European University Institute

DEPARTMENT OF ECONOMICS

EUI Working Papers

ECO 2010/11
DEPARTMENT OF ECONOMICS

FORECASTING NONLINEAR AGGREGATES AND AGGREGATES WITH TIME-VARYING WEIGHTS

Helmut Luetkepohl

EUROPEAN UNIVERSITY INSTITUTE, FLORENCE
DEPARTMENT OF ECONOMICS

Forecasting Nonlinear Aggregates and Aggregates with Time-Varying Weights

HELMUT LUETKEPOHL

EUI Working Paper **ECO** 2010/11

ISSN 1725-6704

Printed in Italy
European University Institute
Badia Fiesolana
I – 50014 San Domenico di Fiesole (FI)
Italy
www.eui.eu
cadmus.eui.eu

March 21, 2010

Forecasting Nonlinear Aggregates and Aggregates with Time-varying Weights

Helmut Lütkepohl[1]
Department of Economics, European University Institute
Via della Piazzola 43, I-50133 Firenze, ITALY
email: helmut.luetkepohl@eui.eu

Abstract. Despite the fact that many aggregates are nonlinear functions and the aggregation weights of many macroeconomic aggregates are time-varying, much of the literature on forecasting aggregates considers the case of linear aggregates with fixed, time-invariant aggregation weights. In this study a framework for nonlinear contemporaneous aggregation with possibly stochastic or time-varying weights is developed and different predictors for an aggregate are compared theoretically as well as with simulations. Two examples based on European unemployment and inflation series are used to illustrate the virtue of the theoretical setup and the forecasting results.

Key Words: Forecasting, stochastic aggregation, autoregression, moving average, vector autoregressive process

JEL classification: C32

[1]I thank Ralf Brüggemann, Andrea Silvestrini and Giacomo Sbrana for commenting on an earlier draft of this paper.

1 Introduction

Forecasting macroeconomic variables is a standard task of many applied economists. A number of economic research institutes and central banks publish macroeconomic forecasts regularly. Because most macroeconomic variables are contemporaneous aggregates it is not surprising that a huge literature exists on forecasting contemporaneously aggregated variables (see Lütkepohl (2009) for a recent survey). Different predictors have been compared in this literature. Most results refer to linear aggregation with time-invariant aggregation weights.

On the other hand, in practice, time-varying aggregation weights are very common. Suppose, for example, that one wants to construct an unemployment rate series for the European Union (EU) by aggregating the unemployment rates of the individual EU countries. The overall unemployment rate, u_t^{EU}, is a weighted average of the individual unemployment rates, $u_t^{(i)}$,

$$u_t^{EU} = \sum_{i=1}^{N} w_{it} u_t^{(i)}. \tag{1.1}$$

Here N denotes the number of member states and the aggregation weights, signified as w_{it}, are related to the relative size of the workforce of country i. Because the relative country sizes change over time, the weights in (1.1) also have to be time-varying. In fact, the number of EU countries has changed over the years and is meant to change further in the future when new member states enter the EU. Hence, one may use for N the maximum number of states and assign zero weights to those countries which were no members at a particular time t. Even for periods with a constant number of member states the weights are related to the workforce in the different member states which is best viewed as stochastic. Hence, the weights are random variables. In other words, the EU unemployment rate is a contemporaneous aggregate with stochastic aggregation weights.

There are many other series where stochastic aggregation weights are used. For example, there are several proposals for aggregating gross domestic product (GDP) or its growth rates based on stochastic weights. Suppose y_t^{EMU} denotes euro-area GDP and $y_t^{(i)}$ is the corresponding figure for country i. Then Winder (1997) computes the aggregate growth rates for the European Monetary Union (EMU) countries as

$$\Delta \log y_t^{EMU} = \sum_{i=1}^{N} \frac{y_{t-1}^{(i)}/e_{TB}^{(i)}}{y_{t-1}^{EMU}} \Delta \log y_t^{(i)}, \tag{1.2}$$

where $e_t^{(i)}$ denotes the exchange rate of country i in period t and TB signifies a fixed base period. Although no exchange rate is needed for the period since 1999, when the euro was introduced as common currency, the exchange rate is required for the pre-EMU period. Winder considers a fixed exchange rate for the sample period. Alternatively, Beyer, Doornik and Hendry (2001) propose the aggregate

$$\Delta \log y_t^{EMU} = \sum_{i=1}^{N} \frac{y_{t-1}^{(i)}/e_{t-1}^{(i)}}{y_{t-1}^{EMU}} \Delta \log y_t^{(i)}, \tag{1.3}$$

which uses a flexible exchange rate for the pre-EMU period. The aggregation weights

$$\frac{y_{t-1}^{(i)}/e_{TB}^{(i)}}{y_{t-1}^{EMU}} \quad \text{and} \quad \frac{y_{t-1}^{(i)}/e_{t-1}^{(i)}}{y_{t-1}^{EMU}}$$

are obviously stochastic in both cases.

Yet another proposal for aggregating euro-area data is due to Anderson, Dungey, Osborn and Vahid (2007) who suggest to account for the structural adjustments in some of the member states prior to the EMU. They consider aggregation weights depending on the distance of the economic conditions in a particular country from those of core countries such as Germany. More precisely, they specify a distance measure for monthly data for country i,

$$d_{i,t} = \min \left(\frac{|y_t^{(i)} - y_t^{core}|}{|y_{1979M3}^{(i)} - y_{1979M3}^{core}|}, 1 \right),$$

where March 1979 is used as the reference period because it marks the time where the European Monetary System began. They define the weight of country i in period t to be

$$w_{it} = w_{i,F} \times (1 - d_{i,t}),$$

where $w_{i,F}$ is the weight of country i when full integration is achieved. Thus, also in this aggregation scheme the weights are stochastic.

Similar ideas can be used to determine other real as well as monetary euro-area aggregates. Further examples of aggregates based on stochastic weights include price levels and associated inflation rates as well as real macroeconomic variables such as consumption and investment. Despite the wide-spread use of aggregates with stochastic weights, the theoretical literature has primarily focussed on forecasting contemporaneous aggregates with fixed, time-invariant weights.

There are also situations were the aggregate is a nonlinear function of disaggregate components. For example, if the disaggregate components are modelled in logs, an exponential transformation may be necessary to obtain the aggregate from the log components. As an example of this type, Hubrich (2005) compares different predictors for euro-area consumer price inflation based on aggregated and disaggregated variables in logs.

In this study a general framework is presented for comparing alternative forecasts for nonlinear aggregates including aggregates with time-varying, possibly stochastic weights. The basic idea is adapted from Hendry and Hubrich (2010) who propose to incorporate disaggregate information in forecasting an aggregate by using a system of variables comprised of the aggregate and some disaggregate series. A Monte Carlo study is used to investigate the small sample properties of the forecasts considered in the theoretical framework and empirical forecast comparisons for European unemployment and inflation variables are used to illustrate the theoretical results.

The study is organized as follows. In the next section the basic setup is presented. Theoretical results on alternative predictors for nonlinear aggregates with possibly stochastic weights are discussed in Section 3. A small sample experiment for investigating practical aspects of using disaggregate information for forecasting an aggregate with stochastic weights is reported in Section 4 and empirical forecast comparisons are discussed in Section 5. Conclusions are drawn in Section 6.

The following general notation is used. The lag operator is denoted by L, that is, for a stochastic process or time series y_t, $Ly_t = y_{t-1}$. The normal distribution with mean (vector) μ and variance (covariance matrix) Σ is denoted by $\mathcal{N}(\mu, \Sigma)$. The sets of integers and positive integers are signified as $\mathbb{Z}$ and $\mathbb{N}$, respectively. The $(K \times K)$ identity matrix is denoted by I_K. Moreover, the following abbreviations are used: AR for autoregressive or autoregressive process, VAR for vector autoregressive, MA for moving average, ARCH for autoregressive conditional heteroskedasticity, iid for independently identically distributed, OLS for ordinary least squares, GLS for generalized least squares, DGP for data generating process, MSE for mean squared error, RMSE for root mean squared error, HICP for harmonized index of consumer prices, AIC for the Akaike information criterion and SC for the Schwarz-Rissanen information criterion (see, e.g., Lütkepohl (2005, Section 4.3) for precise definitions and discussion of these criteria).

2 Aggregates

Let $y_t = (y_{1t}, \ldots, y_{Kt})'$ be the vector of disaggregate component series and the aggregate of interest is $a_t = f_t(y_t)$. Here $f_t(\cdot)$ is a scalar aggregation function which may be time-varying. In fact, it is straightforward to extend the discussion to a vector process a_t. The assumption of a univariate process is made for clarity and notational convenience. A typical situation is that $f_t(y_t) = w_t' y_t$, where $w_t = (w_{1t}, \ldots, w_{Kt})'$ is a vector of stochastic weights.

The process a_t is assumed to be (covariance) stationary with AR and MA representations,

$$a_t = \sum_{i=1}^{\infty} \theta_i a_{t-i} + v_t, \quad t \in \mathbb{Z}, \tag{2.1}$$

and

$$a_t = \sum_{i=0}^{\infty} \psi_i v_{t-i}, \quad t \in \mathbb{Z}, \tag{2.2}$$

respectively, where v_t is a martingale difference white noise sequence with respect to $\mathcal{A}_t = \{a_t, a_{t-1}, \ldots\}$ with time-invariant variance σ_v^2. The martingale difference property implies that the conditional mean $E(v_t|\mathcal{A}_{t-1}) = 0$.

Notice that for an aggregate $a_t = w_t' y_t$ covariance stationarity holds under quite general conditions for w_t and y_t. For example, if the two processes are both covariance stationary and independent. If they are dependent, as may be the case in some of the motivating examples in Section 1, time-invariance of the first and second order moments is also not very restrictive as long as the two processes are covariance stationary individually. Some properties of specific aggregates with stochastic weights are derived in Appendix B.

