Ruhr Economic Papers

Roland Döhrn, Christoph M. Schmidt
and Tobias Zimmermann

Inflation Forecasting with Inflation Sentiment Indicators

#80

RWI ESSEN

Ruhr Economic Papers

Published by

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstraße 150, 44801 Bochum, Germany

Technische Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstraße 12, 45117 Essen, Germany

Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI Essen)
Hohenzollernstrasse 1/3, 45128 Essen, Germany

Ruhr Economic Papers #80

Responsible Editor: Christoph M. Schmidt

ISSN 1864-4872 (online) – ISBN 978-3-86788-087-9

Ruhr Economic Papers

#80

Roland Döhrn, Christoph M. Schmidt
and Tobias Zimmermann

Inflation Forecasting with Inflation Sentiment Indicators

RWI
ESSEN

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

ISSN 1864-4872 (online)
ISBN 978-3-86788-087-9

Roland Döhrn, Christoph M. Schmidt and Tobias Zimmermann*

Inflation Forecasting with Inflation Sentiment Indicators

Abstract

In this paper we argue that future inflation in an economy depends on the way people perceive current inflation, their inflation sentiment. We construct some simple measures of inflation sentiment which capture whether price acceleration is shared by many components of the CPI basket. In a comparative analysis of the forecasting power of the different inflation indicators for the US and Germany, we demonstrate that our inflation sentiment indicators improve forecast accuracy in comparison to a standard Phillips curve approach. Because the forecast performance is particularly good for longer horizons, we also compare our indicators to traditional measures of core inflation. Here, the sentiment indicators outperform the weighted median and show a similar forecasting power as a trimmed mean. Thus, they offer a convincing alternative to traditional core inflation measures.

JEL Classification: E30, E31, E37, C53

Keywords: Inflation forecasting, monetary policy

December 2008

* Roland Döhrn and Tobias Zimmermann, RWI Essen; Christoph M. Schmidt, RWI Essen, Ruhr-Universität Bochum, IZA, Bonn, and CEPR, London. – The authors thank Kai Carstensen, Jonas Dovern and Simeon Vosen for helpful comments to earlier versions of this paper, and Michael Kind, Claudia Lohkamp and Waltraud Lutze for their technical assistance. We are grateful to Steve Reed, Rob McClelland and Ken Stewart at the BLS for giving us the opportunity to work with the CPI research data. – All correspondence to Roland Döhrn, RWI Essen, Hohenzollernstr. 1-3, 45128 Essen, Germany, e-mail: doehrn@rwi-essen.de.

1. Introduction

In recent years, methodological aspects of inflation forecasting have increasingly attracted attention in empirical economics. One strand of the literature focuses on the determinants of future inflation such as monetary aggregates (e.g. Carstensen 2007) or the output gap (e.g. Stock and Watson 1999). The other concentrates on finding an adequate measure of the trend component of inflation which is expected to be a better predictor for overall inflation. A prominent role in this second group of papers plays *core inflation*, which is understood as the permanent part of inflation that is not being influenced by random short term fluctuations. A quite familiar concept for measuring core inflation is to exclude prices for energy and unprocessed food from the recorded basket of goods, since their prices tend to be very volatile without any trend (Gordon 1975). Alternatively, median based indicators or trimmed mean measures have been proposed (Bryan et al. 1994), as well as smoothed versions of these indicators (Rich and Steindel 2005). Moreover, another suggestion is to incorporate co-integration restrictions (Smith 2004).

Though having some aspects in common with median-based measures of core inflation, our paper focuses on another aspect of price trends. We construct an indicator capturing whether a given inflation rate is the result of similar price increases for many items in the goods basket, or whether it results from price hikes for a few relatively important goods (e.g. furniture or cars). Since the concrete way an overall increase in the price level is coming about will influence how it is perceived by consumers and firms, we label our indicator *inflation sentiment*. The same inflation rate may have different consequences for the future, depending on the distribution of price increases of individual items. When many prices are on the rise, inflation climate may have changed. Producers might be inclined to pass through higher costs because everybody does it, and workers might feel inflation to be more severe and struggle more fiercely for higher wages. As a consequence inflation will tend to increase further in the future.

