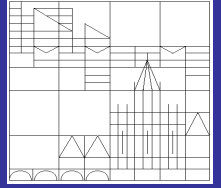




University of Konstanz
Department of Economics



Forecasting Contemporaneous Aggregates with Stochastic Aggregation Weights

Ralf Brüggemann and Helmut Lütkepohl

Working Paper Series
2011-23

<http://www.wiwi.uni-konstanz.de/workingpaperseries>

Forecasting Contemporaneous Aggregates with Stochastic Aggregation Weights¹

Ralf Brüggemann

Department of Economics, University of Konstanz
Box 129, D-78457 Konstanz, GERMANY
email: ralf.brueggemann@uni-konstanz.de

and

Helmut Lütkepohl

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

April 21, 2011

Abstract. Many contemporaneously aggregated variables have stochastic aggregation weights. We compare different forecasts for such variables including univariate forecasts of the aggregate, a multivariate forecast of the aggregate that uses information from the disaggregate components, a forecast which aggregates a multivariate forecast of the disaggregate components and the aggregation weights, and a forecast which aggregates univariate forecasts for individual disaggregate components and the aggregation weights. In empirical illustrations based on aggregate GDP and money growth rates, we find forecast efficiency gains from using the information in the stochastic aggregation weights. A Monte Carlo study confirms that using the information on stochastic aggregation weights explicitly may result in forecast mean squared error reductions.

Key Words: Aggregation, autoregressive process, mean squared error

JEL classification: C32

¹Corresponding author: Ralf Brüggemann, Department of Economics, University of Konstanz, Box 129, D-78457 Konstanz, GERMANY; email: ralf.brueggemann@uni-konstanz.de

1 Introduction

Many economic variables which are contemporaneous aggregates of a number of disaggregate variables have time-varying aggregation weights. For example the European Union (EU) growth rate is an average of the growth rates of the individual member states weighted by the relative shares of overall output. The EU unemployment rate is the weighted average of the individual member states' unemployment rates with weights being the relative shares of the respective labor forces. As another example consider North American output which is the sum of the outputs of the northern American countries weighted by the exchange rates. In these examples the aggregation weights are actually best thought of as stochastic.

Despite the stochastic nature of the weights of many aggregates, most previous studies on forecasting contemporaneously aggregated variables focus on aggregation with fixed, time-invariant weights. Examples are Ansley, Spivey and Wroblewski (1977), Tiao and Guttman (1980), Wei and Abraham (1981), Kohn (1982), Lütkepohl (1984a, 1984b, 1986, 1987). See also the survey by Lütkepohl (2010). These studies suggest that taking into account disaggregate information is theoretically helpful for reducing the forecast mean squared error (MSE). However, specification and estimation uncertainty may reduce or even reverse the gains, in particular, when higher-dimensional multivariate models are fitted to disaggregate data. Therefore some studies also compare aggregates of univariate forecasts of the disaggregate components and find that such forecasts may outperform aggregated multivariate forecasts. Also parameter reduction methods such as subset vector autoregressions as in Hubrich (2005) or factor models as in Hendry and Hubrich (2011) have been considered in this context. The results are not uniform across studies and depend to some extent on the data generation process (DGP). Overall there is evidence that taking into account disaggregate information can improve forecast efficiency for contemporaneous aggregates with fixed weights if methods are used which limit the estimation and specification uncertainty. Empirical studies confirming this conclusion are, for instance, provided by Marcellino, Stock and Watson (2003), Espasa, Senra and Albacete (2002) and Carson, Cenesizoglu and Parker (2010).

The fact that many aggregates have time-varying weights was recognized by Lütkepohl (2011) who developed a general framework for comparing predictors for such aggregates based on aggregate and disaggregate information. In that framework the process generating the time-varying, possibly stochastic weights is not explicitly considered, however. Hence, any information in that process is ignored or taken into account only indirectly for forecasting purposes. In practice, such an approach has its advantages if the aggrega-

tion weights are not available or unobservable. On the other hand, there are also many cases where past aggregation weights are available. In this study we focus on that case and investigate whether it is worthwhile to take the information in the weights explicitly into account in forecasting.

There are a number of different predictors which can be used in this situation. For instance, one may model the disaggregate series and the aggregation weights separately, forecast them and then aggregate the forecasts using the predicted weights or one may construct a joint model for the disaggregate series and the aggregation weights and aggregate the forecasts from such a model. Obviously, this forecasting strategy may quickly result in very high-dimensional models even if only a few disaggregate components are involved and in practice we often have to deal with very large panels of disaggregate components, as the aforementioned examples suggest. Hence, one may consider forecasting all components and aggregation weights with univariate models and then aggregating these forecasts. A range of other possibilities may be useful and in Section 2 some of them will be discussed.

The main objective of this study is to check whether taking into account the information in stochastically varying aggregation weights is potentially beneficial for forecasting. Therefore we will focus on a small number of plausible predictors and compare their forecasting efficiency on a limited set of example series. We find that taking the information in the aggregation weights explicitly into account may indeed help improving the forecasts in a MSE sense. It is not the purpose of this study to suggest a universally optimal predictor but rather to point out that there is a source of information which may be worthwhile to consider. We are fully aware that in practice for each specific forecasting problem the most suitable predictors may be different.

The structure of the study is as follows. In Section 2 some possible predictors for contemporaneous aggregates with stochastic weights are presented and discussed. In Section 3 a small set of real life examples is investigated and it is demonstrated that taking explicitly into account the information in the stochastic aggregation weights helps improving forecast efficiency. In Section 4 a small Monte Carlo study is presented which explores the potential for forecast efficiency gains in a controlled environment. Finally, Section 5 concludes.

The following notation is used throughout. E denotes the expectation operator and E_τ denotes the corresponding conditional operator which conditions on information up to period τ . The natural logarithm is denoted by \log and Δ is the differencing operator. We use the following abbreviations: AR for autoregressive, VAR for vector autoregressive, iid for independently, identically distributed, DGP for data generation process, MSE and RMSE for mean squared error and root mean squared error, respectively, GDP for

gross domestic product, PPP for purchasing power parity, US for United States of America, NAFTA for North American Free Trade Agreement.

