

# Forecasting Euro-Area Variables with German Pre-EMU Data\*

Ralf Brüggemann

Helmut Lütkepohl

Massimiliano Marcellino

June 28, 2006

## Abstract

It is investigated whether Euro-area variables can be forecast better based on synthetic time series for the pre-Euro period or by using just data from Germany for the pre-Euro period. Our forecast comparison is based on quarterly data for the period 1970Q1 - 2003Q4 for ten macroeconomic variables. The years 2000 - 2003 are used as forecasting period. A range of different univariate forecasting methods is applied. Some of them are based on linear autoregressive models and we also use some nonlinear or time-varying coefficient models. It turns out that most variables associated with the monetary sector of the economy can be better predicted based on German data while aggregated European data are preferable for forecasting variables from the real economy. These results suggest that it may be reasonable to consider the monetary German pre-EMU data also for studying economic problems for the Euro-area.

*Keywords:* Aggregation, forecasting, European monetary union, constructing EMU data

*JEL classification:* C22, C53

---

\*This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 'Economic Risk'. Ralf Brüggemann: Humboldt-Universität zu Berlin, Department of Economics, Spandauer Str. 1, D-10178 Berlin, Germany, email: brueggem@wiwi.hu-berlin.de; Helmut Lütkepohl: Department of Economics, European University Institute, Via della Piazzola 43, I-50133 Firenze, Italy, email: helmut.luetkepohl@iue.it. Massimiliano Marcellino: IGIER, Bocconi University, Via Salasco 3/5, I-20136 Milan, Italy, email: massimiliano.marcellino@uni-bocconi.it

# 1 Introduction

The creation of the European Monetary Union (EMU) makes it necessary to investigate some economic issues for the Euro-area as a whole rather than for individual member countries. For this purpose synthetically constructed, aggregated data for the pre-EMU period are often used. That approach has a number of drawbacks, however. Marcellino (2004) lists the following: (1) Some data may not be available at the desired frequency for some of the member states and, hence, have to be interpolated. (2) If seasonally adjusted data are used, various different working day adjustments in the seasonal adjustment procedures are necessary for different countries and institutional changes may have to be accounted for. (3) The method of aggregation will have an important impact on the series available for analysis. Different aggregation methods are discussed by Winder (1997) and Beyer, Doornik & Hendry (2001) whereas Bosker (2006) shows that the aggregation method can have a substantial impact on the parameter estimates of a Euro-area money demand function. One may also add that substantial adjustment processes were required in some countries prior to the EMU to satisfy the Maastricht criteria. These adjustments may have induced structural changes in the generation processes of some variables. To avoid such problems Brüggemann & Lütkepohl (2006) use a different approach. They combine German data until 1998 with EMU data from 1999 onwards. They argue that using German data makes sense because Germany roughly satisfied the Maastricht criteria already at the time when the conditions for entering the EMU were announced. Thus, no substantial adjustment processes were necessary in Germany. Therefore the question arises if using past German data rather than artificially aggregated EMU data may be beneficial in forecasting EMU aggregates. In other words, it is of interest to determine whether German data prior to the EMU period contain as much or even more information on the EMU period than aggregated Euro-area data.

In this context it may be worth recalling some implications of contemporaneous linear aggregation for time series. If the series are generated by autoregressive integrated moving average (ARIMA) processes then the aggregate is also an ARIMA process. Its orders may be higher, however, than for the individual series and its structure more complicated and more difficult to capture by simple linear forecasting models (see, e.g., Lütkepohl (2005) for a summary of results). In fact, if there is considerable heterogeneity in the DGPs of the variables to be aggregated, then forecasting the aggregate directly will usually be less efficient than aggregating disaggregate univariate or multivariate forecasts. Thus, aggregation can have a substantial impact on forecast accuracy. Of course, this does not even take into account possible structural change in some of the individual country series which would be translated into the aggregate and, hence, may make forecasting more difficult. If the aggregate series is characterized by structural change, fitting time series models which do not take that into account may lead to poor forecasts. If the corresponding German series is more similar to the post-EMU series, using German data may clearly have advantages.

Using EMU data from 1970 onwards, Marcellino (2004) finds that nonlinear models often beat linear forecasting models in terms of accuracy. This result may be a consequence of data problems that originate in the aggregation over data from countries where substantial adjustment processes were going on and which are captured by nonlinearities. In that case, using German data for the pre-EMU period may result in superior linear forecasts and this is what we will investigate in the present study. If linear models based on German pre-EMU data were superior to models based on

artificially aggregated EMU data, this would also be an indication that German data may be useful for modelling the Euro-area economy, as argued by Brüggemann & Lütkepohl (2006). On the other hand, a different view was expressed by Nautz & Offermanns (2006) who find that the Euro/Dollar exchange rate can be predicted better based on synthetic pre-EMU data than on German pre-EMU data.

We will try to shed light on the problem which data is suitable for economic analysis by investigating the forecasting performance of models based on different time series for a range of different Euro-area variables. In fact we use a subset of the variables from Marcellino (2004) and exclude those variables for which a direct German analog does not exist. Forecasts for the EMU period are based on different linear and nonlinear or time-varying coefficient univariate models. It turns out that for variables such as GDP which need major adjustments to obtain an EMU proxy from the German figures, using the synthetic EMU data is preferable. On the other hand, for variables like interest rates and price indices which have similar levels for Germany and the EMU countries, using German rather than aggregated EMU data for the pre-EMU period has advantages.

The structure of our study is as follows. In the next section the data are described and in Section 3 the forecasting methods are presented. The forecasting results are discussed in Section 4 and Section 5 concludes.

## 2 Data

For the forecasting comparison we use data on a subset of the macroeconomic variables that have been analyzed in Marcellino (2004). The variables include real GDP (YER) and its components, real personal and real government consumption (PCR and GCR), investment and inventories (ITR and SCR). Nominal and financial variables are the GDP deflator (YED), a consumer price index (CPI), short- and long-term interest rates (STN and LTN) and the real exchange rate (EER). For some of the variables used in Marcellino (2004), it is not obvious how they should be constructed in a meaningful way based on German data. Examples include variables that are related to trade like imports, exports and the trade balance. We do not include those variables in our study.

For each of the ten variables we have quarterly data for the period from 1970Q1 to 2003Q4. Apart from aggregated European data we have three variants that use German data for the pre-EMU period.

The aggregated European data is constructed by using official data for the Euro-area as provided by the OECD and the ECB, whenever it is available. Euro-area time series data for YER, PCR, GCR, ITR, SCR and YED are taken from the OECD statistical compendium for the whole sample period 1970Q1-2003Q4. For the other time series, official European series are not provided before the beginning of the 90s. In these cases, we have used time series from the database constructed by Fagan, Henry & Mestre (2001)<sup>1</sup> until official data is available (see Table 1 for data sources). Official interest rate time series for the Euro-area are, e.g., published by the ECB only from 1994 onwards. In this case, we take data for the period from 1970Q1-1993Q4 from the Fagan et al. (2001) database.

---

<sup>1</sup>The database is available from the ECB website.