Suppose that y_t^* is an M-dimensional (covariance) stationary stochastic process related to y_t. It may be equal to y_t or it may consist of a subset of the components of y_t. In fact, it may also contain transformations of the components of y_t or, more generally, any variables that are related to y_t. It is assumed that the joint process $x_t = (a_t, y_t^{*\prime})'$ is stationary with VAR and MA representations, respectively,

$$x_t = \sum_{i=1}^{\infty} A_i x_{t-i} + e_t, \quad t \in \mathbb{Z}, \tag{2.3}$$

and

$$x_t = \sum_{i=0}^{\infty} \Xi_i e_{t-i}, \quad t \in \mathbb{Z}, \tag{2.4}$$

where the A_i's and Ξ_i's are $((M+1)\times(M+1))$-dimensional coefficient matrices with $\Xi_0 = I_{M+1}$ and e_t is a $(M+1)$-dimensional martingale difference with respect to $\mathcal{X}_t = \{x_t, x_{t-1}, \dots\}$ with time-invariant, nonsingular covariance matrix $E(e_t e_t') = \Sigma_e$.

Nonsingularity of the covariance matrix Σ_e excludes that any of the components of x_t is a linear combination of the remaining components. Thus, if $a_t = w'y_t$ is an aggregate of y_t with fixed, time-invariant weights, $y_t^* = y_t$ is not possible. However, as long as at least one component series of y_t is missing from y_t^*, time-invariant aggregation is covered. Thus, the present framework can be used for studying forecasts of aggregates with fixed weights w. In that case y_t^* may, for instance, consist of the last $K-1$ components of y_t. Defining

$$F = \begin{bmatrix} & w' \\ 0 & I_{K-1} \end{bmatrix}$$

and $x_t = Fy_t$, the first component of x_t is the aggregate a_t. Because F is nonsingular, no loss in information is incurred by considering the transformed process x_t instead of y_t. Hence, any linear transformation of forecasts based on y_t can equivalently be obtained from forecasts of x_t.

If the aggregation weights are stochastic, considering the aggregate a_t in addition to the disaggregate component series y_t, that is, choosing $y_t^* = y_t$ leads to an extension of the information set because the information in the process w_t is also incorporated. These considerations imply that the aforementioned framework is useful not only for studying stochastic aggregation but any form of aggregation which implies a nonsingular joint process x_t, e.g., whenever e_t has a nonsingular covariance matrix. Thus, this framework covers general nonlinear aggregates of y_t.

The framework is useful even if some of the disaggregate component series are not available or if they are not known precisely. In that case one may just include all available component series in y_t^*. Also one may include series in y_t^* which are potentially useful indicators for the aggregate. For example, one may include factors which are computed from a large set of disaggregate component series. Moreover, the framework can be used if the weights are deterministic but still have changed over time, e.g., due to a structural break or a new member in some union.

Although the basic framework is quite general in many respects, it is restrictive in some dimensions. For example, there are no deterministic terms in the DGPs. They have been dropped for convenience in the present setup because they are not of much interest for the theoretical comparison of predictors. Since future values of deterministic terms are known, they do not contribute to forecast uncertainty. Hence, there is no need to consider them

in comparing predictors. Excluding them in the theoretical analysis is therefore not restrictive. Of course, they have to be included when models are fitted to actual data and this will be done in Sections 4 and 5.

As already mentioned earlier, the assumption that the aggregate a_t is a univariate variable is just a convenience. It is not very restrictive because allowing a_t to be a vector is a straightforward extension.

A more severe limitation of the present framework is the assumption of stationarity. Clearly many economic variables have stochastic trends, i.e., they are integrated and may have other nonstationarities. If unit roots are the only source of nonstationarity, this feature can be accounted for by considering differenced variables. Most of the results discussed in the following also hold for integrated processes. I will comment on this issue later. For the initial comparison of different predictors in the next section, the stationarity assumption is helpful because it simplifies the exposition and analysis.

Assuming that the residuals of the DGPs under consideration are martingale differences is a limitation which may be restrictive because it excludes potentially interesting nonlinearities. It is also made for convenience because it implies that optimal predictors are easy to determine. As will be seen in the next section, the predictors under scrutiny will still be best *linear* predictors if the white noise processes in the DGPs are at least serially uncorrelated and not necessarily martingale differences. In that case, the results on the relative efficiencies of the predictors are maintained.

3 The Predictors

If forecasts of a_t are desired, there are at least two obvious predictors. Suppose a forecast h periods ahead (an h-step forecast) is of interest in period τ. The optimal, minimum mean squared error (MSE) forecast based on present and past values of a_t is the conditional expectation,

$$a_{\tau+h|\tau} = E(a_{\tau+h}|a_\tau, a_{\tau-1}, \dots) = \sum_{i=1}^{\infty} \theta_i a_{\tau+h-i|\tau} = \sum_{i=h}^{\infty} \psi_i v_{\tau+h-i}, \qquad (3.1)$$

where $a_{\tau+h-i|\tau} = a_{\tau+h-i}$ for $i \geq h$. Thus, the AR form can be used for computing forecasts recursively for $h = 1, 2, \dots$, while the last expression on the right-hand side of (3.1) is useful for assessing the forecast error. The corresponding quantity is

$$a_{\tau+h} - a_{\tau+h|\tau} = \sum_{i=0}^{h-1} \psi_i v_{\tau+h-i} \sim \left(0, \sigma_a^2(h)\right), \qquad (3.2)$$

where

$$\sigma_a^2(h) = \sigma_v^2 \sum_{j=0}^{h-1} \psi_j^2 \tag{3.3}$$

is the forecast error variance which is also the forecast MSE. The fact that the conditional expectation can be expressed as in (3.1) is a consequence of the assumption that v_t is a white noise martingale difference sequence.

The second obvious predictor for a_t is based on the full process x_t. Let

$$x_{\tau+h|\tau} = E(x_{\tau+h}|x_\tau, x_{\tau-1}, \dots) = \sum_{i=1}^{\infty} A_i x_{\tau+h-i|\tau} = \sum_{i=h}^{\infty} \Xi_i e_{\tau+h-i} \tag{3.4}$$

be the optimal predictor based on present and past values of x_t. Its forecast error covariance or MSE matrix is

$$\Sigma_x(h) = E[(x_{\tau+h} - x_{\tau+h|\tau})(x_{\tau+h} - x_{\tau+h|\tau})'] = \sum_{j=0}^{h-1} \Xi_j \Sigma_e \Xi_j'. \tag{3.5}$$

The first component of $x_{\tau+h|\tau}$ is an h-step forecast of $a_{\tau+h}$ which will be denoted by $a^o_{\tau+h|\tau}$ in the following to distinguish it from $a_{\tau+h|\tau}$. The corresponding forecast error has zero mean and variance or MSE

$$\sigma_o^2(h) = \mathbf{i}'\Sigma_x(h)\mathbf{i}, \tag{3.6}$$

where $\mathbf{i} = (1, 0 \dots, 0)'$ is a $(M+1)$-dimensional vector.

In the literature on forecasting aggregates, a third predictor is occasionally considered. It is based on aggregating univariate forecasts of the disaggregate component series in y_t (e.g., Lütkepohl (2009)). In the present framework, where neither y_t nor the aggregation function $f_t(\cdot)$ are assumed to be known, such a predictor is not available in general. Therefore, it is not considered in the following although there are cases where it may be useful for practical purposes.

A well-known forecasting result for linearly transformed vectors implies that

$$\sigma_o^2(h) \leq \sigma_a^2(h) \tag{3.7}$$

(see, e.g., Lütkepohl (1987)). Moreover, stationarity of a_t implies that $a^o_{\tau+h|\tau}$ and $a_{\tau+h|\tau}$ become identical for $h \to \infty$ and, hence,

$$\lim_{h\to\infty} \sigma_o^2(h) = \lim_{h\to\infty} \sigma_a^2(h) = E(a_t^2). \tag{3.8}$$

Thus, the two predictors are equally efficient for long-term forecasting.

In fact, they may also be equal, and, hence, equally efficient under any loss function, for short forecast horizons. This will hold under very restrictive conditions for the DGP of y_t and the weighting scheme w_t. This fact is known already from the time-invariant $w_t = w \neq 0$ case. For that case, by Proposition 4.1 and Corollary 4.1.1 of Lütkepohl (1987),

$$a^o_{\tau+h|\tau} = a_{\tau+h|\tau} \text{ for all } h \in \mathbb{N} \quad \Leftrightarrow \quad w'\Phi(L) = \psi(L)w', \tag{3.9}$$

where $\Phi(L) = \sum_{i=0}^{\infty} \Phi_i L^i$ and $\psi(L) = \sum_{i=0}^{\infty} \psi_i L^i$ are the MA operators of y_t and a_t, respectively. The condition in (3.9) is quite restrictive. For example, if the components of y_t are independent and, hence, $\Phi(L)$ is a diagonal operator, equality of the predictors requires that all components with nonzero weight have identical MA operators and are, hence, homogeneous processes as far as their autocorrelation properties are concerned (e.g., Lütkepohl (1987, Corollary 4.1.2)). In other words, if $\Phi(L)$ is diagonal and the diagonal elements of $\Phi(L)$ are denoted by $\phi_{kk}(L)$, $k = 1, \ldots, K$, and all elements of w are nonzero, $\phi_{kk}(L) = \psi(L)$ must hold for $k = 1, \ldots, K$ for the predictors based on the disaggregated process and the aggregated process to be identical.

These considerations show that equality of $a^o_{\tau+h|\tau}$ and $a_{\tau+h|\tau}$ for all h holds only under very restrictive conditions. However, for practical purposes, using the predictor based on the aggregated process directly may be justified even if predictor equality holds only approximately. Even that requires a very special DGP of the disaggregate process y_t, however, and, hence, for the case of a time-invariant aggregation scheme, gains from using disaggregated information for short-term forecasting seem likely in typical situations. For long-term forecasts of stationary variables the disaggregate information is not helpful, however, because $a^o_{\tau+h|\tau}$ and $a_{\tau+h|\tau}$ will be similar or identical for large h.