2 Possible Predictors

Suppose $y_t = (y_{1t}, \dots, y_{Kt})'$ is the vector of disaggregate component series and the aggregate of interest is $a_t = w_t' y_t$, where $w_t = (w_{1t}, \dots, w_{Kt})'$ is a vector of stochastic (time-varying) weights. Furthermore, suppose that y_t and w_t are generated by stochastic processes or possibly by a joint stochastic process. In the empirical section it will be assumed that all DGPs are AR or VAR processes which can at least be approximated well by finite order versions. For discussing the predictors to be used later, such an assumption is not required, however. The following h -step predictors at origin τ will be considered:

Univariate forecast Direct forecast of the univariate process a_t :

$$a_{\tau+h|\tau} = E(a_{\tau+h} | a_\tau, a_{\tau-1}, \dots) = E_\tau(a_{\tau+h}).$$

This predictor serves as a benchmark. It does not use any disaggregate information. If such information is useful then forecasts based on it should improve on this predictor.

Multivariate linear forecast Linear forecast taking into account disaggregate information:

$$a_{\tau+h|\tau}^o = E(a_{\tau+h} | a_\tau, a_{\tau-1}, \dots, y_\tau, y_{\tau-1}, \dots),$$

that is, a multivariate model is fitted to $(a_t, y_t)'$ and used for forecasting. The first component of the vector forecast is $a_{\tau+h|\tau}^o$. As in Lütkepohl (2011), the forecast may be based on selected components of y_t only rather than the full disaggregate vector.

Aggregation of multivariate forecasts Forecast based on multivariate predictions of disaggregate components and weights:

$$a_{\tau+h|\tau}^{mult} = E(w_{\tau+h} | w_\tau, w_{\tau-1}, \dots)' E(y_{\tau+h} | y_\tau, y_{\tau-1}, \dots).$$

Aggregation of univariate forecasts Forecast based on univariate predictions for disaggregate components and weights:

$$a_{\tau+h|\tau}^{uni} = [E_\tau(w_{1,\tau+h}), \dots, E_\tau(w_{K,\tau+h})] \begin{bmatrix} E_\tau(y_{1,\tau+h}) \\ \vdots \\ E_\tau(y_{K,\tau+h}) \end{bmatrix},$$

where $E_\tau(w_{k,\tau+h}) = E(w_{k,\tau+h}|w_{k,\tau}, w_{k,\tau-1}, \dots)$ etc.

The last predictor is included because it may not be possible to construct multivariate forecasts of y_t and w_t if there are many disaggregate components. Of course, other predictors are conceivable. For example, one may use a multivariate forecast of y_t and still predict the components of w_t with univariate models or vice versa. Also, it is possible that the disaggregate components and aggregation weights are related. In that case modelling and forecasting the joint process $(y'_t, w'_t)'$ may be plausible and then computing the aggregate forecast on that basis. Having quickly a very high-dimensional prediction problem is the obvious disadvantage. As mentioned earlier, it is not the objective of this study to find a universally optimal predictor for the case of aggregates with stochastic weights as we believe that the most suitable predictor will depend on the problem at hand. The small selection of predictors described in the foregoing is enough for making our main points. Hence, we limit attention to them.

3 Empirical Examples

Two examples based on real economic data are considered. In the first one forecasts for real GDP growth in the NAFTA are studied, the second one is based on European money stock variables. In both examples only three component series are aggregated. With such a small number of disaggregate components multivariate methods based on VARs are still feasible and may in fact have an advantage over univariate methods. This is the reason why we have chosen these examples although we know that there are many examples in practice where one has many more components.

3.1 NAFTA real GDP Growth

Quarterly data on real GDP for the three NAFTA countries US, Canada and Mexico measured at price levels and PPPs of 2005 are considered. Details on the data sources are given in Appendix A.1. The aggregate series is computed with weights computed based on real GDP shares. In other words, the aggregate NAFTA real GDP growth rate is computed by aggregating growth rates using these weights. More precisely,

$$\Delta \log q_t^{NAFTA} = \sum_i \frac{q_{t-1}^{(i)}}{q_{t-1}^{NAFTA}} \Delta \log q_t^{(i)}, \quad (3.1)$$

where $q_t^{(i)}$ denotes output in country i with $i = \text{US, Canada, Mexico}$. Notice that the weights are based on the output share in the previous period, as in Beyer, Doornik and Hendry (2001), so that for one-step ahead forecasts the weights are actually known. Data is available for the period 1970Q1-2010Q4 although NAFTA started only in 1994. In the following only data from 1985Q1 is used to avoid problems related to structural breaks. The three disaggregate and the aggregate series are plotted in Figure 1 and the aggregation weights are depicted in Figure 2. Apparently the weights vary substantially and are quite persistent. In fact, augmented Dickey-Fuller tests (results not shown) suggest that the Canadian series has a unit root.

We conduct a recursive pseudo out-of-sample forecasting experiment for growth rates of real GDP. Estimation and model selection are repeated for every sample considered. We use data from 1985Q1 onwards and the actual starts of the estimation periods are adjusted according to the presample values needed. We fit AR and VAR processes only and choose the lag order by model selection criteria AIC and SC, the first one being more generous and the second one more restrictive if they differ. The maximum lag length is four in all cases. Potential breaks and outliers in the time series or weights have not been modelled. The end of the initial estimation period varies because we wanted to check the robustness of the results with respect to the forecast period. To check how the recent recession affects the outcome one set of results is reported using data only until 2007Q4 and another one with data until 2010Q4. In both cases there are evaluation periods of different length to investigate the sensitivity of the results with respect to variations in the forecast period. Forecasting horizons are $h = 1$ and $h = 4$. RMSEs relative to the univariate AR forecasts for different evaluation periods are presented in Table 1.