To quantify this fundamental idea of inflation sentiments one might alternatively follow the concept developed by Brachinger (2006). He constructs his index of perceived inflation by re-weighting the components of German CPI according to their frequency of purchases. This approach requires information going beyond the data regularly provided by the CPI statistics, such as on the expenditures on all components of the CPI and the prices per unit purchased. Therefore it might be rather difficult to construct such indices for many countries for purposes of international comparison or to provide a long time series.

We propose a more simple methodology. It makes use of the fact that inflation measurement is standardized to some extent around the world. Most statistical offices collect price data in sufficient detail to provide price indices for the product categories enumerated in the *Classification of Individual Consumption by Purpose* (COICOP). The COICOP serves as a guideline for the disaggregation of private consumption expenditures in the national accounts.[1] For some product categories (e.g. food) prices for quite a number of goods are required to represent the heterogeneous basket. For other categories, e.g. electricity or heating gas, a small number of prices suffices to characterize price changes. We argue that, if the prices of many products (in a category) rise, consumers will perceive inflationary trends more intensely.

We use simple transformations of the data taken from the CPI statistics to construct several indicators of inflation sentiment. First, we calculate the *unweighted median*. By comparing it to the weighted mean of the price changes of the individual goods and services (the current CPI), we get an impression of the skewness of their distribution. If the median is larger than the weighted mean, the overall price trend reflects a relatively large number of similar price changes for individual products. In that situation we would expect consumers to perceive inflation more strongly. Secondly, we test a *diffusion index* measuring the share of prices which grow faster than current CPI. Finally, a *momentum index* is defined as the difference between the share of prices which grow faster than in the previous period, and the share of prices which grow more slowly. These indicators are calculated for quarterly data for the US and Germany.

In analysing the forecasting power of these indicators we follow the technique proposed in their seminal paper by Stock and Watson (1999). Specifically, we estimate several variants of the Phillips curve in which our indicators and, as in the standard procedure, the CPI itself are each used as the relevant price term, respectively. Starting with an initial sample length we employ all indicators for an out-of-sample forecast up to a maximum horizon of two years. Then, the sample is expanded by one quarter and another set of forecasts is made. By continuing this procedure, we generate a series of out-of-sample forecasts which can then be used to evaluate the forecasting power of all candidate approaches.

The remainder of this paper is organized as follows: In the next section we discuss the data, and the indicators that are used to characterize the inflation sentiment. Section 3 presents the econometric approach to forecast inflation and the methods which are employed to evaluate the estimates. In the fourth section we discuss our results. Section 5 concludes.

---

[1] See e.g. http://unstats.un.org/unsd/cr/registry/regcst.asp?Cl=5&Top=2&Lg=1.

2. Measuring inflation sentiment

In our empirical work we rely on three time series of *quarterly* inflation rates (Figure 1) that we construct as simple averages of seasonally adjusted monthly data. Our series for the US covers the period 1978:1 to 2006:4. It uses the CPI-U research data that were calculated by the Bureau of Labour Statistics to account for methodological changes. The number of goods and services included in this data set varies from 60 at the beginning to 212 at the end of the sample period.

[Figure 1: about here]

For Germany, we investigate two CPI series. The first covers the years 1985 to 1998 and relates to West Germany, the second provides data for unified Germany for the years 1993 to 2007[2]. These periods comprise quite different inflationary regimes. The first sub-period is characterized by large fluctuations in inflation rates; they stick very closely to the inflation objective in the more recent sub-period. The German CPI covers about 750 single items, which are aggregated in several steps to gain inflation rates for product categories.

[Table 1: about here]

To be able to derive comparable results for Germany and the US, we utilize these data on a four digit COICOP-level, which allows us to include about 100 product categories in our indices (Table 1). We employ three indicators to measure inflation sentiment.[3] Firstly, for each quarter t we construct the difference, s_t^{Med}, between (unweighted) median inflation, π_t^{Med}, and the headline inflation rate, π_t.

$$s_t^{Med} = \pi_t^{Med} - \pi_t . \qquad (2.1)$$

This indicator measures the skewness of the distribution of the price increases of the individual CPI components. When a high proportion of price components is rising, the unweighted median tends to be above the weighted average. In that case, we expect consumers and price setters to perceive inflation more strongly.