Forecasts based on the aggregated Euro-area data will be compared to different variants of Euro-area data that uses German data for the period before the Euro was introduced. In particular, we use three variants of data that all use German data for the pre-EMU period:

- (1) West-German, German and European data have been joined but no further adjustments have been made afterwards. West-German data have been used until the end of 1990 (1990Q4) and from 1991Q1 to 1998Q4 data for unified Germany are considered. Finally, starting in 1999Q1 European data has been used. The data sources are given in Table 1. We refer to this data as GER0.
- (2) The second variant using German data is based on the series in (1) but in addition a purely statistical adjustment has been made which increases the German data to the EMU level at the beginning of 1999 and adjusts (multiplicatively) the whole series accordingly. An analogous adjustment has been made to join West-German data and data for unified Germany. This kind of approach has also been used by Fagan et al. (2001) for adjusting German data. In the following this data is referred to as GER1.
- (3) In the third variant, the series have been adjusted by estimating the shifts in the time series. To be more precise, for each time series from (1) the shifts are estimated by

$$y_t = \nu_0 + \nu_1 t + \nu_2 s_{91} q_{1t} + \nu_3 s_{99} q_{1t} + u_t, \quad (2.1)$$

where  $y_t$  is the time series,  $u_t$  is an error term and  $s_{91} q_{1t}$  and  $s_{99} q_{1t}$  are step dummies which have the value one from the first quarter of 1991 and 1999 onwards, respectively and are zero elsewhere. Data up to 1999Q4 are used for estimating the shift. Adjusted data for 1970Q1-1990Q4 is obtained by adding  $\hat{\nu}_2 + \hat{\nu}_3$  to the original series. For the period 1991Q1-1998Q4,  $\hat{\nu}_3$  is added to the original series. We refer to this data as GER2.

The two interest rate time series based on German data do not show any obvious breaks or shifts at the time of the Euro introduction. Therefore, we decided not to adjust the interest rate series. Thus the interest rate series in GER0, GER1 and GER2 are identical.

The time series included in the four data sets are depicted in Figure 1 for the real variables and in Figure 2 for the nominal and financial variables. We will sometimes refer to the latter set of variables as monetary variables in the following. Although the large shift in the real GER0 series at the beginning of 1999 to some extent distorts the overall comparability of the series, it can be seen that for YER, PCR, GCR and ITR the general trending behaviour of the European and GER1 series is similar whereas shorter term deviations are also obvious. Moreover, all three inventory series (SCR) based on German data are much more volatile than the corresponding European series. Here the smoothing effect of aggregation becomes apparent.

For the monetary series the variants based on German data are fairly similar, as can be seen in Figure 2. Of course, for the two interest rates (STN, LTN) there is just one such variant. Even for the price series YED and CPI the three series based on German data are closer to each other than to the corresponding European series. Clearly the aggregated European series increase more steeply than the German series, that is, the aggregate inflation in the Euro-area was generally higher than in Germany during the pre-Euro period, as is well known. Also the European and German interest

rates and exchange rates were quite different in the pre-Euro period. Thus, for these series it will be of particular interest to see whether the aggregate or the German series are better suited for forecasting the corresponding series in the Euro-era.

### 3 Forecasting methods and forecast evaluation

In this section we describe the forecasting methods used in our comparison and the criteria for evaluating their relative merits. Both, the methods and the criteria are similar to those adopted by Stock & Watson (1999) and Marcellino (2004), so that we only provide a brief sketch and refer to these papers for additional details.

#### 3.1 Forecasting methods

The objective is to forecast a variable  $y_t$ . The forecasts may be obtained from forecasts of the first differences  $\Delta y_t = y_t - y_{t-1}$  of  $y_t$ . To gain efficiency in describing the forecasting methods we specify auxiliary variables  $y_{t+h}^h$  such that  $y_{t+h}^h = y_{t+h}$  if  $y_t$  is treated as stationary and  $y_{t+h}^h = y_{t+h} - y_t$  if  $y_t$  is integrated. Here  $h$  indicates the forecast horizon.

Using this notation, all forecasting methods considered in the following can be expressed in terms of the general specification

$$y_{t+h}^h = f(Z_t; \theta_{ht}) + \varepsilon_{t+h}, \quad (3.1)$$

where  $Z_t$  is a vector of predictor variables,  $\varepsilon_t$  is an error term, and  $\theta_{ht}$  is a vector of coefficients which may vary over time. Different forecasting methods use different specifications of the  $f$  function, the coefficients  $\theta_{ht}$  and the predictor variables  $Z_t$ . The details will be given shortly.

An  $h$ -step forecast is obtained as

$$\hat{y}_{t+h}^h = f(Z_t; \hat{\theta}_{ht}), \quad (3.2)$$

and the forecast error is

$$e_{t+h} = y_{t+h}^h - \hat{y}_{t+h}^h = y_{t+h} - \hat{y}_{t+h} \quad (3.3)$$

independently of whether  $y_t$  is treated as stationary or integrated. Thus, the forecast errors are directly comparable. This is convenient because we also consider a pre-test forecast where the decision on the stationarity of  $y_t$  is based on a unit root test, which often improves the forecasting performance, see e.g. Diebold & Kilian (2000). Specifically, we use the Elliott, Rothenberg & Stock (1996) DF-GLS statistics, which performed best in the simulation experiments in Stock (1996).

As in Marcellino (2004), forecasts 1, 2 and 4 quarters ahead will be considered. For each forecast horizon,  $h$ , a model of the general type (3.1) is fitted and used for forecasting. In other words, we use an ‘ $h$ -step ahead projection’ or dynamic estimation approach (e.g., Clements & Hendry (1996)). It differs from the standard approach which estimates a one-step ahead model and uses this model to obtain  $h$ -step ahead predictions by iterating that model forward by  $h$  steps. In our framework the  $h$ -step ahead projection approach has the advantage that forecasts from nonlinear

models are directly available and need not be determined by simulation methods. Moreover, the potential impact of specification error in the one-step ahead model can be reduced by using the same horizon for estimation as for forecasting. The resulting forecasts could be slightly less efficient, see e.g. Granger & Teräsvirta (1993, Ch. 8) and Marcellino, Stock & Watson (2005), but the computational savings in our real time exercise with several series are substantial.

We now sketch the methods and models to be used in the forecast comparison.

### 3.1.1 Linear methods

*Autoregression (AR)*. Several studies found that simple linear AR models are very good forecasting tools for economic variables, see e.g. Meese & Geweke (1984), or Marcellino, Stock & Watson (2003) for the Euro-area. For these models the  $f$  function in (3.1) is linear, and  $Z_t$  includes lags of  $y_t$  and a deterministic component. In our comparison the latter can be either a constant or a linear trend. The lag length is either fixed at 4, or it is chosen by AIC or BIC with a maximum of 4 lags. The variable  $y_t$  can be treated as stationary or  $I(1)$ . Moreover, the choice between the two may depend on a pre-test for a unit root. Taking into account all these possibilities, there are 18 models in this class. We label them as AR1 - AR18 as specified in Table 2.

The following strategy is used when pretests for unit roots are applied. If a trend is included in the original model, a trend is maintained if the unit root is rejected whereas only a constant is included in the model for the differences if a unit root cannot be rejected. In contrast, if a model with only a constant is used in the unit root pretest, a constant is included irrespective of the outcome of the unit root test in the models for the levels and the models for first differences.

*No change*. As a possible benchmark we also consider a ‘no change forecast’, that is,  $\hat{y}_{t+h} = y_t$ . Although this model is optimal only if the data generation process is a random walk it was found to outperform even forecasts from large-scale structural models in some forecast comparisons, see e.g. Artis & Marcellino (2001). This method is abbreviated as NOCHG in our forecast comparison (see Table 2).