The results in Table 1 show that the predictors which utilize forecasts of the aggregation weights, $a_{\tau+h|\tau}^{mult}$ and $a_{\tau+h|\tau}^{uni}$, are often superior to those which do not forecast the weights. Of course, for $h = 1$ the weights are known at the time of the forecast because the lagged shares are used (see (3.1)). Therefore it is important to note that efficiency gains are also obtained for forecast horizons $h = 4$. When the lag order is selected by AIC and only data until 2007Q4 are used, that is, the recent crisis period is excluded, $a_{\tau+h|\tau}^{uni}$ provides the smallest RMSEs for three out of four evaluation periods. The superior performance of the predictor which aggregates univariate forecasts of the disaggregate components and the weights may reflect the small sample size used for some of the forecasts. For example, when a long evaluation period starting in 1995Q1 is used, the associated estimation and specification period from 1985Q1-1994Q4 is rather small and leaves only a sample size of $T = 40$

Table 1: RMSEs Relative to Univariate Forecasts for NAFTA GDP Growth
(Total Sample Period: 1985Q1 - 2010Q4)

forecast	$h = 1$		$h = 4$		$h = 1$		$h = 4$	
	AIC	SC	AIC	SC	AIC	SC	AIC	SC
	ev. period: 1995Q1-2007Q4				ev. period: 1995Q1-2010Q4			
$a_{\tau+h \tau}^o$	1.0360	1.0155	0.9553	1.0076	1.0783	1.0968	0.9730	0.9642
$a_{\tau+h \tau}^{mult}$	1.0497	0.9854	0.9591	1.0068	1.0795	1.0167	0.9750	0.9563
$a_{\tau+h \tau}^{uni}$	0.9577	0.9926	0.9247	0.9976	1.0054	1.0076	0.9832	0.9726
	ev. period: 1998Q1-2007Q4				ev. period: 1998Q1-2010Q4			
$a_{\tau+h \tau}^o$	0.9812	1.0178	0.9669	1.0145	1.0613	1.1137	0.9784	0.9626
$a_{\tau+h \tau}^{mult}$	0.9652	0.9750	0.9682	1.0145	1.0456	1.0183	0.9795	0.9544
$a_{\tau+h \tau}^{uni}$	0.9387	0.9843	0.9534	0.9997	1.0064	1.0068	0.9978	0.9711
	ev. period: 2000Q1-2007Q4				ev. period: 2000Q1-2010Q4			
$a_{\tau+h \tau}^o$	1.0084	0.9788	0.9828	1.0233	1.0817	1.1072	0.9838	0.9579
$a_{\tau+h \tau}^{mult}$	0.9812	0.9583	0.9886	1.0203	1.0613	1.0154	0.9858	0.9480
$a_{\tau+h \tau}^{uni}$	0.9376	0.9797	0.9746	1.0057	1.0138	1.0069	1.0088	0.9688
	ev. period: 2003Q1-2007Q4				ev. period: 2003Q1-2010Q4			
$a_{\tau+h \tau}^o$	0.9741	0.9306	0.9725	0.9856	1.1037	1.1466	0.9831	0.9440
$a_{\tau+h \tau}^{mult}$	0.9916	0.9890	0.9793	1.0000	1.0945	1.0430	0.9845	0.9341
$a_{\tau+h \tau}^{uni}$	1.0067	1.0081	0.9787	0.9976	1.0573	1.0230	1.0156	0.9621

Table 2: RMSEs Relative to Univariate Forecasts for European Real M3 Growth (Total Sample Period: 1981Q2 - 2010Q3)

forecast	$h = 1$		$h = 4$		$h = 1$		$h = 4$	
	AIC	SC	AIC	SC	AIC	SC	AIC	SC
	ev. period: 1992Q1-2007Q4				ev. period: 1992Q1-2010Q3			
$a_{\tau+h \tau}^o$	1.0012	0.9861	1.0137	1.0095	0.9859	1.0439	1.0026	1.0061
$a_{\tau+h \tau}^{mult}$	0.9978	0.9804	1.0097	1.0031	0.9877	1.0390	0.9908	1.0015
$a_{\tau+h \tau}^{uni}$	0.9313	0.9267	0.9821	0.9973	0.9412	0.9486	0.9757	0.9957
	ev. period: 1996Q1-2007Q4				ev. period: 1996Q1-2010Q3			
$a_{\tau+h \tau}^o$	0.9946	0.9978	1.0226	1.0211	0.9756	1.0749	1.0058	1.0140
$a_{\tau+h \tau}^{mult}$	1.0377	0.9894	1.0077	1.0115	1.0135	1.0684	0.9832	1.0075
$a_{\tau+h \tau}^{uni}$	0.9388	0.9666	0.9744	1.0013	0.9500	0.9862	0.9678	0.9983
	ev. period: 2000Q1-2007Q4				ev. period: 2000Q1-2010Q3			
$a_{\tau+h \tau}^o$	0.9897	0.9973	1.0154	1.0230	0.9693	1.0875	0.9980	1.0142
$a_{\tau+h \tau}^{mult}$	0.9962	0.9905	0.9923	1.0133	0.9810	1.0824	0.9682	1.0082
$a_{\tau+h \tau}^{uni}$	0.9176	0.9741	0.9629	0.9997	0.9374	0.9947	0.9585	0.9966
	ev. period: 2003Q1-2007Q4				ev. period: 2003Q1-2010Q3			
$a_{\tau+h \tau}^o$	0.9274	0.9704	0.9479	1.0228	0.9260	1.1077	0.9496	1.0107
$a_{\tau+h \tau}^{mult}$	0.9818	0.9614	0.9721	1.0134	0.9669	1.1021	0.9460	1.0062
$a_{\tau+h \tau}^{uni}$	0.8258	0.9023	0.9531	0.9939	0.8957	0.9637	0.9508	0.9917

when forecasts for 1995 are determined. For fitting three-dimensional VARs for such a small sample period may well lead to large estimation uncertainty and reduced forecast precision relative to a predictor which is based exclusively on univariate forecasts. It has to be noted, however, that the situation is slightly different for evaluation periods up to 2010Q4. In that case, the aggregation of multivariate forecasts leads to slightly smaller RMSEs. Actually, for $h = 4$ and lag order selection by SC, $a_{\tau+h|\tau}^{mult}$ results in the smallest RMSEs for all four evaluation periods. In any case, the smallest RMSEs in most cases are obtained for predictors that explicitly utilize information in the aggregation weights.