For the general case, where $a_t = f_t(y_t)$, the discussion in Section 4.2.4 in Lütkepohl (1987) shows that equality of the two predictors, $a^o_{\tau+h|\tau} = a_{\tau+h|\tau}$ for all $h \in \mathbb{N}$, holds if and only if y^*_t does not Granger-cause a_t or, in other words, if there are no lagged y^*_t's in the a_t equation in (2.3). Thus, for the examples considered in the introduction, the forecast based on the disaggregate information is generally superior to $a_{\tau+h|\tau}$ because the components of y_t are included in a_t and, hence, are likely to enter the a_t equation in the model for x_t.

If the variables of interest have unit roots and are, hence, integrated, these general results will still hold because stationarity can be achieved by differencing. Since forecasts for the levels of the variables can be obtained directly from forecasts of the changes it is in fact enough to focus on stationary

processes for deriving theoretical results which also hold for MA processes with unit roots. MA unit roots are permitted, for example, in deriving the superiority of the forecast based on disaggregate components, that is, the result in (3.7). Also the condition for equality of predictors in (3.9) holds if $\Phi(L)$ is noninvertible (see Lütkepohl (1987)). Thus, even if all components are differenced individually without regard of possible cointegration, the previous theoretical analysis remains valid. For practical purposes one may still work with the levels VAR version of the process and avoid the detour via the differences. What does not hold for integrated processes is the convergence of the forecast MSEs in (3.8). In other words, even for long-term forecasting, $a^o_{\tau+h|\tau}$ may be superior to $a_{\tau+h|\tau}$ if a_t has a unit root.

Hendry and Hubrich (2010) consider this framework although they assume time-invariant aggregation weights throughout much of their theoretical analysis. They consider the predictors $a^o_{\tau+h|\tau}$ and $a_{\tau+h|\tau}$ for U.S. inflation. Note, however, that the main focus of Hendry and Hubrich (2010) is an analysis of the impact of various types of model deficiencies.

Although the previous results suggest that theoretically a forecast of the aggregate should always be based on the disaggregate process x_t because the components of y_t are part of the aggregate and, hence, are related to a_t and supposedly should have a Granger-causal relation to a_t, the situation is more difficult in practice, as pointed out by Hendry and Hubrich (2010). The theoretical results refer to correctly specified models with known coefficients. In practice, models for the DGPs involved have to be estimated and specified on the basis of the available sample information. As shown already by Lütkepohl (1987), estimation and specification uncertainty can reverse the ranking of the MSEs of the two predictors. In fact, if a large set of component series is aggregated, specifying a joint multivariate model for the components and the aggregate may not be feasible because of degrees of freedom limitations. Even if just a moderate number of components is considered, fitting a multivariate model may involve estimation of many more parameters and, hence, a much larger estimation uncertainty than fitting a univariate model to the aggregate only. Thereby the forecast based on the disaggregate information may become inferior to the one based on the aggregate directly. Therefore Hendry and Hubrich (2010) propose to limit the number of components included in the multivariate model. For example, for an inflation forecast one may include only the most important sub-index information rather than considering all sectoral indices. Alternatively, one may summarize the information of a large number of components in a small number of factors which are included in the multivariate model. These possibilities will be explored in Section 5 in forecasting euro-area inflation.

Before empirical examples based on economic data are considered, the small sample implications of including disaggregate information in the forecast for an aggregate based on stochastic weights will be investigated in a Monte Carlo study in the next section.

4 Monte Carlo Experiment

The foregoing discussion suggests that forecast improvements for an aggregate will be determined to some extent by the number of disaggregate component series which are included in the information set. Clearly, the optimal number will depend on factors such as the underlying DGP of the disaggregate components, the aggregation weights and the available sample size. Specification and estimation uncertainty will play a crucial role in practice. Therefore I have conducted a small Monte Carlo experiment for investigating the impact of these factors in an ideal environment before real data are analyzed in Section 5.

4.1 Monte Carlo Setup

I use the following simulation setup. The disaggregate component series are generated by a VAR(1) process,

$$y_t = \nu + A_1 y_{t-1} + u_t, \tag{4.1}$$

where $A_1 = \text{diag}(\alpha_1, \ldots, \alpha_K)'$ is a diagonal matrix, $u_t \sim \mathcal{N}(0, I_K)$ is an iid Gaussian white noise process with independent components and identity covariance matrix. The components of y_t are independent. Hence, each component contains additional information. Of course, this does not mean that the additional information is valuable enough for forecasting the aggregate in order to compensate for the additional specification and estimation uncertainty incurred by an increase in the dimension of the model used for forecasting.

The aggregation weights w_t are generated by drawing random vectors w_t^* from a K-dimensional normal distribution,

$$w_t^* \sim \mathcal{N}(\mu, \sigma^2 I_K),$$

replacing all negative elements by zero and then defining

$$w_t = w_t^* \Big/ \sum_{k=1}^{K} w_{kt}^* \, .$$

Table 1: Data Generation Processes Used in Monte Carlo Simulations

DGP	K	ν	A_1	μ'	σ
1	5	0	$\text{diag}(0.8, 0.8, 0.8, 0.8, 0.8)$	$(10, 5, 5, 5, 5)$	0.1
2	5	0	$\text{diag}(0.95, 0.8, 0.5, 0.4, 0.3)$	$(10, 5, 5, 5, 5)$	0.1
3	5	0	$\text{diag}(0.9, -0.9, 0.5, -0.5, 0.5)$	$(10, 5, 5, 5, 5)$	0.1

Thus, for each time period t, the components of w_t are nonnegative and sum to one and, hence, have standard properties of aggregation weights. The mean vector μ is chosen such that the first component has a larger weight on average than the other ones. Again, different weights are not uncommon in practice. I have specified relatively large values of the components of μ and a small standard deviation σ so that the variation in the aggregation weights from one period to the next tends to be small. Small changes in the aggregation weights over time seem to be typical for many aggregates in practice. Thus, the aggregation scheme used in the simulations is meant to mimic aggregates relevant for empirical work.

The DGPs for the disaggregate components, y_t, used in the simulations are summarized in Table 1. The first DGP consists of independent components with identical coefficients. For such a process, using the disaggregate components does not result in forecast efficiency gains if the aggregation weights are time-invariant. Given the relatively small changes in the aggregation weights for DGP 1, the disaggregate information is expected to be of limited value for forecasting the aggregate. In fact, specification and estimation uncertainty may dominate and reduce the efficiency of multivariate forecasts. Notice, however, that in this case a_t is also an AR(1) if the aggregation weights are fixed, whereas a_t is not an AR(1) if the aggregation weights are stochastic, as shown in Appendix B. For our choice of parameter values it is well approximated by an AR(1), however.

DGPs 2 and 3 have inhomogeneous components. Taking them into account may improve the forecast RMSE of the aggregate. In DGP 3 the AR coefficients of two components are even negative. Hence, the serial correlation structure of these components is very different from that of the remaining ones. It follows from the results in Appendix B that for these processes a_t has properties quite different from an AR(1). Although in a panel of component series it may be more common that the components have similar serial correlation properties, there are also cases where the components are quite distinct. Price indices may serve as an example. For instance, the DGP for energy prices may be quite different from that of food prices. Thus, using

a DGP with quite different components may provide relevant insights for applied work. Also in these DGPs the component series are independent so that they individually contain additional information.

The simulations are intended to disclose the value of disaggregate information relative to additional specification and estimation uncertainty incurred by adding component series in the forecasting model. The first component of y_t is always used in y_t^* and then the other components of y_t are added to y_t^* according to their position number. As mentioned previously, the weights attached to the different components differ. The first component of y_t has a larger weight in all three DGPs. Adding it first in y_t^* mimics the situation that the forecaster knows which components have a larger weight in the aggregate and, hence, may be more important. In the empirical examples in Section 5, some information on the importance of the individual components is obviously available. It may be difficult to give a full ranking of the components of y_t according to their importance, however. This is in line with the situation in the three DGPs in Table 1 where the last four components have equal average weight.

Samples of different sizes are used. In each case the simulations of the y_t are started from an initial vector of zero and the first 50 observations are discarded to reduce the impact of the startup values. In the following, T denotes the net sample size used for estimation and forecasting. In other words, additional presample values are used in the estimation and VAR order selection which is based on AIC and SC. The actual presample values considered depend on the maximum lag order allowed for in the selection procedure. Only full VAR models with a constant term are fitted in the Monte Carlo study. The number of replications for each design is 5000. All computations are performed with MATLAB programs.

4.2 Monte Carlo Results

Monte Carlo results for different forecast horizons and sample sizes are presented in Figures 1 - 5. Figure 1 shows results based on DGP 1. It presents the percentage deviations of forecasts based on different numbers of disaggregate components from the RMSE of the corresponding AR forecast for the aggregate. In other words, different y_t^* vectors are considered. Results for different lag orders and forecast horizons $h = 1$ and 6 are shown. For VAR order p, the AR model for the aggregate is also based on order p. The sample size underlying the results in Figure 1 is $T = 100$. Thus, the sample size has a typical order of magnitude for macroeconomic studies.

Given the structure of DGP 1, the situation is close to the one described in Section 3 where no forecast efficiency gains from using disaggregate infor-

mation are obtained. Hence, the RMSE deviations from the RMSE of the univariate forecast for the aggregate are primarily due to estimation uncertainty. Hence, it is not surprising that disaggregation reduces the forecast accuracy. In fact, the RMSEs increase with the number of disaggregate components included in the model. Moreover, increasing the VAR order reduces the estimation precision and, hence, increases the forecast RMSE. The damage from using disaggregate information is slightly reduced when the forecast horizon increases.

Figure 2 shows the impact of the sample size. It is based on the same DGP as Figure 1 but uses sample sizes of $T = 250$. The general picture is the same as in Figure 1. The losses due to estimation are substantially reduced, however. In other words, the forecast RMSEs still increase with the VAR order and the number of disaggregate component series, but now the efficiency loss is much smaller than in Figure 1.

