, the model parameters are treated as random variables. All the model parameters to estimate in the BVAR-model, $\boldsymbol{\theta} = (\boldsymbol{\Pi}, \boldsymbol{\Sigma}, \boldsymbol{\mu})$, are given a prior distribution, $\pi(\boldsymbol{\theta})$, that is not conditioned upon any realised observation. It is used to form the posterior distribution $p(\boldsymbol{\theta}|\mathbf{X}_T)$, that is the distribution of the parameters conditional on the observed data.

Another key element in BVAR-estimaton is the distribution of the observed data conditional on the parameters, that is the likelihood function in Equation 3.2

$$L(\mathbf{X}_T|\boldsymbol{\theta}) = \prod_{t=1}^{T} f(\mathbf{x}_t|\mathbf{X}_{t-1}, \boldsymbol{\theta}) \tag{3.2}$$

By Bayes rule the posterior distribution is proportional to the product of the likelihood function and the prior distribution as shown in Equation 3.3.

$$p(\boldsymbol{\theta}|\mathbf{X}_T) = \frac{L(\mathbf{X}_T|\boldsymbol{\theta})\,\pi(\boldsymbol{\theta})}{\int L(\mathbf{X}_T|\boldsymbol{\theta})\,\pi(\boldsymbol{\theta})\,d\boldsymbol{\theta}} \propto L(\mathbf{X}_T|\boldsymbol{\theta})\,\pi(\boldsymbol{\theta}) \tag{3.3}$$

where $\int L(\mathbf{X}_T|\boldsymbol{\theta})\,\pi(\boldsymbol{\theta})\,d\boldsymbol{\theta} = p(\mathbf{X}_T)$ is the marginal likelihood.

When integrating out the marginal distributions of the parameters, the point estimates can be found, as well as measurements of their precision.

It is apparent that the posterior distribution depends on the data as well as the prior beliefs. The choosing of prior distributions are as such crucial to the Bayesian estimation.

[1]MATLAB® version 8.3.0.532 (R2014a), The MathWorks Inc. Natick, Massachusetts

The frequentist approach is the special case of the Bayesian VAR when the prior distributions are totally uninformative. The point estimates of the marginal posterior distributions converge to the Ordinary Least Squares estimate if the so called diffuse prior, $\pi(\boldsymbol{\theta}) \propto |\boldsymbol{\Sigma}|^{\frac{-(n+1)}{2}}$, is being used. One common way of choosing priors is to have prior information on the coefficients using the normal distribution while using diffuse prior on the error covariance matrix. This is called normal-diffuse priors (see Kadiyala and Karlsson, 1997). Villani (2009) shows that in the Steady State VAR the normal-diffuse priors become:

$$p(\boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{\frac{-(n+1)}{2}} \tag{3.4}$$

$$\boldsymbol{\mu} \sim N_n(\boldsymbol{\theta}_{\boldsymbol{\mu}}, \boldsymbol{\Omega}_{\boldsymbol{\mu}}) \tag{3.5}$$

$$vec(\boldsymbol{\Pi}) \sim N_{mn^2}(\boldsymbol{\theta}_{\boldsymbol{\Pi}}, \boldsymbol{\Omega}_{\boldsymbol{\Pi}}) \tag{3.6}$$

where $vec(\boldsymbol{\Pi}) = (\boldsymbol{\Pi}_1, \boldsymbol{\Pi}_2, \dots, \boldsymbol{\Pi}_m)'$.

The prior distribution of the steady state $N_n(\boldsymbol{\theta}_{\boldsymbol{\mu}}, \boldsymbol{\Omega}_{\boldsymbol{\mu}})$ is specified in Section 4, while the distribution of the lag coefficients $N_{mn^2}(\boldsymbol{\theta}_{\boldsymbol{\Pi}}, \boldsymbol{\Omega}_{\boldsymbol{\Pi}})$ is specified by the Minnesota Prior beliefs (see Litterman, 1986).

### 3.1.1 Minnesota Prior beliefs

Doan et al. (1984) and Litterman (1986) proposed the now called Minnesota Priors, that are well suited to avoid over-parametrisation in a VAR-model as they restrict the lag-structure by imposing tightness on the parameters. The first mean-lag ($diag(\boldsymbol{\theta}_{\boldsymbol{\Pi}_1})$) is chosen to be 0.9 for variables modelled in levels, and 0 for variables modelled in first difference. This takes into account the presumed trending behaviour of the macroeconomic variables. The rest of the lag-coefficients in $\boldsymbol{\theta}_{\boldsymbol{\Pi}}$ are presumed as 0.

To specify the variance-covariance matrix of the coefficients, $\boldsymbol{\Omega}_{\boldsymbol{\Pi}}$, shrinkage hyperparameters are used, which form the tightness structure. Equation 3.7 shows the specification:

$$
diag\left(\mathbf{\Omega_{\Pi}}\right) = \begin{cases} \left(\dfrac{\lambda_1}{m^{\lambda_3}}\right)^2 & \text{for own lags} \\ \\ \left(\dfrac{\lambda_1\lambda_2}{m^{\lambda_3}}\right)^2 \dfrac{\sigma_q^2}{\sigma_r^2} & \text{for lags of variable q in equation r} \end{cases} \tag{3.7}
$$

Where the number of lags is denoted by $m$. The hyperparameters $(\lambda_1, \lambda_2, \lambda_3)$ impose tightness according to Table 3.1. $\sigma_q^2$ is the variance of residuals from an univariate AR(m) parameter estimations for variable $q$. As such $\frac{\sigma_q^2}{\sigma_r^2}$ controls for differences in scales and units of measurement for the variables.

| Hyperparameter | Description | Allowed Range | Value in this study |
|---|---|---|---|
| $\lambda_1$ | Overall Shrinkage | $\lambda_1 > 0$ | 0.2 |
| $\lambda_2$ | Cross-variable Shrinkage | $0 < \lambda_2 \leq 1$ | 0.5 |
| $\lambda_3$ | Lag-decay | $\lambda_3 > 0$ | 1 |

Table 3.1: Minnesota hyperparameter description and choice

The hyperparameter values in Table 3.1 are the same as the original Minnesota priors (see Litterman, 1986), following standard practice. They assume to be tighter around zero for cross-variable lags. Also, the priors are increasingly tighter around zero for each lag, following the prior notion of diminishing effects.