### 3.1.2 Time-varying methods

Under the heading time-varying methods we summarize models with time-varying coefficients which could also be classified as nonlinear models. For ease of reference to other studies, notably Stock & Watson (1999) and Marcellino (2004), we call them time-varying models because this is the classification used in the reference articles.

*Time-varying autoregression (TVAR)*. In the time-varying AR models the function  $f$  in (3.1) is again linear and  $Z_t$  contains a constant and lags of  $y_t$ . Now the coefficients evolve according to the following multivariate random walk model (see e.g. Nyblom (1989)):

$$\theta_{ht} = \theta_{ht-1} + u_{ht}, \quad u_{ht} \sim iid(0, \lambda^2 \sigma^2 Q), \quad (3.4)$$

where  $\sigma^2$  is the variance of the error term  $\varepsilon_t$ ,  $Q = (E(Z_t Z_t'))^{-1}$ , and we inspect several values of  $\lambda$  : 0 (no evolution), 0.0025, 0.005, 0.0075, 0.01, 0.015, or 0.020. We consider three possible setups, (1) a fixed specification with 2 lags of  $y_t$  in  $Z_t$  and  $\lambda = 0.005$ , (2) the model selected by AIC from the set of models with 1, 2 or 4 lags and  $\lambda$  taking on the aforementioned values, (3)

selecting the model by BIC from the same set of possibilities as under (2). In each of these cases, we include either  $y_t$  or  $\Delta y_t$  or we choose between the two on the basis of a unit root pre-test. Thereby we get a total of 9 TVAR models. They are numbered as stated in Table 2. The models are estimated by the Kalman filter.

*Logistic smooth transition autoregression (LSTAR).* The LSTAR( $p$ ) model is of the form

$$y_{t+h}^h = \alpha' Z_t + d_t \beta' Z_t + \varepsilon_{t+h}, \quad (3.5)$$

where  $d_t = 1/(1 + \exp(\gamma_0 + \gamma_1 \zeta_t))$ , and  $Z_t = (1, y_t, y_{t-1}, \dots, y_{t-p+1})'$  if  $y_t$  is treated as stationary or  $Z_t = (1, \Delta y_t, \Delta y_{t-1}, \dots, \Delta y_{t-p+1})'$  if  $y_t$  is viewed as an integrated variable. The smoothing parameters  $\gamma_1$  regulate the shape of parameter change over time. When  $\gamma_1 = 0$  the model becomes linear, while for large values of  $\gamma_1$  the model tends to a self-exciting threshold model (SETAR), see e.g. Granger & Teräsvirta (1993), Teräsvirta (1998) for details. For models specified in levels we consider the following choices for the threshold variable in  $d_t$ :  $\zeta_t = y_t$ ,  $\zeta_t = y_{t-1}$ ,  $\zeta_t = y_{t-3}$ ,  $\zeta_t = y_t - y_{t-2}$ ,  $\zeta_t = y_t - y_{t-4}$ . If the variable under consideration is treated as  $I(1)$ , the threshold variable is  $\zeta_t = \Delta y_t$ ,  $\zeta_t = \Delta y_{t-1}$ ,  $\zeta_t = \Delta y_{t-3}$ ,  $\zeta_t = y_t - y_{t-2}$  or  $\zeta_t = y_t - y_{t-4}$ . Three different values for the lag length  $p$  are considered: 1, 2 and 4. In total the 12 models listed in Table 2 are considered in the comparison. AIC or BIC select from a choice of models with  $p = 1, 2, 4$  and all the specifications of  $\zeta_t$  mentioned in the foregoing. Estimation is carried out by (recursive) non-linear least squares, using an optimizer developed by Stock & Watson (1999).

Overall, there are 19 linear and 21 time-varying or nonlinear models in the forecast comparison exercise. They are listed in Table 2. Initially we have also used other forecasting methods. For example, we have used the exponential smoothing and neural network models which were also considered by Marcellino (2004). The results from those models did not affect our general conclusions, however. Therefore we have eliminated them to make the study more lucid.

### 3.2 Forecast evaluation

To mimic real time situations, for each variable, method and model the unit-root tests, estimation and model selection are repeated each quarter over the forecasting period, which is 2000Q1-2003Q4. Thus, the pre-forecast period for the shortest estimation and specification sample has  $T = 120$  observations and the forecast period consists of 16 quarters. The usual mean squared error (MSE) is used as loss function. For forecast horizon  $h$ , model  $m$  and variable  $n$  with type of data  $j$  we specify

$$\text{MSE}_{n,m,j}^h = \frac{1}{17-h} \sum_{t=T+h}^{T+17-h} (e_{t,n,m,j})^2. \quad (3.6)$$

To simplify the comparison we express each MSE relative to that of a benchmark specification.

## 4 Results

We present selected results in Figures 3 - 5 and Table 3. In Figure 3 the MSEs of 1-step ahead forecasts from the linear AR models are compared for all ten variables. The MSEs are given

relative to an AR model with 4 lags and a constant, specified in levels (AR1 in Table 2) estimated with the aggregated Euro-area data. It turns out that for the real variables where substantial shifts occur when combining German with Euro-area data, using aggregated European data results in better forecasts. This is true for the five variables shown on the left-hand side of Figure 3 (YER, PCR, GCR, ITR, SCR). For almost all the forecasting models the solid bar is the smallest for these variables. Thus, whichever model is chosen, forecasts based on aggregated European data tend to be better than those based on German data for these variables.

The situation is clearly different on the right-hand side of Figure 3 which shows the results for the monetary variables. Here the situation for some variables is just the opposite. For example, for the price indices YED and CPI the forecasts based on German data tend to be superior to those based on aggregated Euro-area data. Recall that no adjustments were made for the interest rate series. Therefore the GER1 and GER2 series are identical and, hence, the same is true for the MSEs in Figure 3. For the short-term interest rate (STR) the aggregated Euro-area data have a slight advantage over the German data while the ranking is reversed for the long-term interest rate (LTN). For the exchange rate (EER) it depends on the model used which data set is preferable. The overall best forecasts are obtained with German data, however.

The overall ranking is similar for forecast horizons 2 and 4. Therefore we do not show the detailed results. We conclude that using German data rather than aggregated European data offers no advantages for the real economy variables which require major adjustments in joining the German and European data. In the following we will therefore take a closer look at the monetary variables for which further results are given in Figures 4 and 5 and Table 3.

Clearly, it is possible that more sophisticated models capture the structure of the aggregated DGP better than the linear models. Therefore we compare forecasts based on linear and time-varying or nonlinear models in Figure 4. In that figure the MSEs of the 21 nonlinear models relative to the best linear model are shown for forecast horizon  $h = 1$  based on the EUR and GER0 data. Clearly there is nothing to gain from using our nonlinear models for the aggregated Euro-area data for YED, CPI, LTN and EER. In fact, many of the nonlinear models produce substantially inferior forecasts relative to the best linear forecasts for these variables. Only for the short-term interest rate (STN) the nonlinear methods may produce slightly better forecasts than linear models based on the EUR data.

The situation is somewhat different if the unadjusted German data (GER0) data are considered. Although no dramatic gains are obtained from using nonlinear/time varying coefficient models, small forecast improvements are obtained for YED and EER. In particular, specific LSTAR models improve forecast precision relative to linear AR models for these variables.