The second indicator of inflation sentiment, s_t^{Diff}, captures the difference between the share of prices which grow faster than current CPI, and the share of prices

---

[2] Price data for unified Germany are available since 1991. However, we omitted the years 1991 and 1992 because they were strongly influenced by irregularities in the follow-up of unification.

[3] Note that all inflation sentiment indicators should be stationary by construction. However, augmented Dickey-Fuller tests for the inflation sentiment indicators partly suggest that these variables may not be stationary and, therefore, should be included in differences. We include them in levels for two further reasons: Augmented Dickey-Fuller tests are not valid in small samples under 100 observations. Moreover, KPSS tests, which can also be found in the Appendix, largely confirm the stationarity assumption.

which grow more slowly than current CPI. Thus, we transfer the concept of the diffusion index, which was proposed by Burns and Mitchell (1946), to prices. Similar indicators are used in the context of technical share price analysis, under the term of “advance-decline”. With $\pi_{i,t}$ being the price increase of *i*-th of *N* goods and services in quarter *t* and π_t still being the headline inflation rate, the indicator is formally calculated as:

$$s_t^{Diff} = \tfrac{1}{N}\sum_{i=1}^{N} V^{Diff}(\pi_{i,t}), \quad V^{Diff}(\pi_{i,t}) = \begin{cases} 1 \text{ for } \pi_{i,t} \geq \pi_t \\ 1 \text{ for } \pi_{i,t} < \pi_t \end{cases}. \qquad (2.2)$$

To capture the momentum of inflation growth a third indicator, s_t^{Mom}, calculates the difference between the share of the price series which exhibit an increasing growth rate and the share of prices which show a decreasing growth rate:

$$s_t^{Mom} = \tfrac{1}{N}\sum_{i=1}^{N} V^{Mom}(\pi_{i,t}), \quad V^{Mom}(\pi_{i,t}) = \begin{cases} 1 \text{ for } \pi_{i,t} \geq \pi_{i,t-1} \\ 1 \text{ for } \pi_{i,t} < \pi_{i,t-1} \end{cases}. \qquad (2.3)$$

None of our inflation sentiment indicators does require any *a priori* assessments on the issue of which might be the products with highly volatile prices. Those would have to be excluded, if we were trying to measure core inflation. Instead, in our analysis all individual price series are used to capture whether inflation is broad-based or not. Moreover, our approach does not require any explicit expenditure weights to filter out price movements which are perceived by private households in a particular way. Instead, items are implicitly re-weighted, following the assumption that the number of representative products in a certain expenditure category coincides with the importance of that category for the formation of inflation sentiment.

In our analysis, we also probe the forecasting potential of two prominent core measures, the weighted median π_t^{WMed} and a 20%-trimmed mean π_t^{tr20}.[4] Specifically, we define $s_t^{WMed} = \pi_t^{WMed} - \pi_t$ and $s_t^{tr20} = \pi_t^{tr20} - \pi_t$. Table 2 reports correlation coefficients between each of the various sentiment indicators and the differences between these measures of core inflation and inflation, respectively. All these correlations are quite large, especially those involving the trimmed mean. This can be explained easily. In a situation in which inflation is triggered by only a few items, there is a high probability that they will be trimmed, i.e. excluded from the core measure. This is not likely to happen if inflation is instead supported by many product categories. All in all, sophisticated core measures such as the

[4] For the computation see e.g. Rich and Steindel (2005).

trimmed mean tend to display similar properties as our measures of inflation sentiment.

[Table 2: about here]

3. Estimation and forecast evaluation

In our empirical analysis, we forecast the difference between the actual inflation rate and the average inflation rate over the next *h* periods ($\pi_{t+h}^{h} - \pi_t$), where π_t is observed at time *t* and π_{t+h}^{h} is not.[5] Starting point of our analyses is the conventional Phillips curve, which constructs a relation between future inflation on the left hand side, a current and lagged first-differenced output variable *x* which represents the inflation pressure coming from the real part of the economy, and current and lagged inflation differences on the right hand side, [6]

$$\pi_{t+h}^{h} - \pi_t = \phi + \beta(L)\Delta x_t + \delta(L)\Delta \pi_t + \varepsilon_{t+h}\,. \qquad (3.1)$$