3.2 European M3 Growth

The second example is based on quarterly real money stock M3 series from the three European countries Germany, France and Italy. Details on the data sources are given in Appendix A.2. The weights are computed based on real M3 shares and the aggregate real M3 growth rate is computed by

aggregating growth rates using these weights. Data is available from 1981Q2-2010Q3, that is, our sample period starts well before the introduction of the euro. The series and the weights are plotted in Figures 3 and 4. Again, the aggregation weights vary substantially and show considerable persistence. In fact, in this case unit roots are not rejected in either of the series.

Forecasting is done as in the previous example, that is, recursive pseudo out-of-sample forecasting for growth rates of real M3 is carried out. Estimation and model selection is repeated for every sample considered. Different initial estimation and evaluation periods are used. We use data from 1981Q2 onwards and again the beginnings of the estimation periods depend on the number of presample values needed. We compare results for samples until 2007Q4 and also consider samples with data until 2010Q3. Forecasting horizons are again $h = 1$ and 4. Relative RMSEs are reported with the univariate forecasts as benchmark. We fit only AR and VAR models without accounting for potential breaks or outliers. The model orders are chosen by AIC and SC using a maximum order of four.

The results are similar to those in the previous example in that the predictors based on forecasts for the aggregation weights have smaller RMSEs than the other two predictors. Note that again the one-step ahead forecasts $a_{\tau+h|\tau}^{mult}$ and $a_{\tau+h|\tau}^{uni}$ use known weights whereas the 4-step ahead forecasts use predicted weights. If AIC is used for lag order selection, $a_{\tau+h|\tau}^{uni}$ results in the smallest RMSEs for forecast horizon $h = 4$ in six out of eight evaluation periods considered in Table 2, the exceptions being the evaluation periods 2003Q1-2007Q4 and 2003Q1-2010Q3 where $a_{\tau+h|\tau}^{uni}$ has a marginally larger RMSE than $a_{\tau+h|\tau}^o$ and/or $a_{\tau+h|\tau}^{mult}$. A similar result is obtained with SC where $a_{\tau+h|\tau}^{uni}$ is best in seven out of eight evaluation periods and in this case the exception is the evaluation period 1996Q1-2007Q4 where the direct univariate predictor has about the same RMSE as $a_{\tau+h|\tau}^{uni}$. The multivariate predictor based on disaggregate information in the aggregation weights, $a_{\tau+h|\tau}^{mult}$, is always worse than $a_{\tau+h|\tau}^{uni}$ when SC is used. It may again suffer from the small sample sizes used for the longest evaluation periods. It is also interesting to note that there are never substantial losses due to using the information in the aggregation weights.

Of course, forecast efficiency gains for a specific dataset and sample period may be obtained by chance. One way to explore whether the gains are real is to conduct a Monte Carlo experiment based on DGPs similar to the models used in the empirical studies. This is what we will do next.

4 Monte Carlo Experiment

In the following we first describe our Monte Carlo setup and then discuss the results.

4.1 Monte Carlo Setup

We have used a number of DGPs with parameter values estimated from our example data. In each case, the disaggregate component series are generated by a VAR(p) process,

$$y_t = \nu + A_1 y_{t-1} + \cdots + A_p y_{t-p} + u_t, \quad (4.1)$$

where A_i are (3×3) coefficient matrices and $u_t \sim \mathcal{N}(0, \Sigma_u)$ is an iid Gaussian white noise process with covariance matrix Σ_u . The aggregation weights w_t are generated by a finite order VAR of dimension two because the aggregation weights add up to one. Although we have experimented with a number of other processes as well, we focus on the following two processes to discuss results.

DGP1: The parameters of a VAR(1) for the disaggregate series are estimated from the NAFTA country data for 1985Q1-2007Q4, that is, y_t is meant to mimic $100 \times (\Delta \log q_t^{\text{US}}, \Delta \log q_t^{\text{CAN}}, \Delta \log q_t^{\text{MEX}})'$ with

$$y_t = \begin{pmatrix} 0.540 \\ 0.159 \\ 0.360 \end{pmatrix} + \begin{pmatrix} 0.025 & 0.255 & 0.022 \\ 0.399 & 0.324 & 0.004 \\ 0.238 & -0.126 & 0.313 \end{pmatrix} y_{t-1} + u_t$$

and associated white noise covariance

$$\Sigma_u = \begin{pmatrix} 0.227 & 0.102 & 0.132 \\ 0.102 & 0.280 & 0.091 \\ 0.132 & 0.091 & 1.585 \end{pmatrix}.$$

The moduli of the roots of the characteristic polynomial of the VAR coefficient matrix are 0.527, 0.323 and 0.187. Hence, the process is stable and the persistence is moderate or small, as one would expect for GDP growth rates.

Given that there are three disaggregate components and the aggregation weights add to one, the DGP of the weights is bivariate. It is obtained by fitting a VAR(1) to US and Canadian data, $w_t = (w_t^{\text{US}}, w_t^{\text{CAN}})'$ giving

$$w_t = \begin{pmatrix} 0.082 \\ 0.008 \end{pmatrix} + \begin{pmatrix} 0.903 & -0.011 \\ -0.006 & 0.966 \end{pmatrix} w_{t-1} + u_t^w$$

and white noise covariance matrix

$$\Sigma_u^w = \begin{pmatrix} 1.072 & -0.109 \\ -0.109 & 0.167 \end{pmatrix} \times 10^{-6}.$$

The moduli of the roots of the characteristic polynomial of A_1 in this case are 0.967 and 0.902. Hence, the DGP of the weights is quite persistent, in line with the visual impression of persistent weights in Figure 2. \square