Of course, one may wonder whether the unadjusted German data are the most useful ones for prediction of these variables. Therefore we show the MSEs obtained with the linear models for YED, CPI and EER based on GER1 and GER2 relative to the corresponding GER0 MSEs in Figure 5. Clearly, the adjustments may be beneficial for YED and CPI if only linear models and one-quarter forecasts are compared. On the other hand, adjustments don't seem to help much for forecasting the exchange rate. Thus, it appears that different data treatments and forecasting methods are optimal for different variables. Hence, it is worth taking a closer look at the results for the individual monetary variables and to perform a comparison across all models and data sets

for each of them.

The three best forecasts from the whole range of methods and data sets used for horizons  $h = 1, 2$  and  $4$  for the five monetary variables are listed in Table 3. Note that there are a number of ties where different methods result in the same model and hence the same MSE. For example, the same model may be chosen by AIC and BIC. Also LSTAR models may reduce to linear models. In such cases we list the first model from Table 2 and present those models with the same MSE in the table footnote.

It turns out that the best forecasts for the two price series YED and CPI, the long-term interest rate (LTN) and the exchange rate (EER) are obtained with German data. In fact, for these four variables there is only one instance where a forecast based on European data is among the first three best forecasts. More precisely, the one-quarter ahead forecast for CPI based on an LSTAR model for European data produces the second best MSE for this variable. The situation is just the opposite for the short-term interest rate (STN). For this series the best three forecasts for each of the three forecast horizons are obtained with European data. Thus, it appears that for all but one of the monetary variables using German data is preferable. As we have already observed in Figure 5, the type of German data resulting in the best forecasts also differs for different variables.

For the GDP deflator (YED) the best one-quarter ahead forecasts are obtained with unadjusted German data (GER0) and the best four-quarters ahead forecasts result from using GER2. Linear AR models with a time trend for the levels variable are among the first three best models for each of the three forecast horizons. Only for one-quarter ahead forecasts the LSTAR and TVAR models appear among the first three best methods. Thus, even if linear models are used, the slight shift in the GER0 series in 1999 (see Figure 2) can be accounted for well enough to get quite good forecasts. For longer-term forecasts a simple adjustment with dummy variables is preferable, however.

An adjustment with dummies (GER2) also gives the best forecasts for the CPI for all horizons. For this variable the LSTAR models are usually quite successful although a TVAR model comes in first for the one-quarter horizon. On the other hand, the linear AR models provide slightly less precise forecasts. Notice that, although the second best four-quarters ahead forecast is based on a linear AR model, its MSE is already quite a bit larger than for the best LSTAR forecast.

As mentioned earlier, the best forecasts for the short-term interest rate (STN) are based on aggregated European data. This result may be somewhat surprising given that some authors have argued that the German Bundesbank has dominated European monetary policy in the pre-Euro period to some extent (e.g., Giavazzi & Giovannini (1989), Kirchgässner & Wolters (1993)). Another interesting result for this variable is that the linear AR models dominate among the best forecasting models, although they come in only second for one-quarter ahead forecasting. The difference in MSE to the best nonlinear forecast in this case is a noticeable 20%. However, for two- and four-quarters ahead forecasts the smoothing due to aggregation may in fact have simplified the structure of the DGP in such a way that simple linear models provide a good description of the series. Note also that the best three forecasts for each horizon are obtained from models in first differences which indicates that allowing for a unit root in this series is a reasonable strategy.

The long-term interest rate is best forecast based on German data. Linear AR models provide quite good forecasts. Although they are first only for four-quarters ahead forecasting, they have

MSEs rather close to the best nonlinear forecasts also for horizons  $h = 1$  and  $h = 2$ . It is interesting, however, that for  $h = 1$  the best linear forecasting model is based on the levels variable while first differences are preferable for longer horizons.

In forecasting the exchange rate (EER) the German GER1 series with a simple statistical adjustment gives the best results. This may be due to the fact that the adjustment by dummies shifts the series considerably after German unification (see Figure 2). This shift may be a slight over-adjustment. The nonlinear/time-varying models dominate the picture and result in much better one- and four-quarters ahead forecasts than the best linear models. Thus, there may be nonlinear mechanisms at work in the generation process of this variable and simple linear modelling may not capture these complex dynamics. Note also that a model in first differences is best for one-quarter ahead forecasting whereas models for levels predict better two and four quarters ahead. Our result that German data are preferable to European data for predicting the exchange rate contrasts with the finding by Nautz & Offermanns (2006) that the Euro/US dollar exchange rate can be better forecast with synthetic European data than with German mark/US dollar exchange rate data. Reasons may be that these authors use monthly data for a different exchange rate variable and they base their forecasts on exchange rate equations which involve also other variables rather than just the information in the past exchange rates. Thus, they consider conditional forecasts which also involve a real time series. Moreover, they use only linear models for prediction.

Clearly, there is not a single forecasting method that works best for all variables. It is also not necessarily better to use levels or first differences or decide on the choice with the help of unit root tests. Also, including a linear trend term or just a constant has different implications for the forecast accuracy for different variables. Which method works best depends on the specific variable and its characteristics. Thus, none of the methods can be recommended as a universal forecasting tool. Finding such a tool is not the objective of this study, of course. For us it is of main interest to see whether there is information in the aggregate Euro-area data on the future development of specific variables which is not available in German data for the pre-EMU period.

The overall conclusion from our forecast comparison is that four of the monetary variables can be better forecast based on German data than with aggregated European data. Thus, our forecasting exercise suggests that using German data rather than aggregated Euro-area data for the pre-EMU period may be beneficial also for other purposes. Some of the German series need adjustments when they are joined with the EMU series. Doing this adjustment well can help to improve the forecast quality. Although relatively simple adjustment methods produced good results in our forecast comparison, some attention has to be paid to this issue for each individual variable. There are variables, however, for which using the aggregated European series may have advantages. Clearly, for real variables, where major adjustments are necessary when combining German and EMU data, using the aggregated European data seems to have advantages. Of course, it is possible that more sophisticated ways to join the series may change the picture also for these variables.

## 5 Conclusions

In this study we have investigated whether German data from the pre-Euro period have the same or even more information content for forecasting Euro-area variables in the EMU period than aggre-

gated, synthetic European data. Given the problems related to data aggregation from a heterogeneous set of countries and given that Germany did not have to make major adjustments to satisfy the Maastricht criteria, it is plausible to expect that at least for some variables the German data may be preferable. We have used 10 quarterly macroeconomic variables for the period 1970Q1 - 2003Q4 and have determined forecasts with a range of different linear and nonlinear/time-varying models. The MSE is used for forecast evaluation and comparison. Five variables are from the real economy and the remaining ones are nominal and financial variables. We have used three different ways of joining German and EMU data and we compare forecasts based on those series with forecasts based on synthetically aggregated European data.

For the real economy variables it turns out that using European data results in superior forecasts. This result may be partly due to the difficulties in joining German and EMU data. Such difficulties are not present or less important for the monetary variables. For four out of five variables the forecasts based on German data are clearly superior. Only for the short-term interest rate an aggregated European series is preferable. It depends on the individual series and in some cases also on the forecast horizon which method forecasts best, as one would expect. In some cases nonlinear models do better than linear ones.

Our results suggest that longer time series constructed from German data may be useful also for analyzing Euro-area models. At least for monetary variables for which German data can be joined easily with EMU data, using such time series may be advantageous. Admittedly, we have only used univariate forecasts in this study and it is not clear that we can draw firm conclusions for a multivariate context. The results indicate, however, that using the monetary German data has at least some potential of being useful and, hence, it may be worth trying in future studies.