Relations of this type are widely used to forecast inflation. Stock and Watson (1999), evaluate the forecasting power of various specifications of x_t. In our paper we concentrate instead on evaluating alternative specifications of the inflation measure on the right hand side. While the standard Phillips curve specification uses the output or the employment gap, we capture influences from the real economy by first differences of the unemployment rate or real GDP. [7]

Concerning the inflation measure on the right hand side, we test one baseline and five alternative specifications. The first follows the standard Phillips curve literature and employs past inflation rates. The success of this approach serves as a benchmark for the other five specifications. Subsequently, we replace the current and lagged differences in inflation one by one by current and lagged values of our

---

[5] $\pi_t^h = \frac{4}{h}\log\left(\frac{P_t}{P_{t-h}}\right)$ is the h period inflation rate in the price level P_t reported at an annual rate.

[6] Δ and L represent the difference- and the lag-operator, respectively. Though modelling inflation as $I(1)$ is standard, it is, however, not always consistent with unit-root tests of the inflation series. Unit root-tests for all variables can be found in the appendix.

[7] According to our analyses, first differences of real variables generally perform better than deviations from HP-filtered variables. Two-sided filters additionally violate the out-of-sample assumptions because information is utilized which was not available at the point where the forecast is made. Alternatively, one might use one-sided filters or calculate quasi real-time output gaps by forecasting x first and applying a two-sided filter afterwards. Because of the poor results for the HP-filtered gaps we do not follow this route.

five candidate indicators of inflation sentiment, s_t^j, $j = Med$, $Diff$, Mom, $WMed$, $tr20$, which are described in section 2. Thus, equation 3.1 changes to

$$\pi_{t+h}^h - \pi_t = \phi^j + \beta^j(L)\Delta x_t + \delta^j(L)s_t^j + \varepsilon_{t+h}^j . \qquad (3.2)$$

In all estimates the lag length is chosen to minimize the *Schwartz* information criterion, respectively. This criterion has been used for model selection, since our simulations indicate that a parsimonious specification with relative small lag length produces the smallest out-of-sample forecast errors. The Schwartz criterion punishes additional coefficients more heavily than for instance the *Akaike* information criterion.

In-sample fit is not necessarily a good indicator of predictive power. Therefore, we evaluate the alternative specifications (3.2) on the basis of the out-of-sample forecast accuracy following Stock and Watson (1999). For that purpose, we generate a series of out-of-sample forecasts by estimating our equations for an expanding sample size and forecasting the average change in inflation over the next *h* periods for each of these samples, with *h* ranging from 1 to 8 for the US, and from 1 to 4 for Germany, respectively.

Thus, in any prediction we exclusively use the data available at the start of the respective forecast period. For instance, our first estimation for the US uses the sample 1978:1 to 1984:4 and forecasts inflation for *h* quarters starting with 1985:1. For the second estimate the sample is extended to 1978:1 to 1985:1 and a forecast is constructed of the average change in the annualized inflation rate for *h* quarters starting 1985:2. For Germany the initial sample is 1985:1 to 1991:4 for West Germany and 1993:1 to 1998:4 for re-unified Germany, respectively.

To evaluate the forecasts three tests are used. First of all, we calculate the *root mean squared forecast errors* (RMSFE) and use the Diebold-Mariano test to check whether the differences in the forecast accuracy of the various specifications are significant. Secondly, we employ an *encompassing test* to verify whether forecast generated by one specification adds information to the forecast generated by another, and thirdly, we test for a *forecast breakdown*, probing whether the out-of-sample accuracy differs significantly from the in-sample fit.