DGP2: The second DGP is a VAR(2) for the disaggregate components estimated using data from 1981Q2-2007Q4 for the three M3 growth rates, that is, $y_t = 100 \times (\Delta \log m3r_t^{\text{GER}}, \Delta \log m3r_t^{\text{FRA}}, \Delta \log m3r_t^{\text{IT}})'$, where

$$\begin{aligned} y_t &= \begin{pmatrix} 0.949 \\ 0.421 \\ 0.462 \end{pmatrix} + \begin{pmatrix} 0.095 & 0.150 & 0.124 \\ 0.030 & 0.364 & 0.026 \\ -0.022 & 0.055 & 0.244 \end{pmatrix} y_{t-1} \\ &+ \begin{pmatrix} -0.089 & -0.110 & -0.288 \\ -0.019 & 0.217 & -0.069 \\ -0.071 & 0.048 & 0.062 \end{pmatrix} y_{t-2} + u_t \end{aligned}$$

and

$$\Sigma_u = \begin{pmatrix} 1.526 & -0.215 & -0.036 \\ -0.215 & 1.146 & 0.090 \\ -0.036 & 0.090 & 1.711 \end{pmatrix}.$$

The moduli of the roots of the characteristic polynomial of the VAR operator are 0.655, 0.532, 0.418, 0.418, 0.322 and 0.322 which again means that the persistence is moderate only.

The associated bivariate process for the aggregation weights is also a VAR(2) with parameters estimated from the German and French series, that is, $w_t = (w_t^{\text{GER}}, w_t^{\text{FRA}})'$ and

$$w_t = \begin{pmatrix} 0.050 \\ -0.008 \end{pmatrix} + \begin{pmatrix} 1.043 & -0.056 \\ 0.098 & 1.301 \end{pmatrix} w_{t-1} + \begin{pmatrix} -0.087 & -0.039 \\ -0.076 & -0.304 \end{pmatrix} w_{t-1} + u_t^w$$

with

$$\Sigma_u^w = \begin{pmatrix} 1.458 & -0.859 \\ -0.859 & 1.079 \end{pmatrix} \times 10^{-5}.$$

The moduli of the roots of the characteristic VAR polynomial are 0.970, 0.970, 0.331 and 0.075 which again implies a high persistence in this aggregation weights, although the roots are still a bit away from unity. \square

Table 3: RMSEs Relative to Univariate Forecasts for DGP1

T	forecast	$h = 1$		$h = 2$		$h = 4$	
		AIC	SC	AIC	SC	AIC	SC
50	$a_{\tau+h \tau}^o$	1.0108	1.0095	1.0131	1.0026	1.0098	1.0004
	$a_{\tau+h \tau}^{mult}$	0.9824	0.9955	0.9979	0.9999	1.0020	0.9987
	$a_{\tau+h \tau}^{uni}$	0.9877	0.9871	0.9926	0.9987	0.9989	0.9992
100	$a_{\tau+h \tau}^o$	0.9741	1.0107	0.9925	1.0052	0.9951	0.9994
	$a_{\tau+h \tau}^{mult}$	0.9684	0.9720	0.9898	0.9939	0.9952	0.9983
	$a_{\tau+h \tau}^{uni}$	0.9992	0.9979	0.9990	1.0012	0.9982	0.9999
250	$a_{\tau+h \tau}^o$	0.9739	0.9716	0.9963	0.9940	0.9981	0.9987
	$a_{\tau+h \tau}^{mult}$	0.9715	0.9689	0.9944	0.9928	0.9977	0.9986
	$a_{\tau+h \tau}^{uni}$	0.9953	0.9971	0.9984	0.9991	0.9987	0.9999

We have also considered a number of alternative artificial DGPs obtained by fitting other VARs to the data or based on selected other processes. Some results obtained for those processes will be mentioned in the following. We have chosen DGP1 and DGP2 for a more detailed discussion of the results because they are similar to those for other DGPs.

Samples of sizes $T = 50, 100$ and 250 were generated plus 50 presample observations which were discarded to reduce the impact of starting-up values. Moreover, four values were generated at the end of each sample for the forecast comparison. In the following, T denotes the gross sample size used for estimation and model specification but excluding the values for forecast evaluation. Full AR and VAR models with a constant term are fitted using lag orders selected by AIC and SC with a maximum order four, as in the empirical studies. The number of replications is 5000.

4.2 Monte Carlo Results

RMSEs relative to the direct univariate forecasts for forecast horizons $h = 1, 2$ and 4 are presented in Tables 3 and 4 for DGP1 and DGP2, respectively. As in the empirical studies, the weights are known for 1-step ahead forecasts while they are forecasted when $a_{\tau+h|\tau}^{mult}$ and $a_{\tau+h|\tau}^{uni}$ are used and $h > 1$. The results in Table 3 show that utilizing the information in the aggregation weights explicitly is beneficial for the forecast efficiency, but the gains are very moderate when the weights are actually forecasted ($h = 2, 4$). Also there is very little difference between the RMSEs of $a_{\tau+h|\tau}^{mult}$ and $a_{\tau+h|\tau}^{uni}$ for DGP1.

Table 4: RMSEs Relative to Univariate Forecasts for DGP2

T	forecast	$h = 1$		$h = 2$		$h = 4$	
		AIC	SC	AIC	SC	AIC	SC
50	$a_{\tau+h \tau}^o$	1.0284	1.0171	0.9977	0.9852	1.0281	0.9986
	$a_{\tau+h \tau}^{mult}$	0.9911	1.0144	0.9843	0.9837	1.0096	0.9976
	$a_{\tau+h \tau}^{uni}$	0.9744	0.9680	0.9711	0.9719	1.0016	0.9998
100	$a_{\tau+h \tau}^o$	0.9752	1.0322	0.9783	0.9885	0.9978	1.0001
	$a_{\tau+h \tau}^{mult}$	0.9614	1.0103	0.9678	0.9820	0.9967	0.9979
	$a_{\tau+h \tau}^{uni}$	0.9735	0.9665	0.9745	0.9732	0.9962	0.9977
250	$a_{\tau+h \tau}^o$	0.9546	0.9969	0.9622	0.9845	0.9956	0.9970
	$a_{\tau+h \tau}^{mult}$	0.9476	0.9611	0.9586	0.9646	0.9947	0.9955
	$a_{\tau+h \tau}^{uni}$	0.9772	0.9735	0.9799	0.9752	0.9953	0.9939

Moreover, it does not make much difference whether the order is chosen by AIC or SC. Clearly, for DGP1 the choice between the four predictors does not make much difference for the forecast efficiency for $h = 2$ or 4. This is true for all sample sizes considered. The fact that it holds also for larger samples of size $T = 250$ indicates that the structure of the DGP does not leave much room for improvements even when specification and estimation problems are less important. A similar result was also obtained for a number of other artificial DGPs.