### 3.1.2 Block Exogeneity

In this study, variables in a large economy (Euro-area) are estimated in an endogenous system with variables in a small economy (Sweden). The prior notion would be that the Swedish variables have negligible effects on the Euro-area. By the addition of a near-zero hyperparameter $\lambda_4$, for a special case of the cross-variable parameters in Equation 3.7, can these prior beliefs be included. This is called imposing block exogeneity.

$$
diag\left(\mathbf{\Omega_{\Pi}}\right) = \left(\frac{\lambda_1\lambda_2\lambda_4}{m^{\lambda_3}}\right)^2 \frac{\sigma_q^2}{\sigma_r^2} \tag{3.8}
$$

for lags of small economy variable q in large economy equation r

This study assigns $\lambda_4 = 0.0001$ in Equation 3.8, thus imposing a hard tightness around zero for small economy variables upon large economy variables.

### 3.1.3 Gibbs Sampling

To form a predictive distribution used in forecasting, it is generally necessary to generate $\boldsymbol{\theta}$ through simulations. Often, it is quite difficult to simulate $\boldsymbol{\theta}$ from the joint posterior distribution $p(\boldsymbol{\theta}|\mathbf{X}_T)$ directly. Kadiyala and Karlsson (1997) show that the Gibbs sampling procedure could be viable in this case. If for an arbitrarily $\theta = (\theta_1, \theta_2, \theta_3)$ there are the full conditional distributions, from which it is know how to generate pseudo-random numbers;

$$p(\theta_1|\theta_2, \theta_3) \quad p(\theta_2|\theta_1, \theta_3) \quad p(\theta_3|\theta_1, \theta_2)$$

Then the Gibbs sampler generates the steps in a chain with:

$$\begin{array}{lll} \theta_1^{(i+1)} & \text{from} & p\left(\theta_1 \middle| \theta_2^{(i)}, \theta_3^{(i)}\right) \\ \theta_2^{(i+1)} & \text{from} & p\left(\theta_2 \middle| \theta_3^{(i)}, \theta_3^{(i)}\right) \\ \theta_3^{(i+1)} & \text{from} & p\left(\theta_3 \middle| \theta_1^{(i)}, \theta_2^{(i)}\right) \end{array}$$

This is done iteratively for a number of draws. Villani (2009) proves that for the normal-diffuse priors in a Steady State BVAR-model, the full conditional posteriors for $\boldsymbol{\Sigma}$ is inverted-Wishart and multivariate-normal for $\boldsymbol{\mu}$ and $\boldsymbol{\Pi}$. In this study, simulating $\boldsymbol{\theta}$ from the posterior is done by this procedure, using 10000 draws.

## 3.2 Predictive distribution

Karlsson (2012) describes the process of Bayesian forecasting. The (posterior) predictive distribution $p(\mathbf{x}_{T+1:T+H}|\mathbf{X}_T)$ is central in this process, which denotes the distribution on future data points $\mathbf{x}_{T+1:T+H}$ conditional on the observed data $\mathbf{X}_T$. It contains all information about the future events. The forecaster is the one to decide which aspects of the distribution that is of interest for the actual forecast.

To find the predictive distribution, one key element is the distribution of future events conditional on the data and the parameters $p\left(\mathbf{x}_{T+1:T+H}|\mathbf{X}_T,\boldsymbol{\theta}\right)$. This distribution corresponds to the likelihood function in Equation 3.2, which is represented in Equation 3.9.

$$p\left(\mathbf{x}_{T+1:T+H}|\mathbf{X}_T,\boldsymbol{\theta}\right)=\prod_{t=1}^{T+H} f\left(\mathbf{x}_t|\mathbf{X}_{t-1},\boldsymbol{\theta}\right) \tag{3.9}$$

When extending the likelihood function to include the future observations $\mathbf{x}_{T+1:T+H}$ one gets the distribution above. By Bayes rule the predictive distribution can be specified as

$$p\left(\mathbf{x}_{T+1:T+H}|\mathbf{X}_T\right)=\frac{p\left(\mathbf{x}_{T+1:T+H},\mathbf{X}_T\right)}{p\left(\mathbf{X}_T\right)}=\frac{\int p\left(\mathbf{x}_{T+1:T+H}|\mathbf{X}_T,\boldsymbol{\theta}\right)L\left(\mathbf{X}_T|\boldsymbol{\theta}\right)\pi\left(\boldsymbol{\theta}\right)d\boldsymbol{\theta}}{\int L\left(\mathbf{X}_T|\boldsymbol{\theta}\right)\pi\left(\boldsymbol{\theta}\right)d\boldsymbol{\theta}} \tag{3.10}$$

By substituting in the posterior distribution $p\left(\boldsymbol{\theta}|\mathbf{X}_T\right)$ from Equation 3.3 one gets

$$p\left(\mathbf{x}_{T+1:T+H}|\mathbf{X}_T\right)=\int p\left(\mathbf{x}_{T+1:T+H}|\mathbf{X}_T,\boldsymbol{\theta}\right)p\left(\boldsymbol{\theta}|\mathbf{X}_T\right)d\boldsymbol{\theta} \tag{3.11}$$

As shown, the predictive distribution in Bayesian forecasting contains the uncertainty of the future observations described by $p\left(\mathbf{x}_{T+1:T+H}|\mathbf{X}_T,\boldsymbol{\theta}\right)$ and the uncertainty of the parameters described by the posterior distribution $p\left(\boldsymbol{\theta}|\mathbf{X}_T\right)$.

Through a simulation process the predictive distribution can be obtained. Villani (2007) uses the following steps:

1. Simulate a $\boldsymbol{\theta}$ from posterior distribution $p\left(\boldsymbol{\theta}|\mathbf{X}_T\right)$ by Gibbs sampling.

2. Simulate the VAR-process h-steps forward conditional on the generated $\boldsymbol{\theta}$ obtaining a sequence draw $\mathbf{x}_{T+1:T+H}$ from $p\left(\mathbf{x}_{T+1:T+H}|\mathbf{X}_T,\boldsymbol{\theta}\right)$.

3. Repeat the preceding steps until convergence to the predictive distribution is obtained.

The point forecast is then approximated through the median of the simulated distribution, as a summarizing measure. The predictive densities obtained can be used to study different aspects of the forecast, not least in its evaluation.

### 3.2.1 Forecast Evaluation

As stated earlier, this study is making forecasts for real Swedish GDP growth. The forecast horizon is 1 to 8 steps ahead ($H = 8$). As the study handles quarterly data, that means forecasting up to 2 years from the starting point. The evaluation of the forecast accuracy is done by a dynamic forecasting scheme.

1. Create a sequential forecast eight steps forward starting 2004Q1, based on the preceding data.

2. Include the observation in 2004Q1 in the estimation, forecasting eight steps forward starting 2004Q2.

3. Repeat until last sequential forecast starting 2013Q1.

The aim is to compare forecasts done with the model including the ESI and with models excluding the ESI. Two different measurements are used to evaluate the forecast performances. The first is to measure how the point forecasts differ from the actual outcomes, which is summarized by the Root Mean Square Error (RMSE). In this scheme, each model will obtain eight RMSEs, one for each step in the forecasting horizon. A smaller RMSE indicates a better forecasting performance.