## References

- Artis, M. & Marcellino, M. (2001). Fiscal forecasting: The track record of IMF, OECD and EC, *Econometrics Journal* **4**: 20–36.
- Beyer, A., Doornik, J. A. & Hendry, D. F. (2001). Reconstructing historical euro-zone data, *Economic Journal* **111**: 308–327.
- Bosker, E. M. (2006). On the aggregation of eurozone data, *Economics Letters* **90**: 260–265.
- Brüggemann, R. & Lütkepohl, H. (2006). A small monetary system for the euro area based on German data, *Journal of Applied Econometrics* p. forthcoming.
- Clements, P. M. & Hendry, D. F. (1996). Multi-step estimation for forecasting, *Oxford Bulletin of Economics and Statistics* **58**: 657–684.
- Diebold, F. X. & Kilian, L. (2000). Unit root tests are useful for selecting forecasting models, *Journal of Business and Economic Statistics* **18**: 265–273.
- Elliott, G., Rothenberg, T. J. & Stock, J. H. (1996). Efficient tests for an autoregressive unit root, *Econometrica* **64**: 813–836.

- Fagan, G., Henry, J. & Mestre, R. (2001). An area-wide model for the Euro area, *Working Papers 42*, European Central Bank.
- Giavazzi, F. & Giovannini, A. (1989). *Limiting Exchange Rate Flexibility, The European Monetary System*, MIT Press, Cambridge, MA.
- Granger, C. W. J. & Teräsvirta, T. (1993). *Modelling non-linear economic relationships*, Oxford: Oxford University Press.
- Kirchgässner, G. & Wolters, J. (1993). Does the DM dominate the Euro market? An empirical investigation, *Review of Economics and Statistics* **75**: 773–778.
- Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*, Springer-Verlag, Berlin.
- Marcellino, M. (2004). Forecasting EMU macroeconomic variables, *International Journal of Forecasting* **20**: 359–372.
- Marcellino, M., Stock, J. H. & Watson, M. W. (2003). Macroeconomic forecasting in the Euro Area: Country specific versus area-wide information, *European Economic Review* **47**(1): 1–18.
- Marcellino, M., Stock, J. H. & Watson, M. W. (2005). A comparison of direct and iterated multistep AR methods for forecasting macroeconomic series h-steps ahead, *Journal of Econometrics*.
- Meese, R. & Geweke, J. (1984). A comparison of autoregressive univariate forecasting procedures for macroeconomic time series, *Journal of Business and Economic Statistics* **2**: 191–200.
- Nautz, D. & Offermanns, C. J. (2006). Does the Euro follow the German Mark, *European Economic Review* **50**: 1279–1295.
- Nyblom, J. (1989). Testing for constancy of parameters over time, *Journal of the American Statistical Association* **84**: 223–230.
- Stock, J. H. (1996). VAR, error correction and pretest forecasts at long horizons, *Oxford Bulletin of Economics and Statistics* **58**: 685–701.
- Stock, J. H. & Watson, M. W. (1999). A comparison of linear and non-linear univariate models for forecasting macroeconomic time series, in R. Engle & R. White (eds), *Cointegration, causality, and forecasting: A festschrift in honor of Clive W.J. Granger*, Oxford: Oxford University Press, pp. 507–552.
- Teräsvirta, T. (1998). Modelling economic relationships with smooth transition regressions, in A. Ullah & D. E. A. Giles (eds), *Handbook of Applied Economic Statistics*, New York: Marcel Dekker, pp. 507–552.
- 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 1: Variables in our data set

Variable	Name	Data Sources	
		Aggregated Euro-Area Data	German/European
YER	real GDP	OECD: 70Q1-03Q4	OECD: 70Q1-03Q4
PCR	real personal consumption	OECD: 70Q1-03Q4	OECD: 70Q1-03Q4
GCR	real government consumption	OECD: 70Q1-03Q4	OECD: 70Q1-03Q4
ITR	investment	OECD: 70Q1-03Q4	OECD: 70Q1-03Q4
SCR	inventories	OECD: 70Q1-03Q4	OECD: 70Q1-03Q4
YED	GDP deflator	OECD: 70Q1-03Q4	OECD: 70Q1-03Q4
CPI	Consumer price index	FHM: 70Q1-94Q4	Buba: 70Q1-98Q4
		ECB: 95Q1-03Q4	ECB: 99Q1-03Q4
STN	short-term interest rate	FHM: 70Q1-93Q4	Buba: 70Q1-98Q4
		ECB: 94Q1-03Q4	ECB: 99Q1-03Q4
LTN	long-term interest rate	FHM: 70Q1-93Q4	Buba: 70Q1-98Q4
		ECB: 94Q1-03Q4	ECB: 99Q1-03Q4
EER	real exchange rate	FHM: 70Q1-92Q4	OECD: 70Q1-98Q4
		ECB: 93Q1-03Q4	ECB: 99Q1-03Q4

*Note:* OECD: Data taken from OECD statistical compendium. FHM: Data taken from Fagan et al. (2001). ECB: Data taken from European Central Bank monthly bulletin. Buba: Data taken from database of the Deutsche Bundesbank.

Table 2: Forecasting methods and models

<b>Linear methods</b>	
	<i>Autoregressive models (18 models)</i>
AR1	AR(4) with constant, levels
AR2	AR(4) with linear trend, levels
AR3	AR(4) with constant, first differences
AR4	AR(4) with linear trend, first differences
AR5	AR(4) with constant, pretest for unit root
AR6	AR(4) with linear trend, pretest for unit root
AR7	AIC-AR with constant, levels
AR8	AIC-AR with linear trend, levels
AR9	AIC-AR with constant, first differences
AR10	AIC-AR with linear trend, first differences
AR11	AIC-AR with constant, pretest for unit root
AR12	AIC-AR with linear trend, pretest for unit root
AR13	BIC-AR with constant, levels
AR14	BIC-AR with linear trend, levels
AR15	BIC-AR with constant, first differences
AR16	BIC-AR with linear trend, first differences
AR17	BIC-AR with constant, pretest for unit root
AR18	BIC-AR with linear trend, pretest for unit root
NOCHG	No change forecast (1 model)
<b>Time-varying methods</b>	
	<i>Random coefficient AR models (9 models)</i>
TV1	random coefficient AR(2) with constant, levels
TV2	random coefficient AR(2) with constant, first differences
TV3	random coefficient AR(2) with constant, pretest for unit root
TV4	random coefficient AIC-AR with constant, levels
TV5	random coefficient AIC-AR with constant, first differences
TV6	random coefficient AIC-AR with constant, pretest for unit root
TV7	random coefficient BIC-AR with constant, levels
TV8	random coefficient BIC-AR with constant, first differences
TV9	random coefficient BIC-AR with constant, pretest for unit root
	<i>Logistic smooth transition (12 models)</i>
LST1	LSTAR(2), transition variable $y_t$ , levels
LST2	LSTAR(2), transition variable $y_t$ , first differences
LST3	LSTAR(2), transition variable $y_t$ , pretest for unit root
LST4	LSTAR(2), transition variable $y_t - y_{t-4}$ , levels
LST5	LSTAR(2), transition variable $y_t - y_{t-4}$ , first differences
LST6	LSTAR(2), transition variable $y_t - y_{t-4}$ , pretest for unit root
LST7	AIC-LSTAR, levels
LST8	AIC-LSTAR, first differences
LST9	AIC-LSTAR, pretest for unit root
LST10	BIC-LSTAR, levels
LST11	BIC-LSTAR, first differences
LST12	BIC-LSTAR, pretest for unit root