The situation is a bit different for DGP2. In Table 4 it can be seen that forecasting the aggregation weights can make a difference and, in fact, improve forecast RMSEs at least for small forecast horizons. In other words, there are sizable gains for $h = 2$ but not for $h = 4$. Of course, one would not expect substantive efficiency gains for larger forecast horizons given that the persistence in the disaggregate components is low. This is nicely reflected in the results for $h = 4$ which are again very similar for all four predictors.

We have also generated an aggregate by using a VAR(1) for the disaggregate components and weights generated by a random walk to explore the impact of persistence in the aggregation weights. The results were more like those in Table 3. In other words, forecasting the aggregation weights improves the forecast efficiency but at a very small margin. Thus, for practical purposes there is very little to recommend one of the predictors over the others.

Generally we found a number of DGPs where the difference between the four predictors in terms of RMSE were small. However, we never found cases

where taking into account the information in the aggregation weights was harmful. In other words, for none of our DGPs the forecasts $a_{\tau+h|\tau}^{mult}$ and $a_{\tau+h|\tau}^{uni}$ were substantially inferior to the other two predictors. Thus, the risk of loosing efficiency by using predictors for the aggregation weights is small while there is a chance for MSE improvements.

5 Conclusions

In this study we have considered forecasting contemporaneous aggregates with stochastic aggregation weights. We have pointed out that such aggregates are quite common in practice and that taking into account the information in the weights may lead to better forecasts in a MSE sense. We have compared four predictors for such variables: (1) a standard direct univariate AR forecast which is based only on the past of the aggregate series, (2) a multivariate linear VAR forecast of the aggregate which takes into account information from the disaggregate components, (3) a forecast which aggregates a multivariate forecast of the disaggregate components and the aggregation weights and (4) a forecast which is based on aggregating univariate AR forecasts for the individual disaggregate components and the aggregation weights. In two empirical examples we have shown that the last two forecasts may lead to lower forecast MSEs than the first two forecasts. In other words, using the information in the stochastic aggregation weights explicitly may indeed improve forecast efficiency. In a Monte Carlo study we have confirmed that such efficiency gains are not just spurious but are a consequence of the stochastic structure of the DGP, although the efficiency gains are not large.

There are a number of related problems which we have not addressed in this study but which may be of interest for future work. First, we have investigated the potential for forecast efficiency gains by using the information in the aggregation weights only for a very small set of empirical examples. A larger scale investigation may shed light on the general potential for gains in forecast precision and perhaps for which aggregates they can be expected. Second, we have considered a rather limited number of possible predictors in our comparison. While they are sufficient to demonstrate that there is scope for improving efficiency by using information in the stochastic aggregation weights, there are a number of other predictors that seem to be natural competitors and may further improve the exploitation of the information in the aggregation weights. For example, one may consider modelling the joint DGP of the disaggregate components and the aggregation weights or one may combine univariate forecasts for the weights with multivariate forecasts for the disaggregate components or vice versa. Another strand of related

research may consider the precise stochastic structure of the aggregate for given DGPs of the weights and disaggregate components. In general this is not likely to be an easy problem because the aggregate is a product of two multivariate processes. A very limited set properties under rather special condition for such processes are provided in Appendix B of Lütkepohl (2011). More general results may well be helpful in assessing the potential for forecast improvements in the aggregation weights. These issues are left for future research.

A Data Sources

A.1 NAFTA GDP Data

The real GDP series denoted as $q_t^{(i)}$ are taken from Thomson Datastream and correspond to seasonally adjusted gross domestic product measured at constant 2005 PPPs in millions of US Dollar as reported by the OECD. The Datastream mnemonics for the US, Canada and Mexico are USOCFGVOD, CNOCFGVOD, MXOCFGVOD, respectively. Growth rates and weights are computed as described in Section 3.1.

A.2 European M3 Data

Germany: Seasonally adjusted monthly values of nominal money supply M3 (in billions of EUR) as reported by the Deutsche Bundesbank are taken from Thomson Datastream (Mnemonic: BDM3....B). Quarterly values correspond to observations of the last month in the respective quarters. Real M3 is obtained by using the GDP deflator with base year 2005 (Datastream Mnemonic: BDONA001E). German unification effects are accounted for by regressing the growth rate of real M3 on a constant, four lags and a unification dummy that takes on value 1 in 1990Q3 and 0 elsewhere. The estimated effect on the growth rates is 0.143 and thus the pre-unification figures are multiplied by 1.143.

France: Seasonally non-adjusted monthly values of nominal money supply M3 (in millions of EUR) as reported by the Banque de France are taken from Thomson Datastream (Mnemonic: FRM3....A). The data has been seasonally adjusted by the X12-ARIMA method and converted into billions of EUR. Quarterly values correspond to observations of the last month in the respective quarters. Real M3 is obtained by using the GDP deflator with base year 2005 (Datastream Mnemonic: FRONA001E).

Italy: Seasonally non-adjusted monthly values of nominal money supply M3 (in millions of EUR) as reported by the Banca d'Italia are taken from Thomson Datastream (Mnemonic: ITM3....A). The data has been seasonally adjusted by the X12-ARIMA method and converted into billions of EUR. Quarterly values correspond to observations of the last month in the respective quarters. Real M3 is obtained by using the GDP deflator with base year 2005, which is obtained by rebasing a price deflator that corresponds to the base year 2000 (Datastream Mnemonic: ITESGDDFE).