The other measurement is called the log predictive density score (LPDS). By evaluating the (logarithm of) predictive density function at the actual outcome, one gets a measurement of the predictive density sharpness (see Mitchell and Wallis, 2011)

$$LPDS_{t,h} = log\, p_t\left(x_{t+h}\right) \tag{3.12}$$

While studying point forecasts is an intuitive approach, they are only pure summaries of the actual forecast density and could be seen as quite crude measurements (or elementary based on the point of view). The log predictive density score however could been seen as a measurement of forecast density accuracy. By taking the average of the LPDS:s, one obtains a simple summary of how close the forecast densities fit the actual outcomes. A larger average LDPS indicates a better forecasting performance.

To make relative comparisons between forecasts, some statistical tests are commonly made. The differences between the mean square errors (MSE) of forecasts as well as differences between the LDPS could serve as loss-functions within the framework of Diebold-Mariano tests for comparing predictive accuracy (see Diebold and Mariano, 1995). However, Clark and McCracken (2001) show that the assumption of standard distributions of the tests may be violated for recursive one-step ahead forecasts by nested models (which is the case in this study). They later extend the results to the multistep forecast case (see Clark and McCracken, 2005). For the case of Mean Square Error-based Diebold-Mariano tests, Clark and West (2007) has constructed an adjusted test to deal with this issue. Both the regular test and the adjusted test will be presented. In the case of Log Predictive Score-based tests, no such evaluation or argument of a violated assumption has been made. Also, the Clark and West test is not suitable in this case, as their adjustment is based on pure prediction errors and not density sharpness measures. In this case, only the Diebold-Mariano test will be conducted in absence of any obvious alternative. It should also be noted that, in practice, one chooses the model which has best forecasting performance in the measurements (commonly the RMSE) regardless of the occurrence of significant differences based on statistical tests.

# 4 Study Design

The study handles quarterly data from 1996 to 2014. Four macroeconomic variables are included in accordance to Equation 4.1 and Table 4.1.

$$\mathbf{x}_t = \begin{pmatrix} y_t^{EA} & \Delta ESI_t^{EA} & y_t & \Delta ETI_t \end{pmatrix}' \tag{4.1}$$

| | |
|---|---|
| $y^{EA}$ | Euro area GDP growth |
| $\Delta ESI^{EA}$ | Euro area Economic Sentiment Indicator (first difference) |
| $y$ | Swedish GDP growth |
| $\Delta ETI$ | Swedish Economic Tendency Indicator (first difference) |

Table 4.1: Variable Notations

The GDP growth variables are seasonally-adjusted real GDP quarter-to-quarter growth in percent. The indicator variables are the average of the monthly observations each quarter, measured in first difference (e.g $ESI_t^{EA} - ESI_{t-1}^{EA}$). All data has been obtained through publicly available and free data sources; Swedish GDP growth from Statistics Sweden [1], the ETI from The National Institute of Economic Research[2] and the Euro area variables from Eurostat [3].

A BVAR-model is formed, where the normal-diffuse priors are chosen in accordance to Section 3.1, with Minnesota priors on the lag-coefficients specified in Section 3.1.1. The prior distribution of the steady states are independent normal, with intervals specified in Table 4.2.

[1]http://www.scb.se/ [Online Accessed:2015-04-01]
[2]http://www.konj.se/ [Online Accessed: 2015-04-01]
[3]http://ec.europa.eu/eurostat/web/main/home [Online Accessed: 2015-04-01]

| | $y_t^{EA}$ | $\Delta ESI_t^{EA}$ | $y_t$ | $\Delta ETI_t$ |
|---|---|---|---|---|
| 95 % prior probability interval | $(0.5, 0.75)$ | $(-0.5, 0.5)$ | $(0.5, 0.75)$ | $(-0.5, 0.5)$ |

Table 4.2: Steady State Prior Intervals

The median of the steady states for the seasonally-adjusted real GDP growth is presumed to be 0.625 % (median of steady state interval). These prior intervals reflects the choices of Stockhammar and Österholm (2014) when they use a mean-adjusted BVAR-model to study the spillover of Euro area shocks upon Swedish GDP Growth.

As the indicators ESI and ETI are indices standardized to mean-value 100 and standard deviations 10, its plausible to assume that the steady states are tightly centered around zero when measuring in first-difference.

Forecasting procedures will be done as proposed in Section 3.2. The RMSE and the LPDS serve as measurements for the forecasting performances. As the purpose is to assess if the inclusion of the ESI improve upon the forecasts, thus arguing for Granger causality, the full model in Equation 4.1 will be compared with the univariate model in Equation 4.2 and the trivariate model in Equation 4.3.

$$\mathbf{x}_t = y_t' \tag{4.2}$$

$$\mathbf{x}_t = \begin{pmatrix} y_t^{EA} & y_t & \Delta ETI_t \end{pmatrix}' \tag{4.3}$$

National confidence indicators as the ETI and real economic variables as the Euro area growth are intuitive choices and common in forecasting Swedish GDP. It is not implausible to assume that the Euro area ESI is closely related to the Swedish ETI and the Euro area GDP growth, and as such it may not contain any additional predictive information of Swedish GDP. A comparison of the forecasting performances between the trivariate model (excluding ESI) and the full model (including ESI) will function as a tool to evaluate the indicator from a predictive standpoint. Also, if the ESI does contain predictive information one would assume that its inclusion would lead to a better forecasting performance than in the univariate case.

# 5 Results

In this section the results will be presented. First, the full model, including all four variables, is estimated. The individual parameters could then be assessed and analysed. Second, the forecasting performances of the univariate, the trivariate and the full model are presented and evaluated.

## 5.1 Model Estimation

As stated earlier, among the parameters to estimate are the unconditional means, $\boldsymbol{\mu}$, for the variables (i.e. the steady states). Figure 5.1 shows the posterior steady states within their timeplots.