Table 3: The three best models for ‘nominal’ and ‘financial’ variables

Variable	$h = 1$			$h = 2$			$h = 4$		
	Model	Data	MSE	Model	Data	MSE	Model	Data	MSE
YED	LST8 <sup>a</sup>	GER0	0.565	AR2	GER0	0.559	AR8	GER2	0.272
	AR8 <sup>b</sup>	GER0	0.672	AR2 <sup>c</sup>	GER1	0.575	AR14	GER2	0.278
	TV5 <sup>d</sup>	GER0	0.702	AR2	GER2	0.594	AR2	GER2	0.308
CPI	TV5 <sup>d</sup>	GER2	0.745	LST7	GER2	0.481	LST10	GER2	0.289
	LST10	EUR	0.765	LST10	GER2	0.498	AR2 <sup>e</sup>	GER0	0.397
	LST7	GER2	0.766	AR2 <sup>e</sup>	GER0	0.569	LST11 <sup>f</sup>	GER0	0.488
STN	LST8 <sup>g</sup>	EUR	0.512	AR4	EUR	0.405	AR16	EUR	0.155
	AR3 <sup>h</sup>	EUR	0.622	AR3 <sup>h</sup>	EUR	0.414	AR4 <sup>i</sup>	EUR	0.171
	TV2 <sup>j</sup>	EUR	0.626	LST2 <sup>k</sup>	EUR	0.417	LST8 <sup>a</sup>	EUR	0.201
LTN	TV5 <sup>d</sup>	GER0	0.704	TV5 <sup>d</sup>	GER0	0.567	AR16	GER0	0.282
	LST5 <sup>l</sup>	GER0	0.708	AR16	GER0	0.581	AR10	GER0	0.303
	AR2	GER0	0.724	AR15 <sup>m</sup>	GER0	0.592	AR4	GER0	0.304
EER	LST2 <sup>k</sup>	GER1	0.764	AR1 <sup>n</sup>	GER1	0.722	LST7	GER1	0.555
	LST2 <sup>k</sup>	GER0	0.777	TV7	GER1	0.756	TV7	GER1	0.589
	AR1	GER1	0.858	LST4	GER1	0.757	AR1	GER1	0.594

*Note:* The comparison includes all linear and non-linear models based on all four data variants, i.e. aggregated European data (EUR) and German data (GER0, GER1 and GER2). MSEs are relative to AR1 benchmark model for aggregated Euro-area data.

<sup>a</sup>LST9 leads to same MSE. <sup>b</sup>AR14, LST11 and LST12 lead to same MSE. <sup>c</sup>AR8 leads to same MSE. <sup>d</sup>TV6, TV8 and TV9 lead to same MSE. <sup>e</sup>AR8 and AR14 lead to same MSE. <sup>f</sup>LST12 leads to same MSE. <sup>g</sup>LST9, LST11 and LST12 lead to same MSE. <sup>h</sup>AR5 and AR6 lead to same MSE. <sup>i</sup>AR10 leads to same MSE. <sup>j</sup>TV3 leads to same MSE. <sup>k</sup>LST3 leads to same MSE. <sup>l</sup>LST6 leads to same MSE. <sup>m</sup>AR17 and AR18 lead to same MSE. <sup>n</sup>AR7 leads to same MSE.

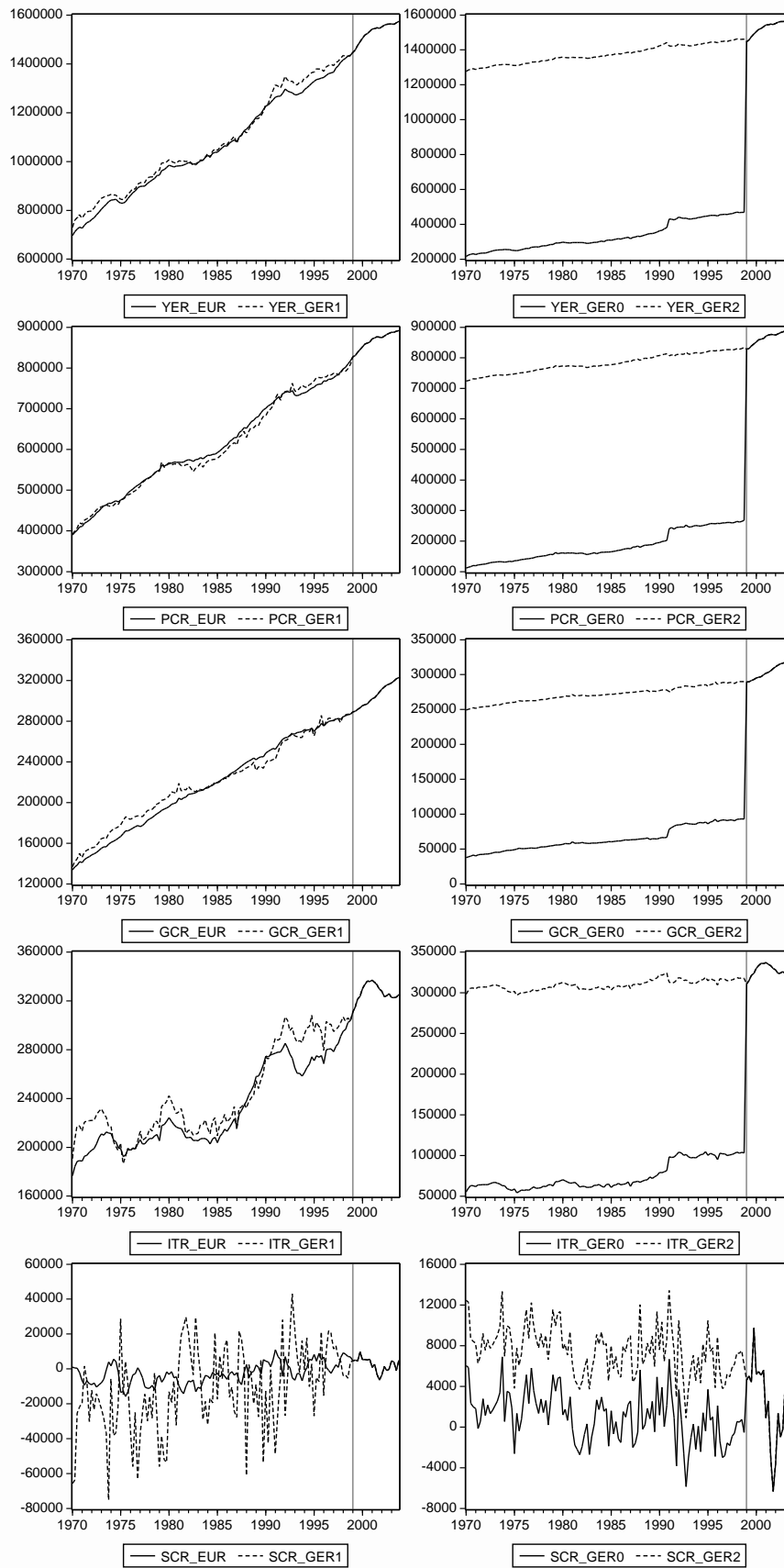


Figure 1: Aggregated Euro-area time series (EUR) and time series based on German data for real variables.