DGP 2 has heterogeneous component series and, hence, without estimation uncertainty, using disaggregate information should reduce the forecast RMSE. In Figure 3 it can be seen that this is indeed the case even when estimated processes are used. The figure is again based on samples of sizes $T = 100$. For forecast horizon $h = 1$, using disaggregate information clearly improves the forecast precision for VAR order 1. The best forecasts are obtained when only two disaggregate components are considered. Using more disaggregate information is still better than forecasting the aggregate directly but leads to reduced forecast precision relative to a model for the aggregate and two disaggregate components. When more lags are considered or larger forecast horizons are of interest, the smallest RMSE is obtained with only one additional disaggregate component. In fact, with more lags, the use of disaggregate information becomes again detrimental. For example, when VAR(4) models are used, adding more than one disaggregate component results in forecasts which are inferior to the corresponding univariate forecast for the aggregate.

For DGP 3 the components are very heterogeneous. This is reflected in Figure 4 in the more substantial gains in forecast precision due to using disaggregate information. Note in particular the change in scale in Figure 4 relative to the other figures. For 1-step forecasts, gains relative to a univariate model for the aggregate are obtained even if large models with four lags and all five disaggregate components (i.e., 6-dimensional models) are considered. Even in this case, including two disaggregate components gives the best 1-step forecasts. With fewer lags, more components improve the forecast efficiency. The situation is somewhat different for 6-step forecasts for this DGP. For such large forecast horizons there is not much to gain from disaggregation except when very few lags are considered.

The previous simulations show the impact of estimation uncertainty. The results in Figure 5 address the effects of model selection. It shows only percentage deviations of RMSEs of 1-step forecasts based on VARs selected by AIC and SC with maximum lag order $p_{\max} = 4$. The sample size is again $T = 100$. Taking into account the change in scaling, the losses and gains for DGPs 1 and 2 are similar to those in Figures 1 and 3 when VAR order $p = 1$ is used. In contrast, the RMSE gains due to disaggregate information in Figure 5 are much smaller for DGP 3 than those for the VAR(1) case in Figure 4 although they are still much better than for VAR order 4, say. Hence, if the actual VAR order is unknown, model selection criteria are clearly helpful for reducing the parameter space and improving forecast efficiency although there will be some losses due to specification uncertainty. For the present DGPs, both AIC and SC deliver roughly the same results. Given the simple VAR(1) structure of the disaggregate components, this outcome is not surprising, of course. For processes with more complicated correlation structure, the results based on AIC and SC may, of course, be different. Because the simulation results suggest that parsimony may be valuable for forecast precision, the SC criterion may be preferable.

In summary, the Monte Carlo results show that, depending on the properties of the DGP, using disaggregate information can be valuable for improving forecast precision for an aggregate. In particular, if the disaggregate components are rather homogenous adding many of them can be detrimental to the forecast accuracy, however. Thus, to obtain actual gains in forecast RMSE over a univariate forecast for the aggregate, limiting the parameter space by considering a small number of disaggregate components and/or a small number of lags is advantageous if samples of typical sizes in macroeconomics are available. Of course, if large samples are available and the disaggregate components are heterogeneous, adding more of them to the model may be a useful strategy for improving forecasts of the aggregate. Finally, depending on the persistence of the DGP, little or no gains can be expected from using disaggregate information for long-term forecasting of stationary DGPs.

5 Empirical Forecast Comparison

I use two datasets to explore the empirical relevance of the previous results. In both cases the aggregates can be thought of as having stochastic weights. The first one consists of euro-area unemployment rates and the second one is based on EMU inflation rates.

5.1 Unemployment

The dataset consists of monthly, seasonally adjusted unemployment rates for the euro-area (a_t) and 10 individual euro-area countries for the period 1995M1-2008M12. Thus, given the relatively large number of component series, I consider only a subset of the 16 EMU countries for which the aggregate a_t is computed. A number of smaller countries have been eliminated because their relative weight in the overall unemployment rate is limited. Details of the data and data sources are given in the Appendix.

I have fitted AR and VAR models to the overall euro-area unemployment rate and subsets of the available country series. Clearly including all 11 series in a full VAR would have severely reduced the degrees of freedom. One of the questions of interest in the following is how many series should be included in a model that accounts for disaggregate information. Hence, systems of different dimensions will be explored. All AR and VAR models are based on the levels unemployment rate series although unit roots are not rejected by standard tests for some country rates. As I have argued in Section 3, unit roots are no problem if AR and VAR models are used for forecasting. Whether they can be utilized to improve forecasts is not the focus of this study. If not accounting for unit roots reduces forecast efficiency, it will affect all predictors although perhaps not to the same extent.

Because the simulation results suggest that parsimonious models may be beneficial for forecasting, I fit full and subset AR and VAR models with an intercept. The maximum order is set at $p_{\max} = 6$ because this order appears to be large enough to capture much if not all of the serial dependence structure for the systems considered. When full AR and VAR models are considered the actual model order is chosen by a model selection criterion. As in the simulations in Section 4, AIC and SC are used for this purpose. Estimation of full models is done by OLS. When subset models are used, the subset restrictions are determined by a sequential elimination strategy for each equation separately when multivariate models are considered. The sequential procedure proceeds by estimating an equation by OLS and eliminating the regressor with the smallest absolute t-value. Then the restricted model is estimated and the next regressor with the smallest absolute t-ratio is eliminated until all absolute t-ratios are greater than 1.96. Thus, a 5% significance level is used as a threshold for eliminating regressors in the subset models. Once all zero restrictions for all the equations of a given model are determined, the restricted model is reestimated by GLS with the OLS covariance estimator of a full model with order $p_{\max}$ (see, Section 5.2.2 of Lütkepohl (2005) for details on the GLS procedure).

Following Hendry and Hubrich (2010), I also consider models with factors

which are computed as principle components from all 10 disaggregate country data. They present another way of obtaining parsimonious models. The factors are computed on the basis of the sample available for model specification and estimation. They are recomputed whenever the sample is extended by an observation. Then VAR models are fitted for the aggregate and one or more factors. Again full and subset VAR models are considered for systems including factors. The previously described specification procedures are used here as well.

Forecasts are computed based on models fitted to increasingly larger samples. The last four years are set aside for the forecast comparison. Thus, the first set of models is based on samples from 1995M1-2004M12 (i.e., 120 observations). Based on these models 1- to 12-step forecasts are produced for the period 2005M1-2005M12. Then one observation is added to the sample and model specification and estimation is repeated for all models. Thereby 48 1-step, 47 2-step etc. out-of-sample forecasts are produced and used for computing RMSEs. Notice that whenever the sample is increased by a new observation, a full new specification and estimation is performed and also new factors are computed. It is also worth emphasizing that multi-step forecasts are computed recursively as in Section 3 (see Equations (3.1) and (3.4)).

Relative forecast RMSEs for a number of different models and forecast horizons $h = 1, 6$ and 12 are presented in Table 2, where the univariate AR model specified with the parsimonious SC is used as a benchmark. In other words, numbers smaller than one in the table indicate that the corresponding model produced a smaller RMSE than the univariate full AR model for a_t for which SC has been used for lag order selection. The following observations emerge from Table 2:

1. Univariate models for the aggregate unemployment rate based on AIC and SC produce the same RMSEs, that is, AIC and SC select the same AR orders (all relative RMSEs of the AIC-AR model are 1).

2. Adding disaggregate individual country variables in the full VAR models helps reducing the RMSEs up to a certain point. The bivariate full VAR models with the aggregate EMU rate and the German rate produce substantially better forecasts than the univariate models. When more variables are added, RMSEs tend to go up, as in the simulations in Section 4. For example, the five-dimensional model comprised of the EMU, German, French, Spanish and Italian rates produce substantially larger RMSEs for all three horizons than the bivariate models.

3. Subset VAR models do not improve forecasts of univariate models and low-dimensional systems. They produce lower RMSEs than the cor-

responding full VAR models in some cases for 4- and 5-dimensional systems. Hence, parsimony considerations become important when the number of component series increases. This result can also be seen in Figure 6, where the RMSEs of subset VARs relative to the corresponding full VARs are shown. The subset models have a clear advantage only for the five-dimensional system.

4. The previous observation suggests that factor models may be useful to include information from a larger number of individual country variables. This is true only partially, however. Bivariate models with one factor produce reasonably small RMSEs when full VAR models are used. Adding a second factor leads to substantially larger RMSEs across all models. In fact, the two-factor models are substantially inferior to the benchmark models. Hence, they do not even improve on direct univariate forecasts of the aggregate. Generally, specifying factor models with subset procedures leads to poor forecasts.

In summary, the forecasting results for the euro-area unemployment rate are in line with the theoretical and simulation results. They confirm that disaggregate information is valuable for improving forecasts as long as the disaggregate models do not get very large. If too much disaggregate information is included in the model, the lack of parsimony may become an obstacle at the specification and estimation stage and the resulting models may lead to substantially inferior forecasts for the aggregate.

5.2 Inflation

The second example considers consumer price indices. The dataset consists of monthly, seasonally unadjusted euro-area HICP series for the period 1990M1-2008M12. The euro-area HICP is obtained from the individual countries' HICPs with weights based on the countries' relative household expenditure shares which are updated annually. Thus, the weights may be viewed as stochastic although they do not change every month. HICPs from 1990M1 are available for the aggregated euro-area and seven individual countries: Austria (A), Finland, France (F), Greece, Italy (I), Netherlands (NE) and Portugal. Only shorter HICP series are available for the other EMU member states. Therefore the first forecasting experiment is based on eight series which include the euro-area HICP. Details on the variables and data sources are again provided in the Appendix. From each price index, say p_t, I have computed an annual inflation rate as $(p_t - p_{t-12})/p_{t-12}$, that is, I use the actual rate and not the approximation based on annual differences of logarithms. Results by Lütkepohl and Xu (2009) suggest that using levels rather than

logarithms of price indices may be preferable for forecasting. All forecasting models are based on the inflation rates rather than the price indices. Thereby I lose one year of data. That is, the full sample for the inflation rates is 1991M1-2008M12.