Differences in forecast accuracy

The RMSFE for each forecast, $\hat{\pi}_{t+h}^{jh}$, is defined as:

$$RMSFE_{jh} = \sqrt{\tfrac{1}{T}\sum_t \left(\hat{\pi}_{t+h}^{jh} - \pi_{t+h}^h\right)^2} = \sqrt{\tfrac{1}{T}\sum_t \left(e_{t+h}^{jh}\right)^2} , \qquad (3.3)$$

where superscript j denotes the candidate forecast model (with 0 indicating the benchmark), e_{t+h}^{jh} is the forecast error made by candidate forecast model j at time t for forecast horizon h, and T denotes the number of forecasts made. Subsequently, we report the *relative* RMSFE (for each horizon h) by dividing the respective RMSFE of each of our alternative specifications by the corresponding RMSFE of the benchmark. If the relative RMSFE is below 1, the alternative specification displays a better forecast performance than the benchmark. To test whether the differences are statistically significant, we employ a DM test (Diebold and Mariano 1995). This test is based on the null hypothesis that two non-nested series of forecasts $\left\{ f_t^{0h} \right\}_{t=1}^{T}$ and $\left\{ f_t^{jh} \right\}_{t=1}^{T}$ are of equal accuracy,

$$E\left(d_{t+h}^{jh}\right) = E\left[\left(e_{t+h}^{jh}\right)^2 - \left(e_{t+h}^{0h}\right)^2\right] = 0\,, \tag{3.4}$$

In this test, the loss function is the difference of the squared forecast errors of the candidate forecasts. Because the sample mean loss differential is asymptotically normally distributed, the large-sample DM test statistic is

$$DM_{jh} = \frac{\bar{d}^{\,jh}}{\sqrt{\hat{\gamma}_d^{jh} / T}}\,, \tag{3.5}$$

where $\bar{d}^{\,jh}$ is the sample mean loss differential and $\hat{\gamma}_d^{jh}$ is the cumulative sample autocovariance up to order h-1.

Encompassing test

Even if a forecast $\left\{ f_t^{jh} \right\}_{t=1}^{T}$ does not outperform the benchmark $\left\{ f_t^{0h} \right\}_{t=1}^{T}$, a combination of these two forecasts could nevertheless help to improve forecast accuracy. Therefore, we consider here the combined forecast $\left\{ f_t^{ch} \right\}_{t=1}^{T}$ estimating λ_{jh} as the corresponding “best” weight.

$$f_t^{ch} = \left(1 - \lambda_{jh}\right) f_t^{0h} + \lambda_{jh} f_t^{jh}\,. \tag{3.6}$$

If the null hypothesis $\lambda_{jh} = 0$ is true, $\left\{ f_t^{0h} \right\}_{t=1}^{T}$ is *conditionally efficient* with respect to $\left\{ f_t^{jh} \right\}_{t=1}^{T}$ (Granger and Newbold 1973; 1986) or *encompassing* $\left\{ f_t^{jh} \right\}_{t=1}^{T}$ (Hendry 1993). In this case, the DM-statistic can be calculated for each period t as

$$d_{t+h}^{jh} = \left(e_{t+h}^{0h} - e_{t+h}^{jh}\right) e_{t+h}^{0h}\,. \tag{3.7}$$

To achieve robust results, both tests described so far require large samples. However, Harvey et al. (1997) recommend a modified test statistic in small samples:

$$MDM_{jh} = \frac{\sqrt{N+1-2h+ h/N\left(h-1\right)}}{\sqrt{N}} DM_{jh} . \tag{3.8}$$

The critical values for this test are taken from the t_{N-1} distribution.

Forecast breakdown test

To evaluate the alternative inflation forecasts further, we also check the models for a *forecast breakdown* (FB). This is defined as a situation in which the out-of-sample forecasting performance of a forecast model is significantly worse than its in-sample fit (Giacomini and Rossi 2006). To implement this check we compare each model's forecasting performance – measured by its mean squared forecast error – to the expected forecast error based on its in-sample-fit.[8] Analytically, a "surprise loss" (sl) at time t is calculated as difference between the out-of-sample loss and the average in-sample loss $\overline{l}_t^{\,jh}$,

$$sl_{t+h}^{jh} = \left(e_{t+h}^{jh}\right)^2 - \overline{l}_t^{\,jh} . \tag{3.9}$$

If forecast model j is reliable, the mean of the associated surprise losses $\overline{sl}^{\,jh}$, taken over all T forecasts, should be close to zero. The standard normally distributed forecast breakdown test statistic is

$$FB_{jh} = \frac{\overline{sl}^{\,jh}}{\sqrt{\hat{\gamma}_{sl}^{jh}/T}} . \tag{3.10}$$

where $\hat{\gamma}_{sl}^{jh}$ is a Newey-West estimator of the variance of the weighted losses. Clearly, the precision of the estimate of the forecast model depends on the length of the sample that is used for estimation. The null hypothesis of a forecast breakdown is rejected at significance level α whenever the forecast breakdown test statistic is larger than the $(1-\alpha)$ -th quantile of a standard normal distribution.