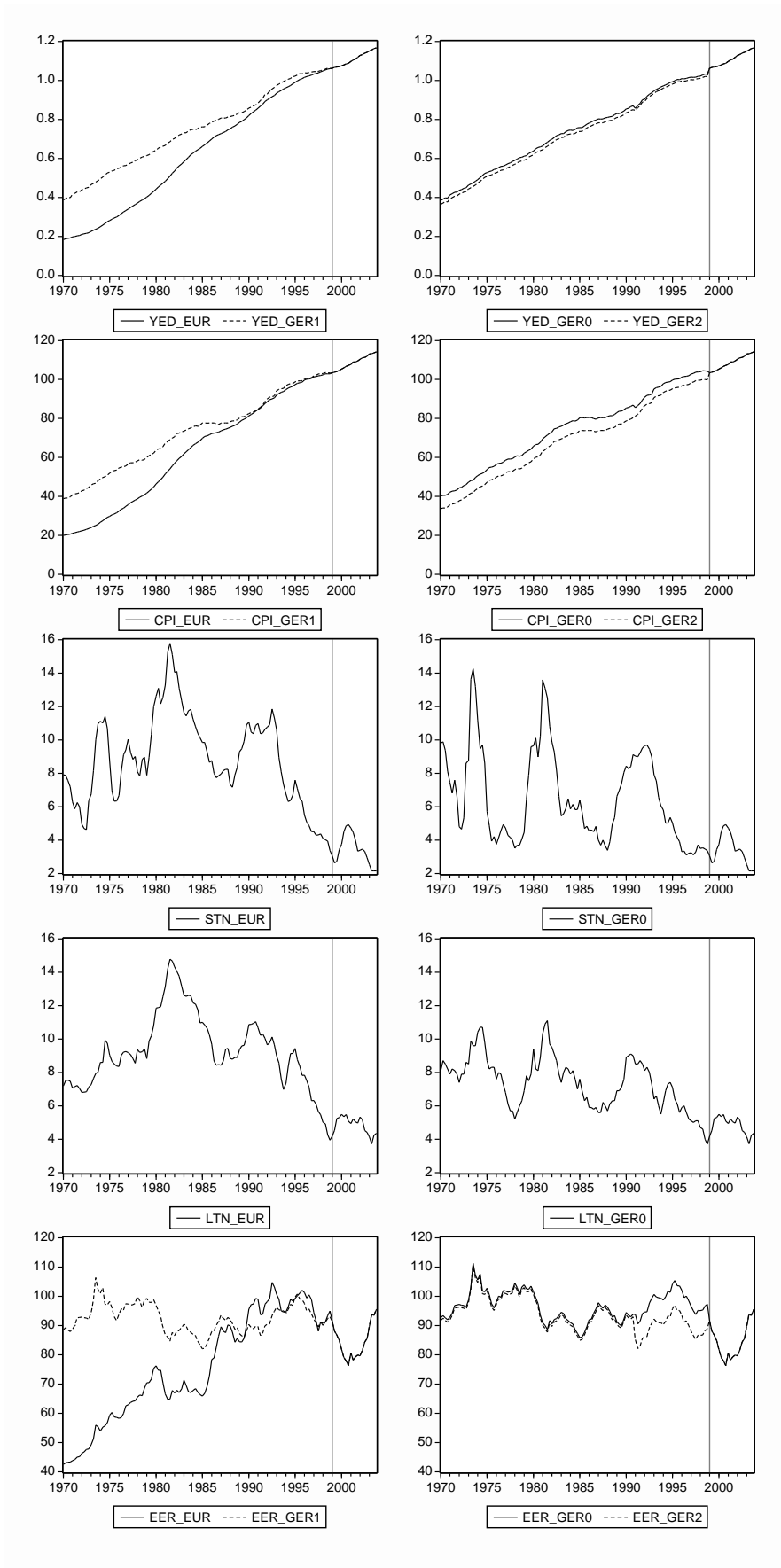


Figure 2: Aggregated Euro-area time series (EUR) and time series based on German data for nominal and financial variables.

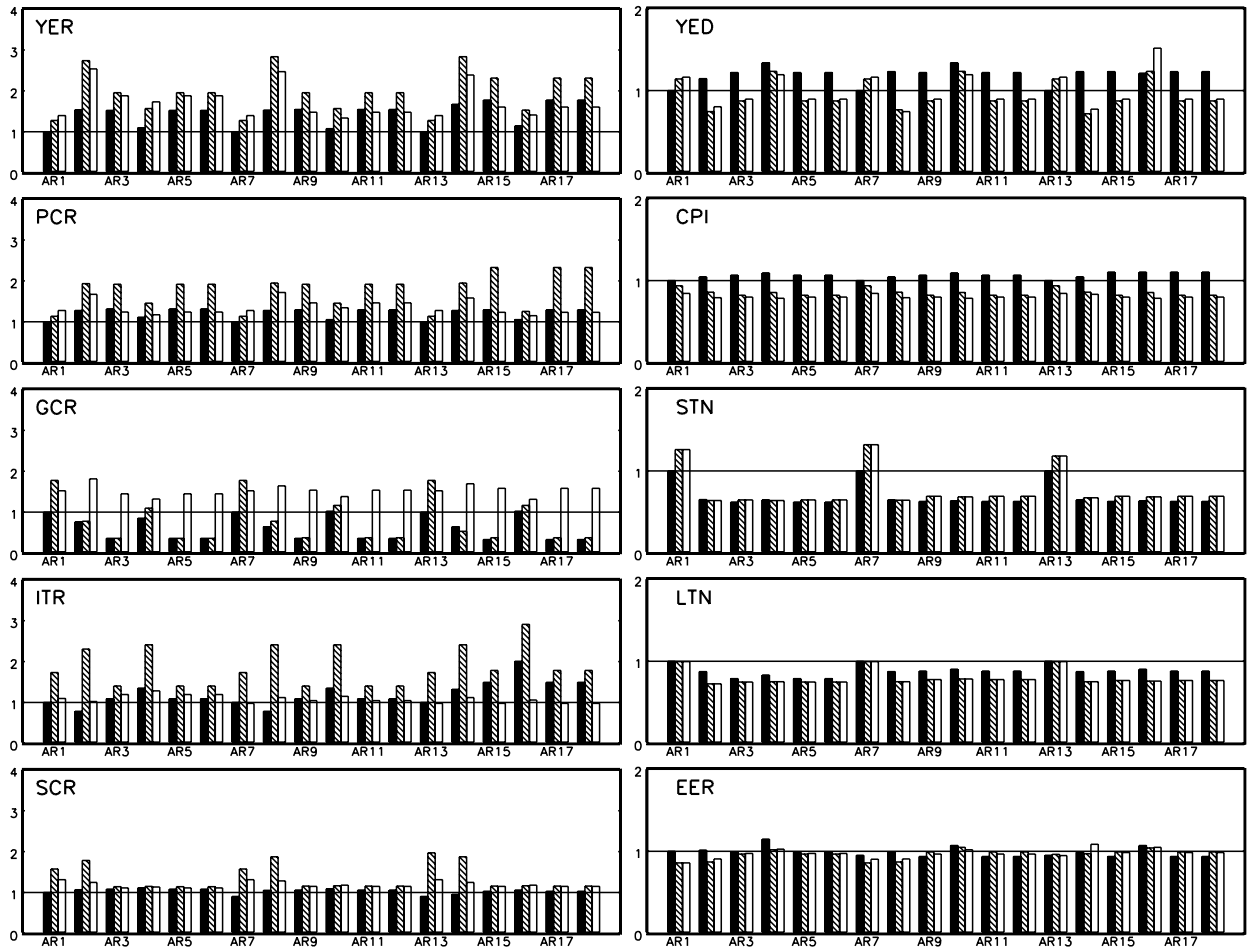


Figure 3: MSEs relative to the AR(4) model with a constant, specified in levels (AR1) estimated with aggregated Euro-area data ( $h = 1$ ). The figure shows results for the 18 linear AR models. Solid and dashed and white bars denote results from models based on aggregated Euro-area data and German data with statistical adjustment (GER1) and German data with adjustment by dummies (GER2), respectively.