There is little seasonality left in the inflation rates. I have still considered higher order models in this case to capture remaining seasonality. More precisely, I have used lag orders of up to 14. No further precautions for seasonality are considered, however. All AR and VAR models used in the following forecast comparison contain a constant but no trends or seasonal dummies.

The modelling and estimation strategy is the same as for the unemployment rates. The last four years are used for the forecast comparison. Thus, 1-step RMSEs are again based on 48 forecasts while 12-step forecast RMSEs are based on only 37 forecasts. Forecast horizons from $h = 1$ to 12 are used as for the unemployment rates.

My forecast experiment bears some similarity to that of Hendry and Hubrich (2010). A main difference is, of course, that I consider European inflation while the latter authors study U.S. inflation. Moreover, I use regionally disaggregated data whereas Hendry and Hubrich (2010) use sub-indices. Also, I do not consider structural change although it may be present due to the introduction of the euro within the sample period.

Forecast RMSEs relative to univariate AR models for the aggregate based on SC for a range of different models are presented in Table 3. Although the combinations of countries used in the table may seem a bit arbitrary, the results are representative for other systems as well. The bivariate systems were chosen because France and Italy are the largest countries in the panel and, hence, are expected to have the greatest impact on euro-area inflation. It turned out that the corresponding bivariate models indeed lead to the smallest 1-step RMSEs. I have experimented with a number of other combinations of country specific inflation rates and found qualitatively similar results to those shown in Table 3. The following observations emerge from the table:

1. Full AR/VAR models tend to outperform the more parsimonious subset VAR models. In fact, only for 1-step forecasts the univariate subset model outperforms the other models. In all other cases, the subset models are inferior to at least one of the other forecasts based on full models.

2. The more profligate full AIC models outperform the more parsimonious competitors for 1-step forecasting and also tend to be superior to the

corresponding SC models for longer-term annual forecasts (12 steps ahead). On the other hand, the 6-month forecasts of the univariate SC benchmark model are hard to beat. Only the bivariate SC-VAR model based on the EMU and French series produced a slightly better RMSE.

3. Increasing the number of disaggregate components tends to improve the 1-step and 12-step forecasts up to a certain point after which the forecast RMSEs increase. For example, for the full models, the AIC-VAR for the overall inflation rate and three country series has the smallest 1-step and 12-step RMSEs whereas adding another component leads to an increase in the RMSEs.

4. The AIC model with one factor produces the lowest 1-step RMSE whereas for larger forecast horizons the factor models do not improve on the forecasts of the univariate benchmark model.

Because the improvements over the univariate benchmark model are in no case impressive, I have also considered using sub-indices instead of disaggregate country data. Euro-area inflation forecasts based on aggregating sub-indices have also been investigated by Hubrich (2005). In contrast to the approach used in the following, she does not include the overall HICP inflation in her multivariate disaggregate model for the sub-indices. As discussed in Section 2, using the aggregate and disaggregate components in one model is covered by the present framework even if the weights of the sub-indices are time-invariant, as long as only a subset of sub-indices is included. The HICP is basically a Laspeyres index based on a fixed basket. Hence, the weights for sub-indices are theoretically constant. In practice, new items are added occasionally, however, which may affect the weighting scheme. The virtue of the present approach is that it permits to take advantage of the information in the available component series without knowing the exact weighting scheme. The dataset consists of the overall price index and eight sub-indices. Details are again presented in the Appendix.

Using this dataset I have produced forecasts for inflation rates in the same way as for the dataset with individual country data. Results are presented in Table 4. Overall they are surprisingly similar to those in Table 3. In other words, generally subset models do not perform well and sizable improvements in short-term forecast accuracy are obtained by using full AIC-VAR models. In contrast to the previous data set, SC-VAR models now deliver the best 12-step forecasts. The optimal 12-step forecast is actually obtained with a bivariate model for the overall index and one factor.

In summary, parsimony has an advantage in particular for longer-term forecasting, where parsimony refers to both the model order and the number

of component series included in the model. Using either a very small number of components or a factor which summarizes the information in a larger set of HICP components leads to forecast improvements. For short-term forecasting the situation appears to be somewhat different in that more heavily parameterized models with larger VAR orders and more components tend to perform better than very small models. However, even for short-term forecasting adding too many components in the models is detrimental to forecast precision. These results are overall in line with those of Hendry and Hubrich (2010) for U.S. inflation. They also found that disaggregate information can help improving inflation forecasts. In contrast, Hubrich (2005) did not find disaggregation helpful for improving forecast accuracy. Her result is based on five disaggregate components and highlights the advantage of including the aggregate and a small subset of disaggregate components in the forecasting model as done in the present study.

6 Conclusions

In this study a framework is presented for analyzing forecasts for aggregated variables obtained by nonlinear and possibly stochastic aggregation. Such aggregates are common in economics. The main advantage of the proposed framework is that the precise form of aggregation and in particular the aggregation weights may be unknown, as is often the case in practice. Using the proposed setup, forecasts based on the aggregate directly and forecasts which take into account at least some of the disaggregate information are compared theoretically and via simulations. It is found that theoretically forecasts that take into account disaggregate information will generally lead to more efficient forecasts in terms of MSE. In practice, including too many disaggregate components can lead to efficiency losses because specification and estimation uncertainty will increase with the number of disaggregate components in the model used for forecasting. Hence, the additional information content in further disaggregate components may be insufficient to compensate for increases in specification and estimation uncertainty.

Two examples are considered to illustrate the practical relevance of the theoretical analysis. The first one is based on the euro-area unemployment rate. It is found that using disaggregate information from the individual member states may result in forecast improvements if the number of component series is small and parsimonious models are fitted. These results are in line with the simulation evidence obtained from artificial DGPs. The second example is based on EMU inflation and also illustrates that the information in disaggregate price indices for different countries or different sectors helps

to improve forecasts as long as small numbers of component series are used for forecasting.

These results raise questions for further research. In particular, it may be of interest to determine the optimal level of disaggregation. Given the results of this paper, it will depend on the DGP of the variables of interest and, hence, optimal disaggregation is ultimately an empirical question. Still, it may be possible to find characteristics of component series which give an indication that they may be helpful in a model for the aggregate. Such questions are left for future research.

Appendix A. Data and Their Sources

All data were extracted from the Statistical Data Warehouse of the European Central Bank. The original source is Eurostat. The precise specifications for the two datasets are as follows.

Unemployment rates

The series are standardized unemployment series as collected by Eurostat. All ages and males and females are included. The series are monthly frequency. They are seasonally adjusted by Eurostat but not working day adjusted. The total sample period is from January 1995 to December 2008. The following eleven series are included in the panel: Euro-area 16 (fixed composition), Austria, Belgium, Germany, Spain, Finland, France, Ireland, Italy, Netherlands, Portugal.

HICP consumer prices

The series are harmonized consumer prices at monthly frequency as collected by Eurostat for the euro-area (changing composition). They are neither seasonally nor working day adjusted. The monthly values of annual inflation rates are obtained from a price index p_t as $(p_t - p_{t-12})/p_{t-12}$. The sample period for the price index data is from January 1990 to December 2008 and, hence, the inflation rates range from January 1991 to December 2008. The base year is 2005. HICP series from the following seven countries are considered: Austria, Finland, France, Greece, Italy, Netherlands and Portugal. In addition overall euro-area HICP is used in the first dataset. The factors are computed from all seven country series.

The following seven series are included in the panel underlying the sectorally disaggregated dataset: HICP - Overall index, Food including alcohol and tobacco, Processed food including alcohol and tobacco, Unprocessed

food, Industrial goods excluding energy, Energy, Services. The factors are computed from the following five series: Processed food including alcohol and tobacco, Unprocessed food, Industrial goods excluding energy, Energy, Services.

Appendix B. Properties of Aggregates with Stochastic Aggregation Weights

Suppose the aggregation weights $w_t \sim (\mu, \Sigma_w)$ are iid vectors which are independent of y_t. Then the aggregate $a_t = w_t' y_t$ has the following properties:

1. $E(a_t) = \mu' E(y_t)$ and

$$\begin{aligned}
\operatorname{cov}(a_t, a_{t-j}) &= E[w_t'(y_t - Ey_t)w_{t-j}'(y_{t-j} - Ey_{t-j})] \\
&= E[\operatorname{tr}\{(y_t - Ey_t)(y_{t-j} - Ey_{t-j})' w_t w_{t-j}'\}] \\
&= \operatorname{tr}\{E[(y_t - Ey_t)(y_{t-j} - Ey_{t-j})']E[w_t w_{t-j}']\} \\
&= \begin{cases} \operatorname{tr}\{\operatorname{cov}(y_t, y_{t-j})(\Sigma_w + \mu\mu')\} & \text{for } j = 0, \\ \operatorname{tr}\{\operatorname{cov}(y_t, y_{t-j})\mu\mu'\} & \text{for } j \neq 0. \end{cases}
\end{aligned}$$

   Thus, a_t is covariance stationary if y_t has this property.

2. If y_t is a finite order MA(q) process, then the same is true for a_t because of the mean and covariance properties given under 1.

3. If $y_t = \rho I_K y_{t-1} + u_t$, where $u_t \sim (0, \sigma_u^2 I_K)$ is white noise, that is, y_t has uncorrelated components with identical correlation structure and $w_t \sim (\mu, \sigma_w^2 I_K)$, then

$$E(a_t a_{t-j}) = \begin{cases} \dfrac{\sigma_u^2}{1-\rho}\left(K\sigma_w^2 + \displaystyle\sum_{k=1}^{K} \mu_k^2\right) & \text{for } j = 0, \\ \dfrac{\rho^j \sigma_u^2}{1-\rho} \displaystyle\sum_{k=1}^{K} \mu_k^2 & \text{for } j \neq 0. \end{cases}$$

   Thus, a_t is not an AR(1), but is close to having AR(1) dynamics if σ_w^2 is small relative to $\sum_{k=1}^{K} \mu_k^2$.