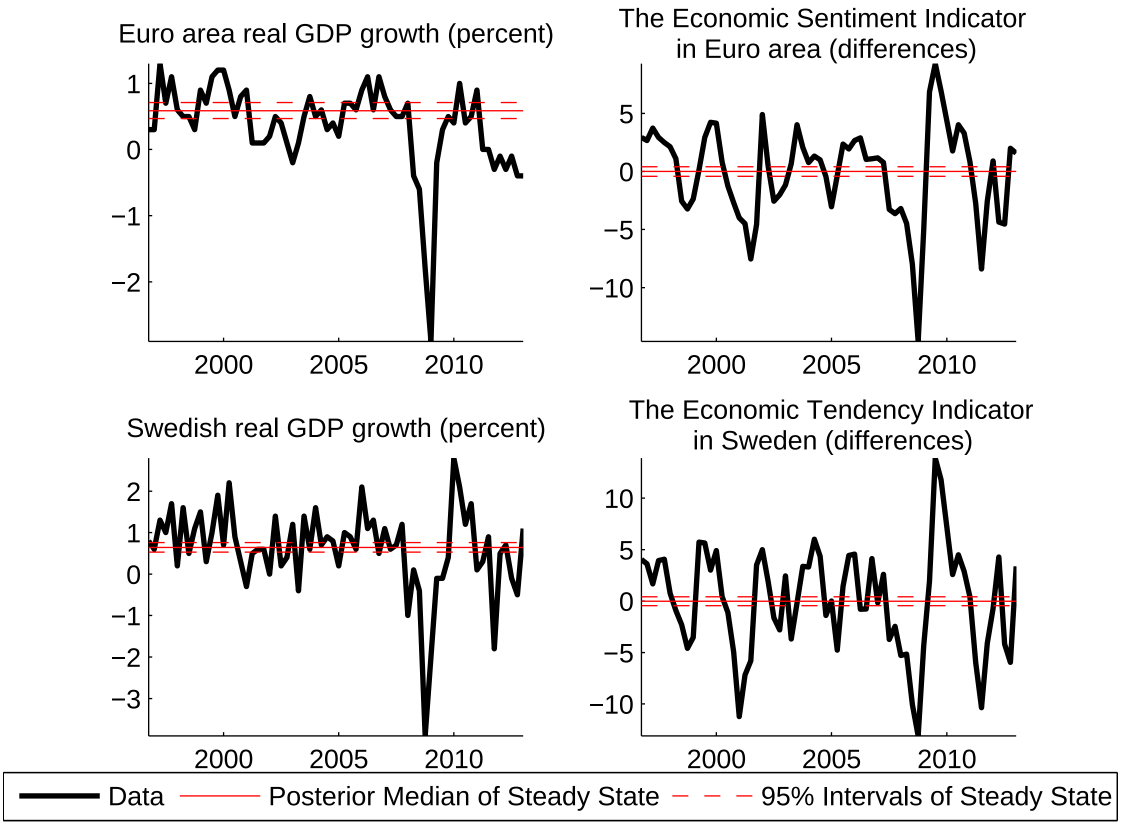
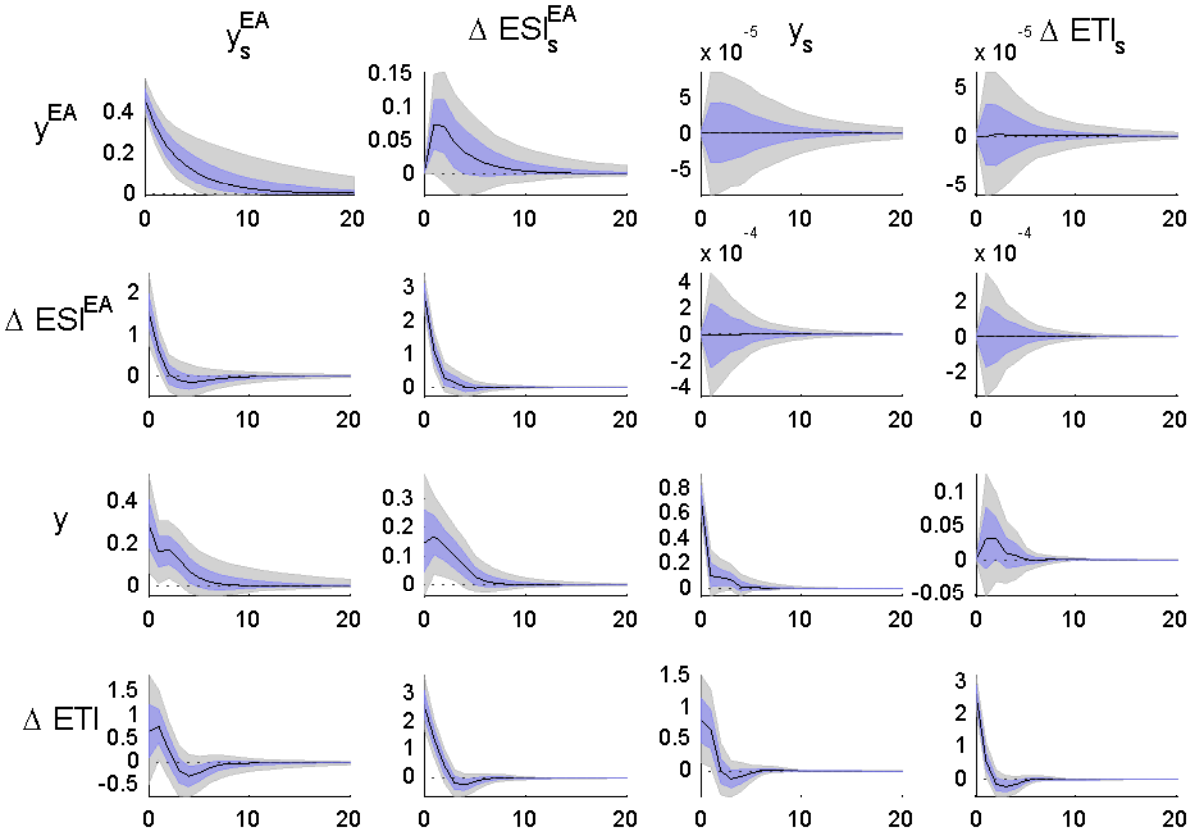



Figure 5.1: Timeplots with posterior Steady State by each variable

When comparing the posterior steady states in the figure with the chosen prior steady states in Table 4.2, it is apparent that they do not differ to any large extent.

The lag-coefficient parameters, $\mathbf{\Pi}$, are estimated and presented in full in the Appendix. For the purpose of this study, analysing the equation with the Swedish GDP growth as the endogenous variable is of most relevance, in particular the coefficients of the ESI-lags. They are all of the expected positive sign, indicating that positive changes in the ESI leads to higher Swedish GDP growth. The first lag has a significant effect on the 10 % level, with parameter-post mean of 0.0433 and standard deviation 0.0255. It is worth noting that the lag-coefficients of the other variables do not have individual significant effects.

Perhaps a more plausible way of studying the effects, not solely relying on the lag-coefficients individually, is through the impulse-response functions. They trace out how a unit increase in one variable effects another over time (all else equal). Figure 5.2 present these functions graphically.