[8] We only perform a one-sided test to reflect the assumption that a loss that is smaller than expected is desirable and therefore does not constitute a forecast breakdown. The forecasting scheme is recursive.

4. Results

4.1 United States

A relatively long time series is available for the US. During the sampling period, the US economy was relatively stable in terms of its underlying structure. The root mean squared forecast errors (RMSFE) of the standard Phillips curves serve as the relevant benchmark throughout the analysis[9]. In the estimates the real side of the US economy is either represented by changes in GDP or in the unemployment rate. Table 3 summarizes the RMSFE for the Phillips curve and compares it to the results of the five alternative specifications featuring inflation sentiment indicators instead of inflation. We can be confident that the results do not depend on the choice of the real side variable: Differences of the unemployment rate and changes in GDP generate more or less the same results concerning the relative RMSFE. In both cases, the s^{Med} - and the s^{Diff} -indicators improve forecast accuracy for all forecast horizons. For short horizons the improvement achieved is about 10% on average. For longer forecast horizons the improvement climbs up to 30%. The s^{Diff} -indicator even performs slightly better than the more established s^{Med} -measure. All differences to the benchmark forecasts are statistically significant, some of them at a 99% level. However, by replacing the inflation differences by the s^{Mom} -indicator, we get statistically worse results.

Since our inflation sentiment indicators are highly correlated with the differences between the two core inflation measures and headline inflation, the forecast potential of these core measures should be very similar to the proposed inflation sentiment indicators. From Table 3 it becomes obvious that the weighted median-indicator, s^{WMed} , performs worse than its un-weighted pendant up to a forecast horizon of roughly one year and better if longer forecast horizons are considered. The trimmed mean-indicator, s^{tr20} , displays high forecast accuracy in particular for longer forecast horizons, one the one hand. On the other hand, the differences in the RMSFE are only marginal and the DM and modified DM test statistics show lower absolute values than in the inflation sentiment indicator models. All in all, both concepts seem to have a similar forecasting power for US inflation.

[Table 3: about here]

Since forecast models using s^{Med} , s^{Diff} , s^{WMed} and s^{tr20} perform better than that using the inflation rate itself, the results of the encompassing tests are not surprising. Table 4 confirms that the null hypothesis that the standard forecasting model

---

[9] Additionally we tested whether the Phillips curve also can outperform a naïve forecast. For the U.S we found the Phillips curve showing smaller forecast errors than e.g. univariate forecasts or using the sample mean as predictor. In most cases, the DM and the MDM test show that the differences are significant. The results are available from the authors upon request.

encompasses a given alternative out of these four candidates is rejected for all forecast horizons at a 95% or 99% level. However, according to the encompassing tests we can confidently consider the standard model to encompass that using the s^{Mom} indicator.

[Table 4: about here]

So far, the inflation sentiment indicators we propose seem to be a useful forecasting tool. However, the question is whether this is true for all periods in our sample. This question is addressed by the forecast breakdown test. Of course, for all forecast horizons the fit is better in-sample than the out-of-sample. However, as shown in Table 5, forecast breakdowns occur in rare cases only. Models featuring inflation sentiment indicators are quite stable at commonly used significance levels. Further calculations show that the forecasts based on the sentiment indicators seem to be more stable than our benchmark, for which forecast breakdowns occur more frequently.

[Table 5: about here]

4.2 Germany

As it can be seen in Figure 1, inflation rates in Germany are less volatile compared to those of the US. This may be the main reason why the out-of-sample accuracy of our benchmark Phillips curve is better in the German case. This holds for both samples considered and for both specifications of the real side of the economy.[10] Nevertheless, taking point estimates of relative RMSFE at face value, also for West Germany most alternative Phillips curve specifications outperform our benchmark in the period from 1985 to 1998 (Table 6). When GDP growth is chosen as real economy variable, the alternative inflation indicators even improve forecast accuracy for all forecast horizons and all information criteria. As an example, consider a one-year-ahead-forecast: Here, the RMSFE can be reduced up to one half when replacing inflation by the s^{Diff}-indicator in the Phillips curve. However, the accuracy gains are only significant for all indicators in the case of three-quarter-ahead-forecasts, and in many cases when one-year-ahead predictions are constructed.