References

- Ansley, C. F., Spivey, W. A. and Wroblewski, W. J. (1977). On the structure of moving average processes, *Journal of Econometrics* **6**: 121–134.
- Beyer, A., Doornik, J. A. and Hendry, D. F. (2001). Reconstructing historical euro-zone data, *Economic Journal* **111**: 308–327.
- Carson, R. T., Cenesizoglu, T. and Parker, R. (2010). Forecasting (aggregate) demand for US commercial air travel, *International Journal of Forecasting* **forthcoming**.
- Espasa, A., Senra, E. and Albacete, R. (2002). Forecasting inflation in the European Monetary Union: A disaggregated approach by country and by sectors, *European Journal of Finance* **8**: 402–421.
- Hendry, D. F. and Hubrich, K. (2011). Combining disaggregate forecasts or combining disaggregate information to forecast an aggregate, *Journal of Business & Economic Statistics* **29**: 216–227.
- Hubrich, K. (2005). Forecasting euro area inflation: Does aggregating forecasts by HICP component improve forecast accuracy?, *International Journal of Forecasting* **21**: 119–136.
- Kohn, R. (1982). When is an aggregate of a time series efficiently forecast by its past?, *Journal of Econometrics* **18**: 337–349.
- Lütkepohl, H. (1984a). Forecasting contemporaneously aggregated vector ARMA processes, *Journal of Business & Economic Statistics* **2**: 201–214.
- Lütkepohl, H. (1984b). Linear transformations of vector ARMA processes, *Journal of Econometrics* **26**: 283–293.
- Lütkepohl, H. (1986). Comparison of predictors for temporally and contemporaneously aggregated time series, *International Journal of Forecasting* **2**: 461–475.

- Lütkepohl, H. (1987). *Forecasting Aggregated Vector ARMA Processes*, Springer-Verlag, Berlin.
- Lütkepohl, H. (2010). Forecasting aggregated time series variables - a survey, *Journal of Business Cycle Measurement and Analysis* **2010/2**: 37–62.
- Lütkepohl, H. (2011). Forecasting nonlinear aggregates and aggregates with time-varying weights, *Jahrbücher für Nationalökonomie und Statistik* **213**(1): 107–133.
- Marcellino, M., Stock, J. H. and Watson, M. W. (2003). Macroeconomic forecasting in the Euro area: Country specific versus area-wide information, *European Economic Review* **47**: 1–18.
- Tiao, G. C. and Guttman, I. (1980). Forecasting contemporal aggregates of multiple time series, *Journal of Econometrics* **12**: 219–230.
- Wei, W. W. S. and Abraham, B. (1981). Forecasting contemporal time series aggregates, *Communications in Statistics - Theory and Methods* **A10**: 1335–1344.

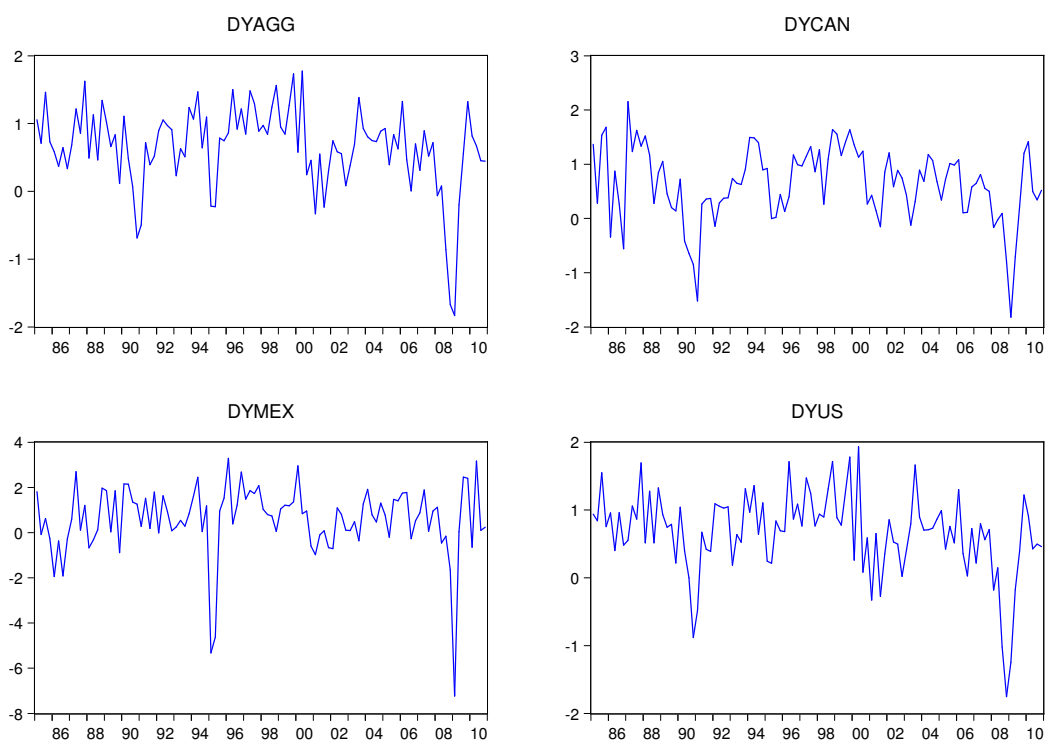


Figure 1: GDP growth rates of NAFTA and NAFTA countries.