Note: Confidence bands denoted by the purple area donotes (68 % interval) and the grey area (95 % interval)

Figure 5.2: Impulse Response Functions in the full model, with variable shocks in columns upon variable in row.

As seen in the figure, the Swedish GDP growth (third row) increases by positive shocks of the other variables. A unit increase of the first difference of the ESI leads to a simultaneous effect, which increases after a few quarters and then diminishes over time. Shocks from the Euro area ESI has a significant effect on Swedish GDP growth. It can be compared to the corresponding Swedish ETI, which does not have a significant effect in this identification scheme.

The four functions in the upper-right corner represents the shocks of the Swedish variables upon the Euro area variables. The absence of such shocks is a good representation of the results of imposing Block Exogeneity described in Section 3.1.5.

## 5.2 Out-of-Sample Forecast

A visually representation of the dynamic sequential forecasts can be done through a cascade plot. Figure 5.3 shows the cascade plot of Swedish GDP growth and its sequential point predictions by the full model.

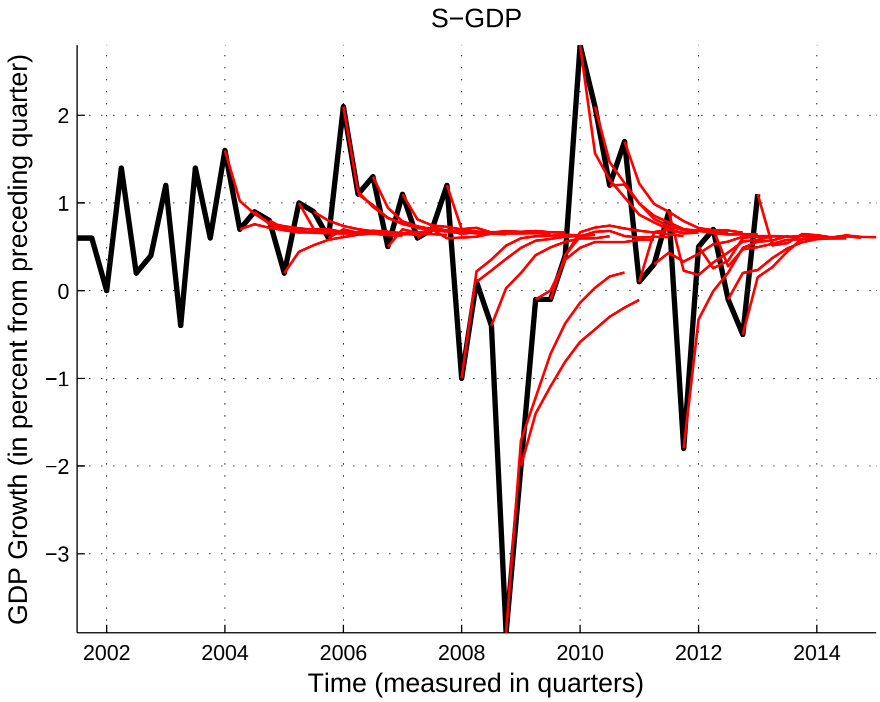



Figure 5.3: Cascade plot of Swedish GDP growth (in percent), plotting each sequential forecast (8-quarters ahead) by the full model

When studying the figure, it is apparent that the forecasts converge to the steady state, which is to be assumed. Also, it is apparent that the forecasts are suffering when it comes to predict the 2008-financial crisis, resulting in large spikes in prediction errors during this period.

Forecasts of three models are compared. Figure 5.4 presents the RMSE for each model and each forecasting horizon, thus comparing the point prediction accuracy.

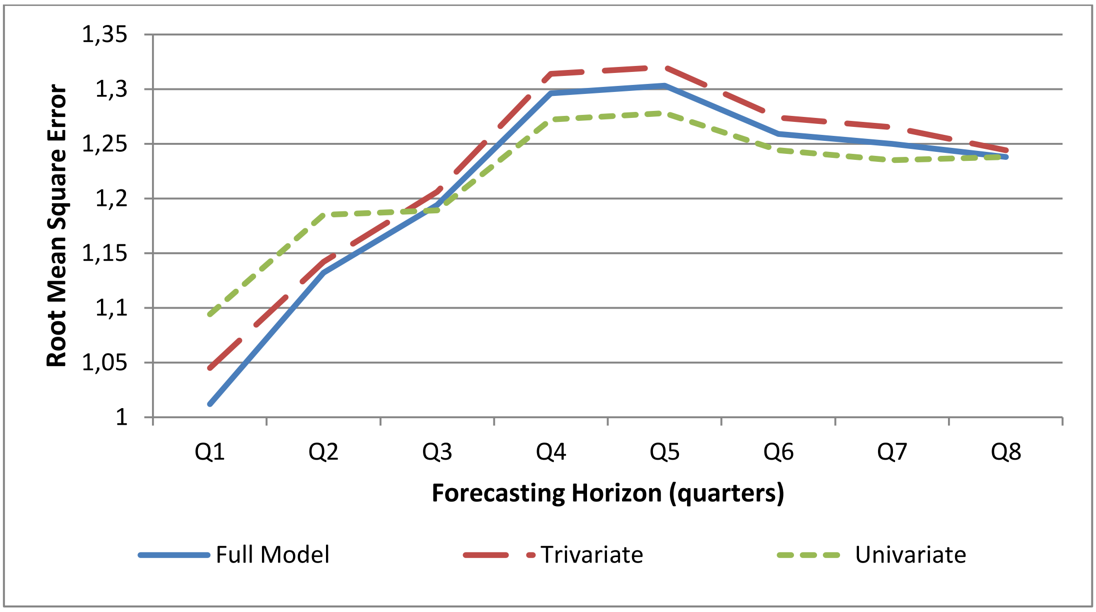



Figure 5.4: Root Mean Square Errors for different forecasting horizon in quarters, plotted for each model

The full model (including ESI) has a lower RMSE for all forecasting horizon than the trivariate model, indicating a better forecast performance short-term and long-term. When forecasting up to two quarters ahead, the differences between the full model and the univariate model is even more apparent. However, in the long-term, the univariate model has the lowest RMSE. The differences between all models are practically eradicated when forecasting 2-years ahead. Also, all models perform best in the short-term. However, the worst performances are made when forecasting four or five quarters ahead, and are improved in even longer horizons. As the forecasts converge to the steady states, such improvements do not have to be remarkable, though they should be noted.