4. If $y_t = \operatorname{diag}(\rho_1, \ldots, \rho_K) y_{t-1} + u_t$, a_t will in general not be an AR(1).

These properties are useful to know in assessing the DGPs used in the simulations in Section 4.

References

Anderson, H., Dungey, M., Osborn, D. R. and Vahid, F. (2007). Constructing historical euro area data, *Discussion Paper*, Australian National University.

Beyer, A., Doornik, J. A. and Hendry, D. F. (2001). Reconstructing historical euro-zone data, *Economic Journal* **111**: 308–327.

Hendry, D. F. and Hubrich, K. (2010). Combining disaggregate forecasts or combining disaggregate information to forecast an aggregate, *Journal of Business & Economic Statistics*.

Hubrich, K. (2005). Forecasting euro area inflation: Does aggregating forecasts by HICP component improve forecast accuracy?, *International Journal of Forecasting* **21**: 119–136.

Lütkepohl, H. (1987). *Forecasting Aggregated Vector ARMA Processes*, Springer-Verlag, Berlin.

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*, Springer-Verlag, Berlin.

Lütkepohl, H. (2009). Forecasting aggregated time series variables - a survey, *Working Paper EUI ECO 2009/17*, European University Institute.

Lütkepohl, H. and Xu, F. (2009). Forecasting annual inflation with seasonal monthly data: Using levels versus logs of the underlying price index, *Working Paper MWP2009/37*, European University Institute.

Winder, C. C. A. (1997). On the construction of European area-wide aggregates - a review of issues and empirical evidence, *IFC Bulletin* **1**: 15–23.

Table 2: Forecast RMSEs for Euro-Area Unemployment Relative to Univariate AR Model Selected by SC

		forecast horizon		
Model type	Model	1-step	6-step	12-step
full AR/VAR	SC-VAR$^{\text{EMU,D}}$	0.9504	0.8227	0.6737
	SC-VAR$^{\text{EMU,D,F,E}}$	0.9773	0.7716	0.6966
	SC-VAR$^{\text{EMU,D,F,I}}$	0.9847	0.8629	0.8801
	SC-VAR$^{\text{EMU,D,F,E,I}}$	1.2519	1.1613	1.1370
	SC-VAR$^{\text{EMU,f1}}$	0.9466	0.8006	0.6818
	SC-VAR$^{\text{EMU,f1,f2}}$	1.1357	1.3319	1.5271
	AIC-AR$^{\text{EMU}}$	1.0000	1.0000	1.0000
	AIC-VAR$^{\text{EMU,D}}$	0.9227	0.8080	0.6508
	AIC-VAR$^{\text{EMU,D,F,E}}$	0.9293	0.7144	0.5847
	AIC-VAR$^{\text{EMU,D,F,I}}$	0.9627	0.8006	0.7905
	AIC-VAR$^{\text{EMU,D,F,E,I}}$	1.0799	0.9854	0.9610
	AIC-VAR$^{\text{EMU,f1}}$	0.9321	0.7949	0.6829
	AIC-VAR$^{\text{EMU,f1,f2}}$	1.0727	1.1313	1.3272
subset AR/VAR	AR$^{\text{EMU}}$	1.0090	1.0324	0.9975
	VAR$^{\text{EMU,D}}$	1.1334	1.1796	0.9522
	VAR$^{\text{EMU,D,F,E}}$	1.0145	0.7167	0.6257
	VAR$^{\text{EMU,D,F,I}}$	1.2888	0.9137	1.0663
	VAR$^{\text{EMU,D,F,E,I}}$	1.2351	0.8028	0.9235
	VAR$^{\text{EMU,f1}}$	1.0505	0.9154	1.0643
	VAR$^{\text{EMU,f1,f2}}$	1.2047	1.0526	1.2614

Notes: Total sample period: 1995M1-2008M12; forecast period: 2005M1-2008M12; maximum lag order $p_{\max} = 6$; EMU – European Monetary Union (euro-area), D – Germany, F – France, E – Spain, I – Italy; f1 – first factor, f2 – second factor.

Table 3: Forecast RMSEs for Euro-Area Inflation Based on Country Specific HICPs Relative to Univariate AR Benchmark Model Selected by SC

		forecast horizon		
Model type	Model	1-step	6-step	12-step
full AR/VAR	SC-VAR$^{\text{EMU,F}}$	1.0088	0.9683	0.9400
	SC-VAR$^{\text{EMU,I}}$	0.9885	1.0028	0.9982
	SC-VAR$^{\text{EMU,F,I,NE}}$	1.0272	1.0473	1.0418
	SC-VAR$^{\text{EMU,F,I,NE,A}}$	1.0428	1.0290	1.0161
	SC-VAR$^{\text{EMU,f1}}$	1.0180	1.0332	1.0039
	SC-VAR$^{\text{EMU,f1,f2}}$	1.0300	1.0615	1.0200
	AIC-AR$^{\text{EMU}}$	0.9608	1.0764	1.0046
	AIC-VAR$^{\text{EMU,F}}$	0.9314	1.0261	0.9911
	AIC-VAR$^{\text{EMU,I}}$	0.9380	1.0624	0.9856
	AIC-VAR$^{\text{EMU,F,I,NE}}$	0.9314	1.0702	0.9666
	AIC-VAR$^{\text{EMU,F,I,NE,A}}$	0.9743	1.0030	0.9721
	AIC-VAR$^{\text{EMU,f1}}$	0.9227	1.0619	1.0392
	AIC-VAR$^{\text{EMU,f1,f2}}$	0.9376	1.0326	0.9996
subset AR/VAR	AR$^{\text{EMU}}$	0.9228	1.0794	1.0032
	VAR$^{\text{EMU,F}}$	0.9825	1.0480	0.9463
	VAR$^{\text{EMU,I}}$	1.0466	1.0404	1.0080
	VAR$^{\text{EMU,F,I,NE}}$	1.1116	1.0526	0.9833
	VAR$^{\text{EMU,F,I,NE,A}}$	1.1405	1.1070	0.9893
	VAR$^{\text{EMU,f1}}$	0.9683	1.1522	1.0781
	VAR$^{\text{EMU,f1,f2}}$	0.9520	1.0762	1.0802

Notes: Total sample period: 1991M1-2008M12; forecast period: 2005M1-2008M12; maximum lag order $p_{\max} = 14$; EMU – European Monetary Union, A – Austria, F – France, I – Italy, NE – Netherlands; f1 – first factor, f2 – second factor.

Table 4: Forecast RMSEs for Euro-Area Inflation Based on Sector Specific HICPs Relative to Univariate AR Benchmark Model Selected by SC

		forecast horizon		
Model type	Model	1-step	6-step	12-step
full AR/VAR	SC-VAR$^{\text{all,F}}$	0.9865	0.9843	0.9641
	SC-VAR$^{\text{all,F,IG}}$	1.0163	1.0070	0.9540
	SC-VAR$^{\text{all,F,IG,E}}$	1.0222	1.0432	0.9722
	SC-VAR$^{\text{all,F,IG,E,S}}$	1.0317	1.0425	0.9834
	SC-VAR$^{\text{all,f1}}$	1.0090	0.9726	0.9505
	SC-VAR$^{\text{all,f1,f2}}$	1.0079	0.9763	0.9594
	AIC-AR$^{\text{all}}$	0.9608	1.0764	1.0046
	AIC-VAR$^{\text{all,F}}$	0.9505	1.0737	0.9924
	AIC-VAR$^{\text{all,F,IG}}$	0.9560	1.1666	1.1089
	AIC-VAR$^{\text{all,F,IG,E}}$	0.9368	1.0768	1.0768
	AIC-VAR$^{\text{all,F,IG,E,S}}$	0.9716	1.1316	1.2106
	AIC-VAR$^{\text{all,f1}}$	0.9450	1.0861	1.0319
	AIC-VAR$^{\text{all,f1,f2}}$	0.9715	1.1051	1.0309
subset AR/VAR	AR$^{\text{all}}$	0.9228	1.0794	1.0032
	VAR$^{\text{all,F}}$	1.2253	1.1581	1.0854
	VAR$^{\text{all,F,IG}}$	1.0146	1.1623	1.0789
	VAR$^{\text{all,F,IG,E}}$	1.0241	1.1721	1.0686
	VAR$^{\text{all,F,IG,E,S}}$	1.1032	1.1488	1.0406
	VAR$^{\text{all,f1}}$	1.1060	1.0674	0.9999
	VAR$^{\text{all,f1,f2}}$	1.0562	1.0974	1.0653

Notes: Total sample period: 1991M1-2008M12; forecast period: 2005M1-2008M12; maximum lag order $p_{\max} = 14$; all – overall HICP, F – Food incl. alcohol and tabacco, IG – Industrial goods excluding energy, E – Energy, S – Services; f1 – first factor, f2 – second factor.