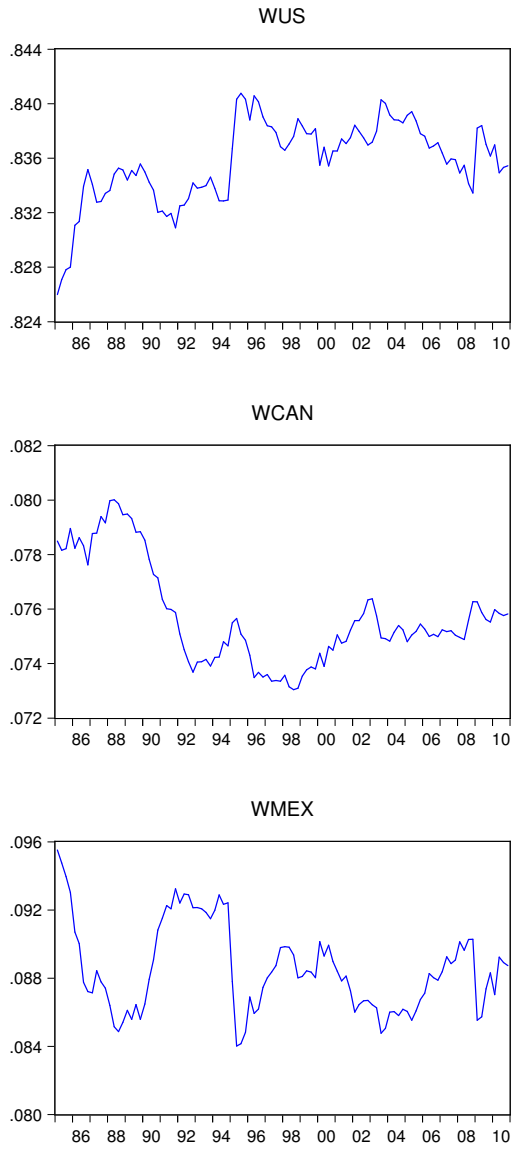


Figure 2: Weights for NAFTA GDP growth rates.

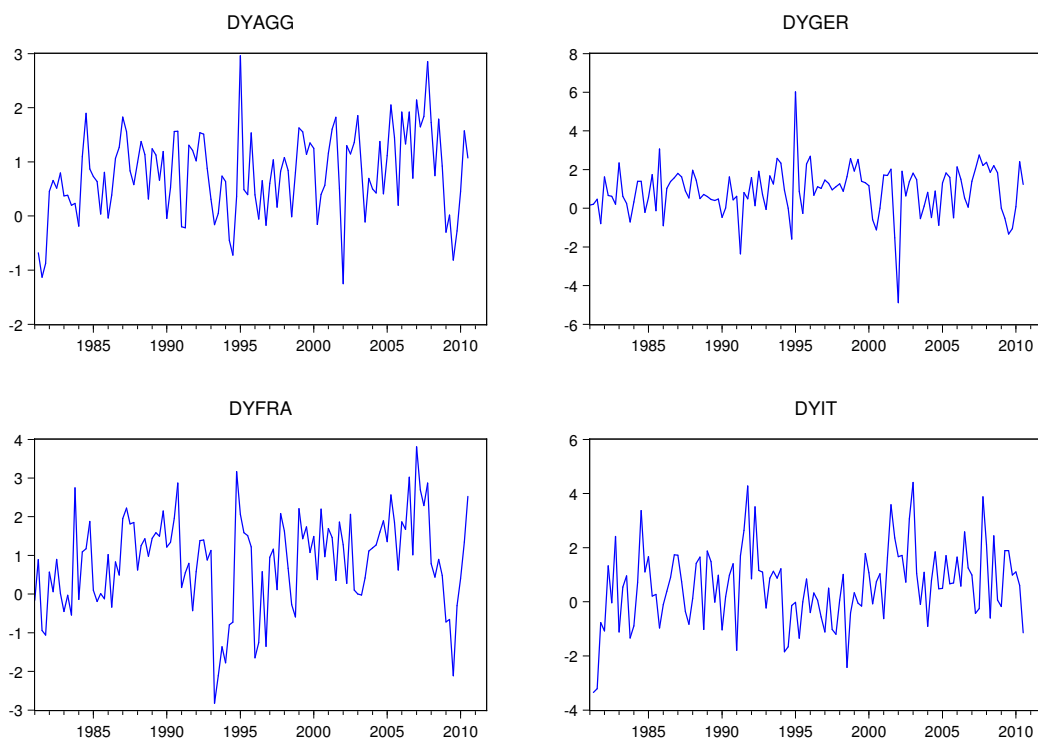


Figure 3: Aggregate and individual real M3 growth rates for Germany, France and Italy.

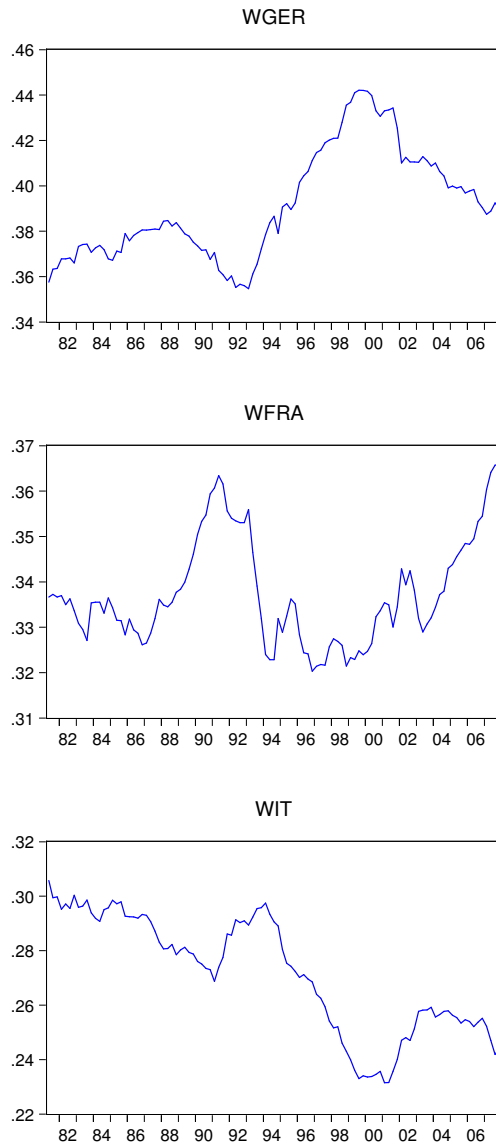


Figure 4: Weights for aggregate real M3 growth rates.