The Diebold-Mariano test can indicate the statistical relevance of the prediction comparisons. Table 5.1 shows the results of the DM-tests for comparing predictive accuracy based on the mean-square errors, by each forecast horizon.

| Full model versus Trivariate model | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| DM-statistic | -2.438** | -1.049 | -0.821 | -0.810 | -0.929 | -1.234 | -1.268 | -1.331 |
| CW-statistic | -2.569** | -1.673 | -1.043 | -1.194 | -1.777* | -1.249 | -1.174 | -1.462 |
| Forecasting horizon | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q8 |
| **Full model versus Univariate model** | | | | | | | | |
| DM-statistic | -2.204** | -1.771* | 0.329 | 1.036 | 1.272 | 1.194 | 1.043 | -0.690 |
| CW-statistic | -2.592** | -2.409** | 0.1146 | 0.790 | 1.260 | 1.218 | 0.9458 | -1.123 |
| Forecasting horizon | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q8 |

Note: ** denotes significance on 5 % level, * denotes significance on 10% level

Table 5.1: Diebold-Mariano and Clark-West tests of comparing predictive accuracy, based on mean square errors (two-sided)

Both tests in both comparisons shows significant differences of the point prediction accuracy in the short-term case. When forecasting one quarter ahead, the full model has a significantly better forecasting performance than both the other models for both tests. It also has a a 10% level significant difference from the univariate model when forecating two quarters ahead by the DM-test and 5% level significant difference by the CW-test. Overall, no significant differences has been found in the long-term case. The exception is a 10% level significant difference from the trivariate model when forecasting 5-quarters ahead by the CW-test.

Again, it is important to repeat that the assumptions of a standard distribution in the DM-test may be violated in the case of nested models, as addressed in Section 3.2.1. However, as the adjusted test shows similar results, there are substantial evidence to believe in the conclusions.

The forecast density sharpness is addressed by the log predictive density scores. They take into account more information from the predictive densities in comparion to the RMSEs, which are only based on point forecasts. Figure 5.5 graphically presents the average LPDS for each model and forecast horizon.

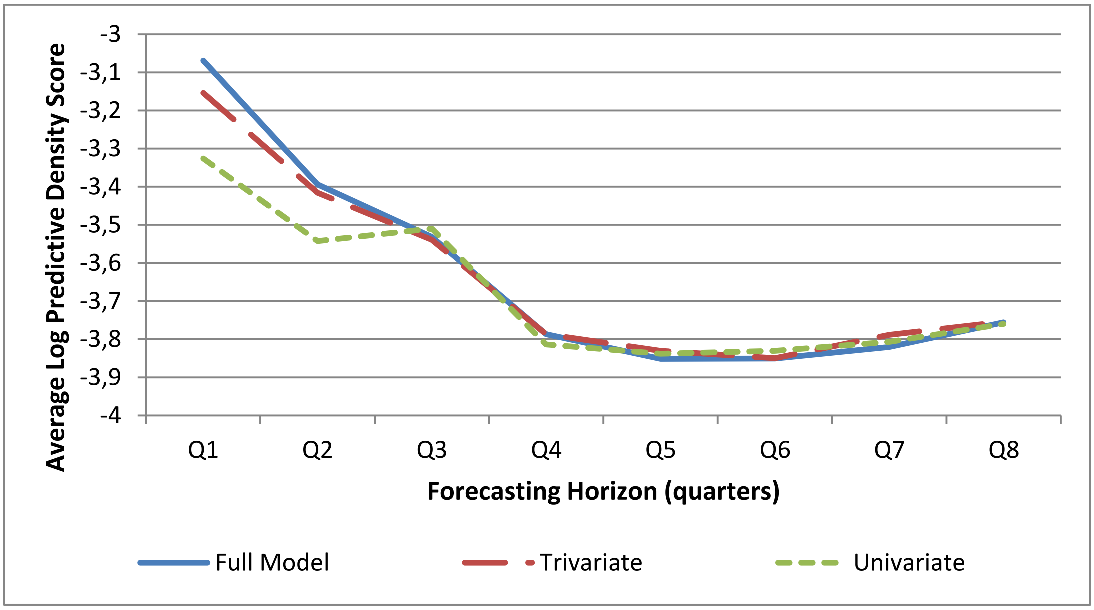



Figure 5.5: Average log predictive density score for different forecasting horizons in quarters, plotted for each model

The figure shows that in short-term, the full model performs the best with the highest average LPDS. In the long-term, any vital differences do not seem apparent. The univariate model has the worst density sharpness in the short-term, which corresponds to the results from the RMSE comparison.

Diebold-Mariano tests are conducted for the log predictive density scores to compare predictive accuracy. Table 5.2 presents the results.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Full model versus Trivariate model | | | | | | | | |
| DM-statistic | 2.315** | 1.104 | 0.186 | -0.007 | -0.562 | -0.004 | -0.674 | -0.068 |
| p-value | 0.026 | 0.279 | 0.853 | 0.995 | 0.577 | 0.997 | 0.505 | 0.947 |
| Forecasting horizon | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q8 |
| Full model versus Univariate model | | | | | | | | |
| DM-statistic | 2.235** | 2.419** | -0.313 | 0.708 | -0.188 | -0.220 | -0.129 | 0.043 |
| p-value | 0.032 | 0.021 | 0.7561 | 0.483 | 0.852 | 0.827 | 0.898 | 0.966 |
| Forecasting horizon | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q8 |

Note: ** denotes significance on 5 % level, * denotes significance on 10% level

Table 5.2: Diebold-Mariano tests of comparing predictive accuracy, based on log predictive density scores (two-sided)

When forecasting one quarter ahead, the full model has a significantly better performance than the trivariate model. It has a significantly better performance in comparison to the univariate model when forecasting one or two quarters ahead. In the long term, no significant differences have been found.

# 6 Discussion

The forecasting evaluations show that in short-term predictions of Swedish GDP growth, the full model including the Economic Sentiment Indicator has the best performance, both in point prediction accuracy and forecast density sharpness. As such there is evidence to suggest that the Euro area Economic Sentiment Indicator contains predictive information about Swedish real GDP growth. This also gets support from the estimated parameters and the impulse-response functions from the full BVAR-model. When forecasting in the long-term however, such evidence is not as prominent. In fact, when forecasting Swedish GDP growth four quarters ahead or more, the RMSE were smaller in the univariate case.

These results suggest that the inclusion of international confidence indicator to forecast national trends in a small open economy could be of value. The harmonisation of business and consumer surveys in the EU could as such not only serve as a basis for aggregate measures or for pure national comparison, but also be used as comparable international leads for national developments.