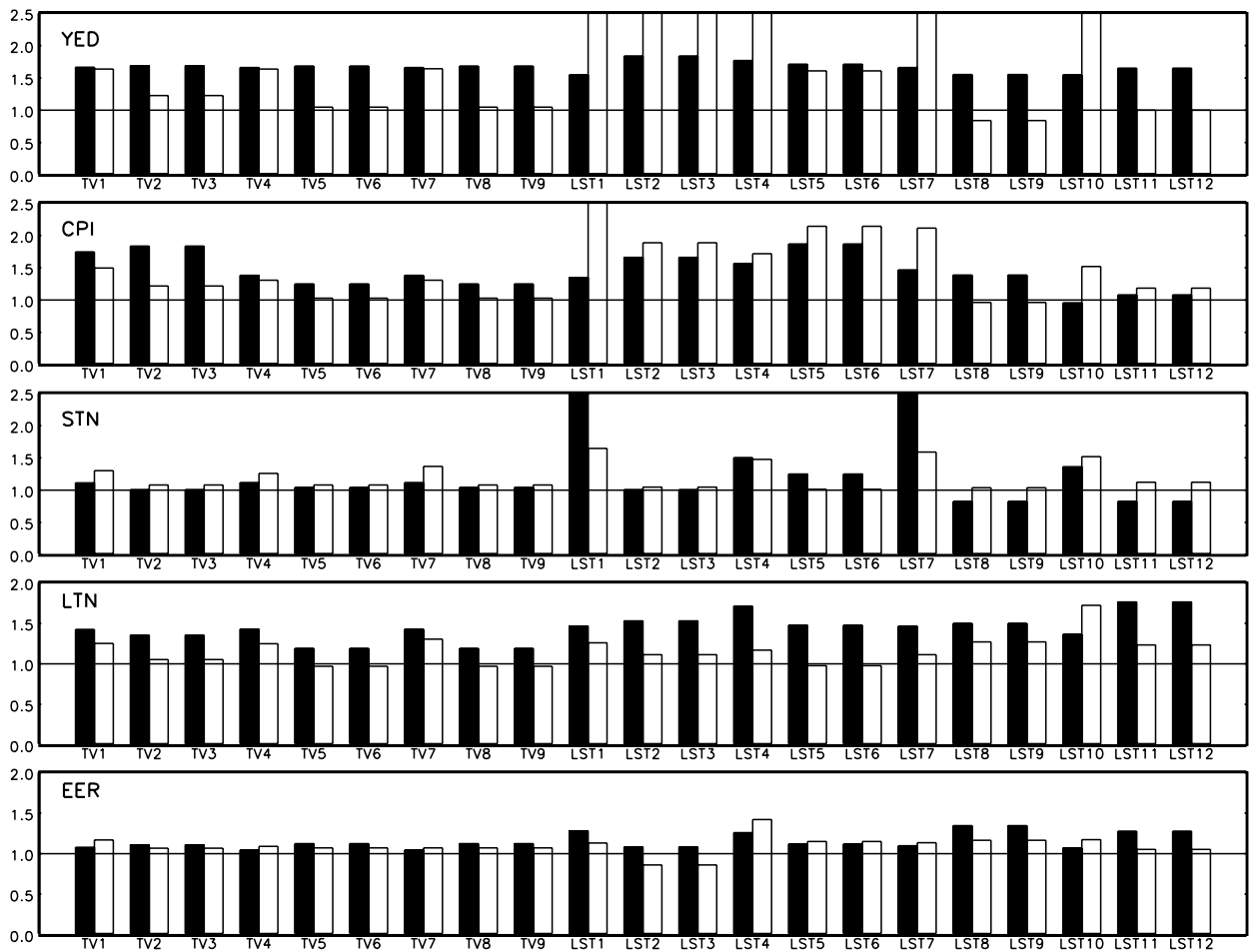


Figure 4: MSEs of time-varying methods relative to the respective best linear AR for ‘nominal’ and ‘financial’ variables ( $h = 1$ ). Solid, white bars denote results from models based on aggregated Euro-area data and white bars denote the corresponding results for German data without adjustment (GER0), respectively.

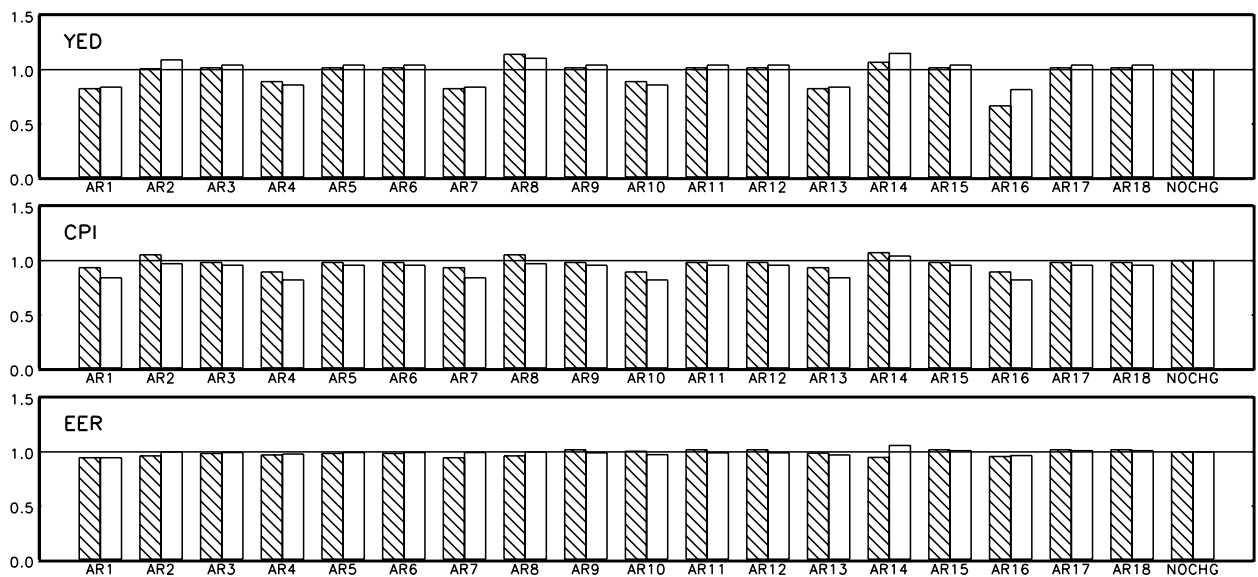


Figure 5: Comparison of adjustment methods: MSEs relative to corresponding GER0 model ( $h = 1$ ) for 19 linear methods. Dashed bar: GER1, white bar: GER2.