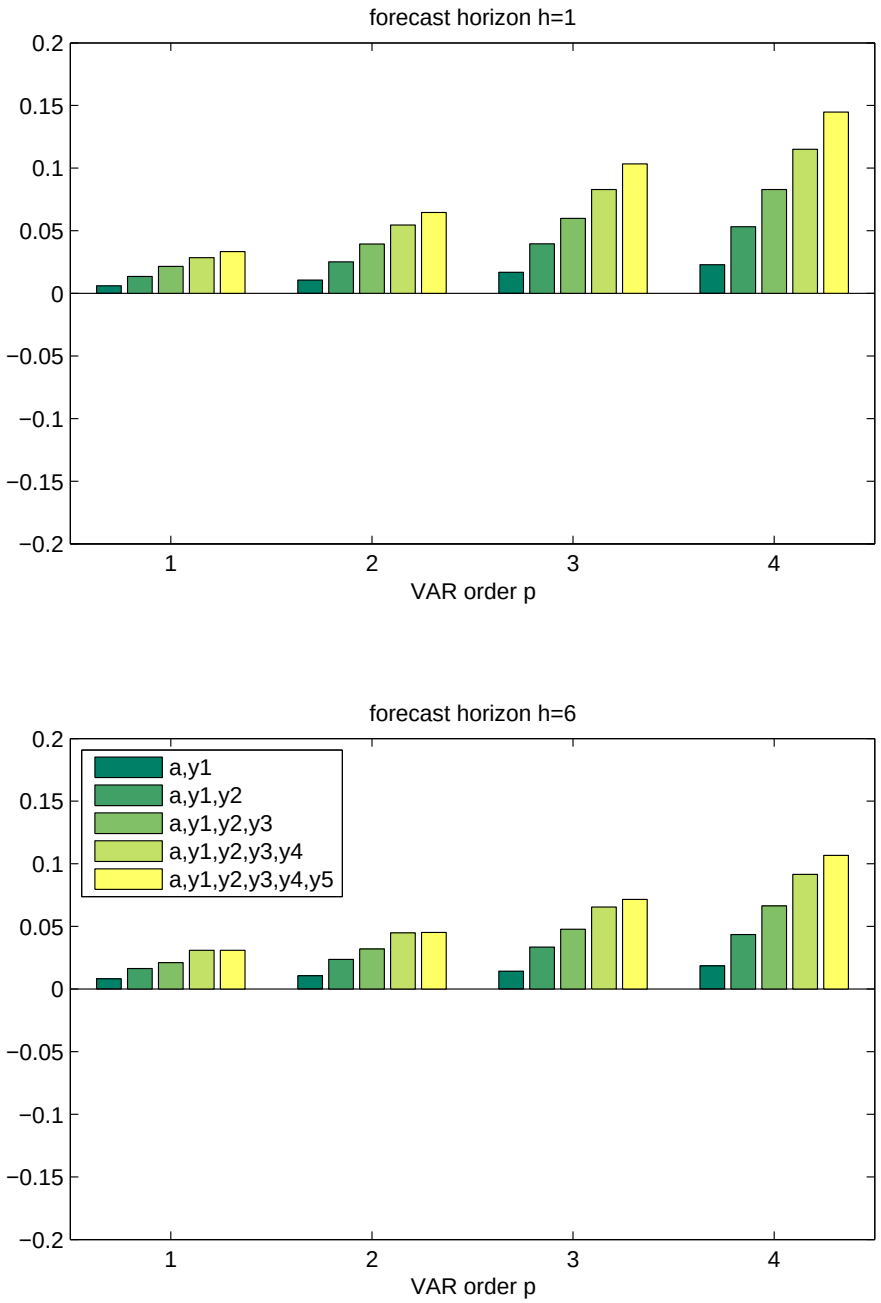

Figure 1: Percentage deviations of forecast RMSEs based on VARs with disaggregate components from RMSE of AR model for aggregate (DGP 1, $T = 100$).

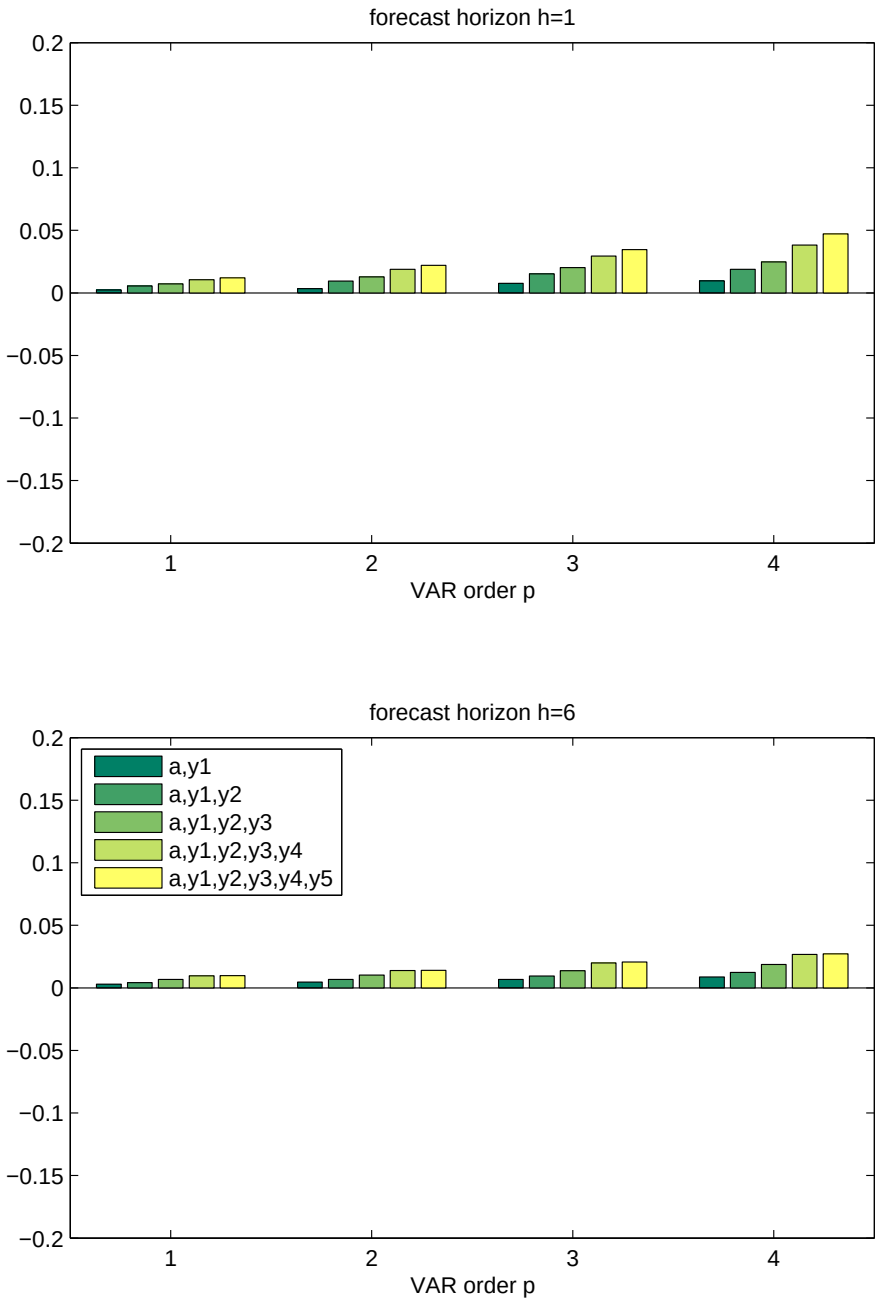

Figure 2: Percentage deviations of forecast RMSEs based on VARs with disaggregate components from RMSE of AR model for aggregate (DGP 1, $T = 250$).

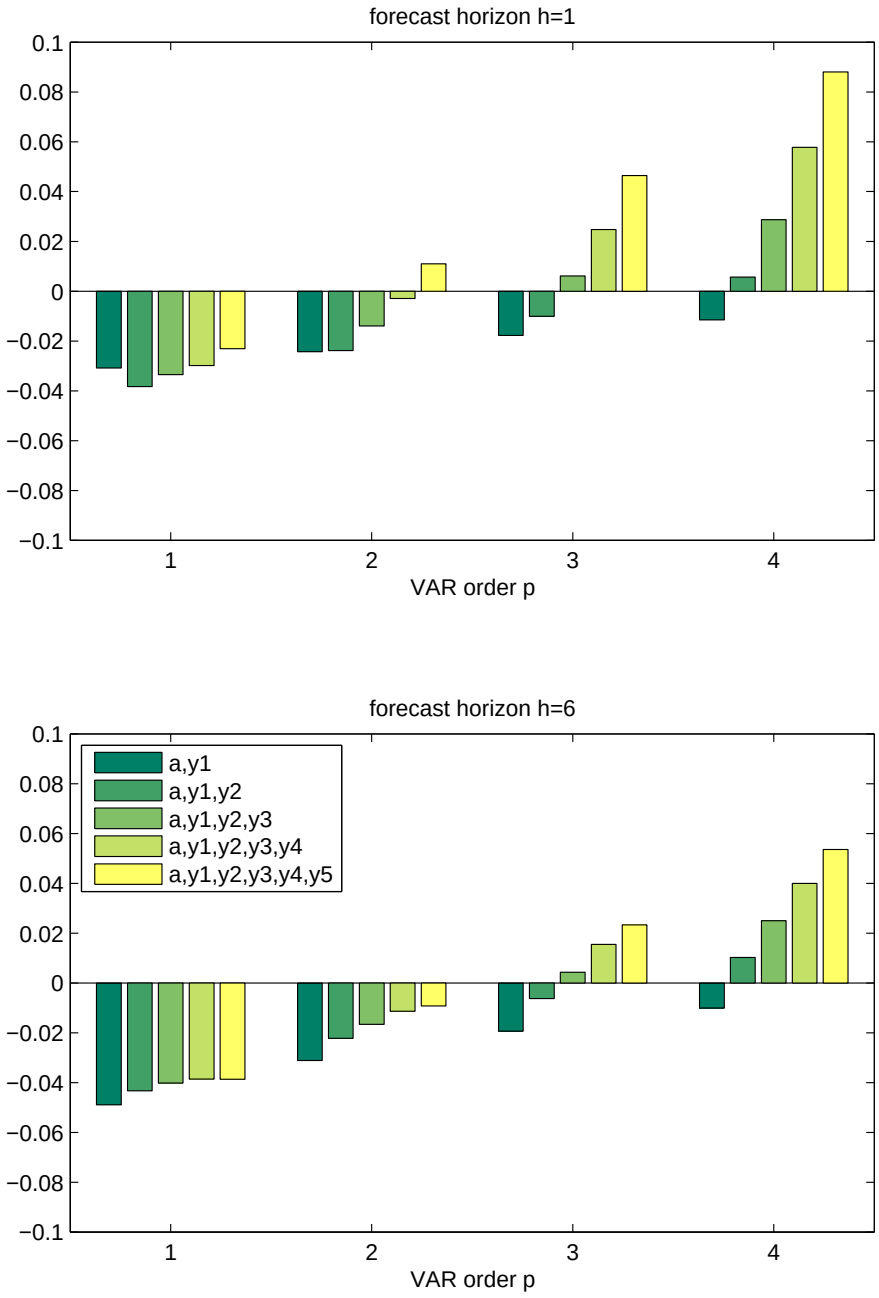

Figure 3: Percentage deviations of forecast RMSEs based on VARs with disaggregate components from RMSE of AR model for aggregate (DGP 2, $T = 100$).