It is also notable that the RMSEs (and to some degree the LPDS) show increased forecasting performances in the longest horizons (6-8 quarters ahead) compared to the mid-horizons (4-5 quarters ahead). This is likely due to the fact that the forecasts converge to the steady state of Swedish GDP growth. The spikes in prediction errors during the financial crisis in 2008 may well have a large impact on the overall results as both the RMSE and the LPDS are quite sensitive to outliers. It could be likely that these spikes were more apparent in mid-early horizons as the economy long-term get closer to the steady state. Simply put, when jumping 2-years ahead one could miss a large amount of the crisis and easier hit the mark and when forecasting ahead from the crisis period, the long-term forecast still converge to the steady state. Using some other forecasting performance measures, which are less sensitive to outliers, would perhaps overcome this issue, and as such be more suitable when forecasting in a post-crisis world. However, the inability to predict the crisis should not be neglected.

There are still work to be done when producing statistical tests for comparing forecast accuracy, not least in the common approach of studying the addition of individual variables through comparing nested models. The Clark and West test is limited to some cases and not to a broader range of loss functions. Even though it arguably can be unsatisfactory not being able to assess significant differences in all these cases, regardless, in reality the model that performs best should be chosen. Furthering such nested model comparisons through a Bayesian approach may allow other comparative measures (like predictive density comparisons) and as such not be needed to rely on statistical tests.

Conclusively, there seem to be reasons to further study international confidence indicators as national leads for macroeconomic variables. As there are many measures, in several sectors and in many countries, the possibilities are numerous. The harmonisation of these indicators in the EU opens the door to continue similar studies.

Studying now-casts, forecasting the not yet observed GDP of today, with international confidence indicators should also be considered. As leads, they could contain information of not yet observed international macroeconomic trends, which in turn could be of value for national forecasting. In these cases, there are often the issue of lack of observations relative to the parameters. The Bayesian approach can be a valuable tool in dealing with the over-fitting issues of a VAR-model and giving economists a wider range of comparing forecast performances.

# Bibliography


R. Ashley, Granger C. W. J, and R. Schamalensee. Advertising and Aggregate Consumption: An Analysis of Causality. *Econometrica*, 48(5):1149–1167, 1980.

H. Berger and P. Österholm. Does Money Growth Granger Cause Inflation in the Euro Area? Evidence from Out-of-Sample Forecasts Using Bayesian VARs. *The Economic Record*, 87(276):45–60, 2011.

J. Bram and S. Ludvigson. Does Consumer Confidence Forecast Household Expenditure? - A Sentiment Index Horse Race. *Economic Policy Review*, 4(2):59–78, 1998.

T. E. Clark. Real-Time Density Forecasts from VARs with Stochastic Volatility. https://www.ecb.europa.eu/events/conferences/shared/pdf/clark_real_time_density.pdf??59df8e9bc2821d4b12d0eb5bb0fe2bfe, 2009. [Online; accessed 13-May-2015].

T. E. Clark and M. W. McCracken. Tests of Equal Forecast Accuracy and Encompassing for Nested Models. *Journal of Econometrics*, 105:85–110, 2001.

T. E. Clark and M. W. McCracken. Evaluating Direct Multistep Forecasts. *Econometric Reviews*, 24(4):369–404, 2005.

T. E Clark and K. D West. Approximately Normal Tests for Equal Predictive Accuracy in Nested Models. *Journal of Econometrics*, 138(1):291–311, 2007.

M. P. Clements and D. F. Hendry. *The Oxford Handbook of Economic Forecasting*. Oxford University Press, 2011. ISBN 0199875510, 9780199875511.

F. X. Diebold and R. S. Mariano. Comparing Predictive Accuracy. *Journal of Business & Economic Statistics*, 13(3):253–263, 1995.

Directorate-General for Economic and Financial Affairs. The Joint Harmonised EU Programme of Business and Consumer Surveys - User Guide. http://ec.europa.eu/economy_finance/db_indicators/surveys/documents/bcs_user_guide_en.pdf, 2014. [Online; accessed 13-May-2015].

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

R. Golinelli and G. Parigi. Consumer Sentiment and Economic Activity - A Cross Country Comparison. *Journal of Business Cycle Measurement and Analysis*, 1(2):147–170, 2004.

C. W. J. Granger. Investigating Causal Relations by Econometric Models and CrossSpectral Methods. *Econometrica*, 37(3):424–438, 1969.

E. P. Howrey. The Predictive Power of the Index of Consumer Sentiment. *Brookings Papers on Economic Activity*, 1(1):175–207, 2001.

K. R. Kadiyala and S. Karlsson. Numerical Methods for Estimation and Inference in Bayesian VAR-Models. *Journal of Applied Econometrics*, 12(2):99–132, 1997.

S. Karlsson. Forecasting with Bayesian Vector Autoregressions. Örebro University School of Business. 2012. ISSN 1403-0586.

R. B. Litterman. Forecasting with Bayesian vector autoregressions - Five years of experience. *Journal of Business and Economic Statistics*, 5:25–38, 1986.

J. Mitchell and K. Wallis. Evaluating Density Forecasts: Forecast Combinations, Model Mixtures, Calibration and Sharpness. *Journal of Applied Econometrics*, 26(6):1023–1040, 2011.

A. Mourougane and M. Roma. Can Confidence Indicators Be Useful to Predict Short Term Real GDP Growth? *Applied Economics Letters*, 10(8):519–522, 2003.

National Institute of Economic Research. User Guide to the Economic Tendency Survey. http://www.konj.se/download/18.2cabf50a141002857ee147b/User-Guide-Economic-Tendency-Survey.pdf. [Online; accessed 13-May-2015].

C. A. Sims. Macroeconomics and Reality. *Econometrica*, 48:1–48, 1980.

P. Stockhammar and P. Österholm. The Euro Crisis and Swedish GDP Growth A Study of Spillovers. *Applied Economics Letters*, 21(16):1105–1110, 2014.

P. Stockhammar and P. Österholm. Effects of US Policy Uncertainty on Swedish GDP Growth. *Empirical Economics*, pages 1–20, 2015. ISSN 0377-7332. doi: 10.1007/s00181-015-0934-y. URL http://dx.doi.org/10.1007/s00181-015-0934-y.

M. Villani. BAYESVAR - A Matlab Package for Bayesian VARs. 2007.

M. Villani. Steady-State Priors for Vector Autoregressions. *Journal of Applied Econometrics*, 24(4):630 – 650, 2009.

# Appendix

## Bayesian VAR-estimation

| Variable | lag | Post. mean | Post. Std.dev. | Post. t-value |
|---|---|---|---|---|
| $y^{EA}$ | 1 | +0.6516 | +0.0995 | +6.5510 |
| | 2 | +0.0256 | +0.0764 | +0.3351 |
| | 3 | +0.0404 | +0.0568 | +0.7107 |
| | 4 | +0.0120 | +0.0446 | +0.2692 |
| $\Delta ESI^{EA}$ | 1 | +0.0241 | +0.0138 | +1.7527 |
| | 2 | -0.0015 | +0.0083 | -0.1864 |
| | 3 | -0.0017 | +0.0056 | -0.3098 |
| | 4 | -0.0003 | +0.0043 | -0.0577 |
| $y$ | 1 | +0.0000 | +0.0001 | +0.0051 |
| | 2 | -0.0000 | +0.0000 | -0.0133 |
| | 3 | -0.0000 | +0.0000 | -0.0096 |
| | 4 | -0.0000 | +0.0000 | -0.0108 |
| $\Delta ETI$ | 1 | +0.0000 | +0.0001 | +0.0051 |
| | 2 | -0.0000 | +0.0000 | -0.0133 |
| | 3 | -0.0000 | +0.0000 | -0.0096 |
| | 4 | -0.0000 | +0.0000 | -0.0108 |
| Constant | | +0.5883 | +0.0620 | +9.487 |
| Sigma for errors | 0.4892 | | | |

Bayesian Model Estmation - Endogeneous variable $y^{EA}$

| Variable | lag | Post. mean | Post. Std.dev. | Post. t-value |
|---|---|---|---|---|
| $y^{EA}$ | 1 | +0.0630 | +0.4292 | +0.1467 |
| | 2 | -0.2159 | +0.2540 | -0.8500 |
| | 3 | -0.1062 | +0.1755 | -0.6050 |
| | 4 | -0.0159 | +0.1343 | -0.1188 |
| $\Delta ESI^{EA}$ | 1 | +0.3653 | +0.1063 | +3.4364 |
| | 2 | -0.0661 | +0.0773 | -0.8560 |
| | 3 | +0.0351 | +0.0574 | +0.6109 |
| | 4 | -0.0112 | +0.0433 | -0.2594 |
| $y$ | 1 | -0.0000 | +0.0003 | -0.0013 |
| | 2 | +0.0000 | +0.0001 | +0.0270 |
| | 3 | +0.0000 | +0.0001 | +0.0040 |
| | 4 | +0.0000 | +0.0001 | +0.0073 |
| $\Delta ETI$ | 1 | +0.0000 | +0.0001 | +0.0084 |
| | 2 | +0.0000 | +0.0000 | +0.0042 |
| | 3 | -0.0000 | +0.0000 | -0.0137 |
| | 4 | -0.0000 | +0.0000 | -0.0026 |
| Constant | | -0.0009 | +0.2096 | -0.0044 |
| Sigma for errors | 3.4209 | | | |

Bayesian Model Estmation - Endogeneous variable $\Delta ESI^{EA}$

| Variable | lag | Post. mean | Post. Std.dev. | Post. t-value |
|---|---|---|---|---|
| $y^{EA}$ | 1 | +0.0852 | +0.1438 | +0.5928 |
| | 2 | +0.0792 | +0.0852 | +0.9299 |
| | 3 | +0.0010 | +0.0581 | +0.0180 |
| | 4 | -0.0113 | +0.0445 | -0.2531 |
| $\Delta ESI^{EA}$ | 1 | +0.0433 | +0.0255 | +1.6954 |
| | 2 | +0.0065 | +0.0154 | +0.4234 |
| | 3 | +0.0045 | +0.0106 | +0.4240 |
| | 4 | +0.0018 | +0.0080 | +0.2191 |
| $y$ | 1 | +0.1689 | +0.1280 | +1.3198 |
| | 2 | +0.0585 | +0.0789 | +0.7416 |
| | 3 | +0.0590 | +0.0583 | +1.0130 |
| | 4 | -0.0174 | +0.0465 | -0.3751 |
| $\Delta ETI$ | 1 | +0.0126 | +0.0182 | +0.6927 |
| | 2 | +0.0069 | +0.0108 | +0.6344 |
| | 3 | +0.0002 | +0.0073 | +0.0266 |
| | 4 | +0.0016 | +0.0055 | +0.2971 |
| Constant | | +0.6447 | +0.0589 | +10.9529 |
| Sigma for errors | 0.8701 | | | |

Bayesian Model Estmation - Endogeneous variable $y$

| Variable | lag | Post. mean | Post. Std.dev. | Post. t-value |
|---|---|---|---|---|
| $y^{EA}$ | 1 | +0.0674 | +0.5977 | +0.1128 |
| | 2 | -0.1960 | +0.3600 | -0.5445 |
| | 3 | -0.1342 | +0.2575 | -0.5210 |
| | 4 | -0.0990 | +0.1939 | -0.5107 |
| $\Delta ESI^{EA}$ | 1 | +0.2513 | +0.1136 | +2.2114 |
| | 2 | +0.0198 | +0.0673 | +0.2941 |
| | 3 | -0.0380 | +0.0468 | -0.8130 |
| | 4 | -0.0066 | +0.0351 | -0.1876 |
| $y$ | 1 | +0.6000 | +0.3592 | +1.6703 |
| | 2 | -0.1030 | +0.0758 | -1.3591 |
| | 3 | -0.0628 | +0.0562 | -1.1168 |
| | 4 | -0.0287 | +0.1082 | -0.2654 |
| $\Delta ETI$ | 1 | +0.2169 | +0.1014 | +2.1389 |
| | 2 | +0.0069 | +0.0108 | +0.6344 |
| | 3 | +0.0002 | +0.0073 | +0.0266 |
| | 4 | -0.0305 | +0.0441 | -0.6902 |
| Constant | | -0.0138 | +0.2189 | -0.0629 |
| Sigma for errors | 4.0951 | | | |

Bayesian Model Estmation - Endogeneous variable $\Delta ETI$

| Mean of contemporaneous coefficients | | | | |
|---|---|---|---|---|
| | $y^{EA}$ | $\Delta ESI^{EA}$ | $y$ | $\Delta ETI$ |
| $y^{EA}$ | 2.0645 | 0.0000 | 0.0000 | 0.0000 |
| $\Delta ESI^{EA}$ | -1.2167 | 0.3458 | 0.0000 | 0.0000 |
| $y$ | -0.5639 | -0.0688 | 1.2946 | 0.0000 |
| $\Delta ETI$ | 0.7623 | -0.3319 | -0.3270 | 0.3716 |

| St.Dev of contemporaneous coefficients | | | | |
|---|---|---|---|---|
| | $y^{EA}$ | $\Delta ESI^{EA}$ | $y$ | $\Delta ETI$ |
| $y^{EA}$ | 0.2047 | 0.0000 | 0.0000 | 0.0000 |
| $\Delta ESI^{EA}$ | 0.3238 | 0.0355 | 0.0000 | 0.0000 |
| $y$ | 0.3785 | 0.0503 | 0.1383 | 0.0000 |
| $\Delta ETI$ | 0.3926 | 0.0619 | 0.1981 | 0.0376 |

Contemporanous effects