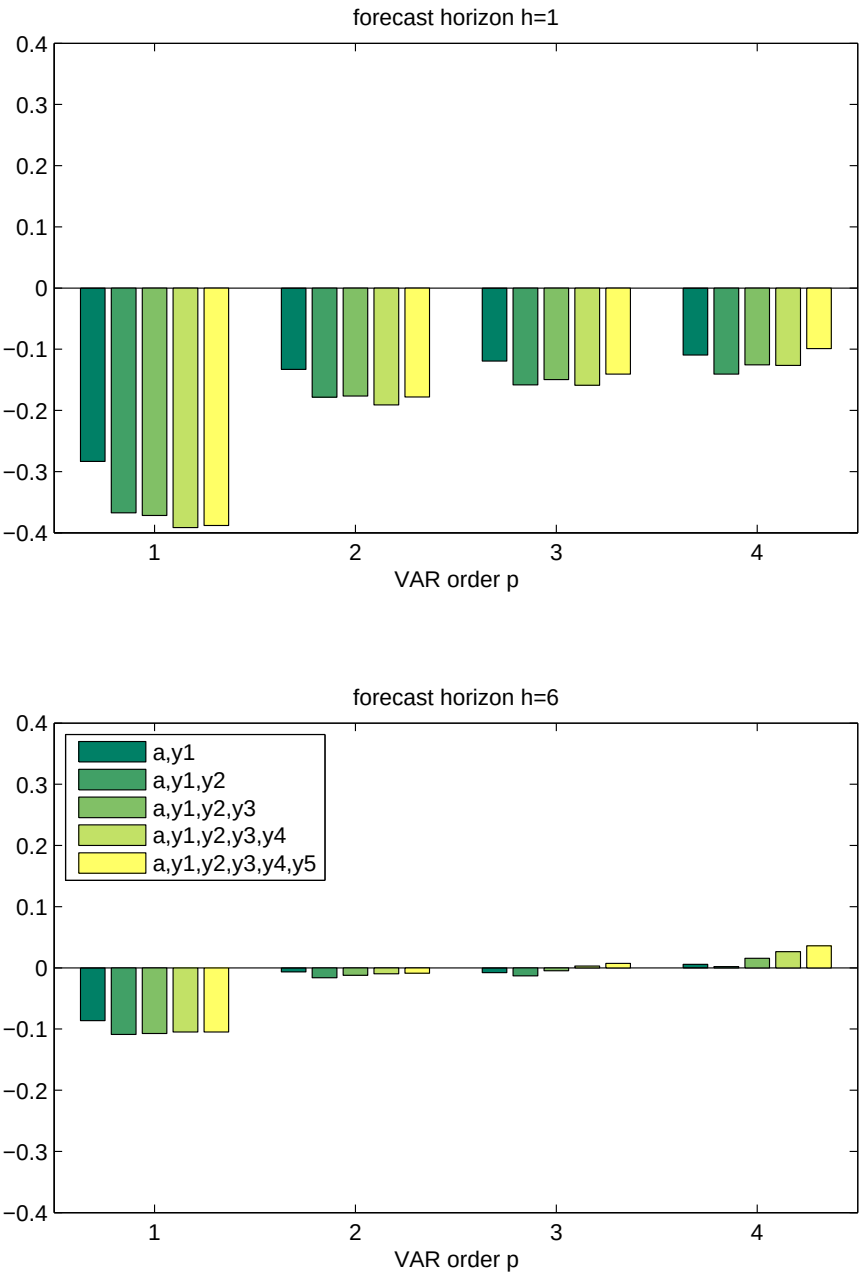

Figure 4: Percentage deviations of forecast RMSEs based on VARs with disaggregate components from RMSE of AR model for aggregate (DGP 3, $T = 100$).

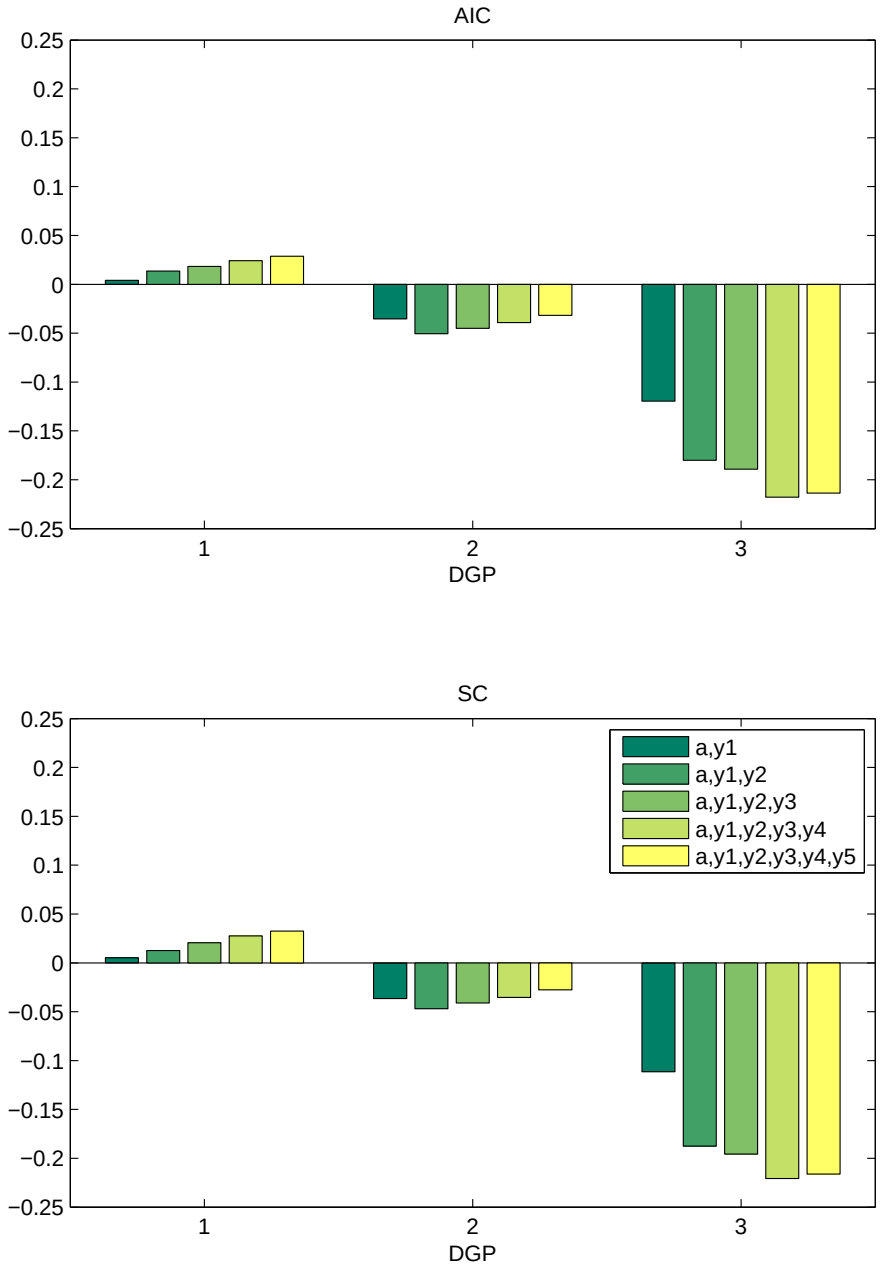

Figure 5: Percentage deviations of 1-step forecast RMSEs based on VARs with disaggregate components selected by AIC and SC from RMSE of AR model for aggregate ($T = 100$).

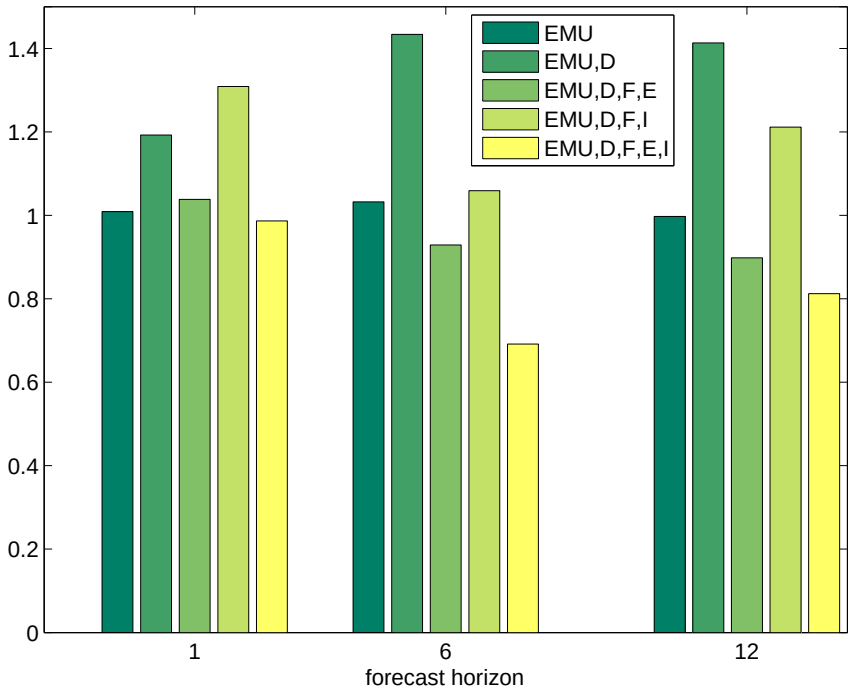

Figure 6: RMSEs of subset AR/VAR models for unemployment relative to full SC-AR/VAR models.


European
University
Institute
European University Press