[Table 6: about here]

As in the case of the US, the s^{Med}-indicator and the s^{Diff}-indicator apparently tend to produce better forecasts than the s^{Mom}-indicator. Indicators which are

[10] On the other hand, the low volatility in inflation rates makes it much harder to beat a naïve forecast.

based on one of the two core inflation measures improve forecast accuracy, too, but (in terms of point estimates) they seem to be inferior to the s^{Med} - and the s^{Diff} -indicators in most cases. Table 6 also documents that, issues of statistical significance aside, all indicators perform better in combination with GDP growth than with changes in unemployment.

The second sub-period relates to unified Germany. We should not be too surprised, if results turn out to be different. On the one hand, it is quite an open question, whether the inflation sentiment among East Germans and those among West Germans, respectively, coincide, and what might happen as they become entangled with one another. On the other hand, much of the sampling period was relatively stable with respect to inflation and, thus, intuitively it should be difficult to outperform the standard Phillips curve approach. Indeed, none of our candidate forecast models is performing significantly better than the standard model. By contrast, for several forecast horizons some of the candidates, in particular that employing s^{Mom} perform significantly worse. This holds especially for models involving the unemployment rate as a predictor.

Again, we also employ encompassing tests to assess whether our candidate variables make a contribution to improved inflation forecasts. Since according to our point estimates, all indicators perform better than the inflation rate in the first subsample, the tests might state that these variables add significant information. Table 7 documents that the null hypothesis is rejected for a large share of the candidates at all forecast horizons. The associated significance levels are particularly high with regard to three-quarter-ahead forecasts. Less uniform are the results for the more recent sub-period, in which inflation sentiment indicators do not outperform the standard Phillips curve approach in most cases.

Yet, the encompassing tests show that in some cases inflation sentiment indicators can add forecast-relevant information, in particular the s^{Med} - and the s^{Diff} -indicator. Again, the alternative indicators lose their appeal when they are combined with unemployment instead of GDP growth. In line with the forecast results, the s^{Mom} -indicator and the core measures do not add significant forecast-relevant information to the inflation rate, the combination of the s^{tr20} -indicator and GDP being an exception. All in all, also in periods where inflation is quite stable, simple inflation sentiment indicators might improve forecast accuracy, whereas this does not seem to be the case for familiar core measures of inflation.

[Table 7: about here]

Since the inflationary regime in Germany seems to differ between the (partly overlapping) sub-periods, the forecast breakdown can be expected to provide interesting insights on the out-of-sample performance of the candidate forecast models (Table 8). For the first German sample, the FB test statistics exhibit relative large negative values, i.e. the null hypothesis that the out-of-sample errors are not worse than those in sample is not rejected. This is particularly true for the s^{Med} - and the s^{Diff} -indicator. For the second sub-period, however, FB tests indicate forecast breakdowns for the s^{Med} -, the s^{Diff} - and the s^{tr20} -indicator-based models. This evidence is consistent with our finding that one might fruitfully employ sentiment indicators for the purpose of forecasting inflation, emphasizing that their contribution will tend to be smaller in stable times.

[Table 8: about here]

5. Conclusions

In this paper we construct several indicators capturing whether a given inflation rate is the result of price increases for many components in the CPI basket or rather the consequence of price hikes for a relatively small number of goods and services with a high weight in the basket. Since inflation is supposed to be perceived more intensively in the first case, we label our indicators *inflation sentiment*. We also demonstrate that simple sentiment indicators are highly correlated with differences of familiar core measures and headline inflation.

With regard to the US we find that inflation sentiment indicators tend to improve the accuracy of inflation forecasts as measured by the RMSFE by roughly 20%, compared to a standard Phillips curve approach. The differences are significant according to the Diebold-Mariano and modified Diebold-Mariano tests for forecast horizons up to eight quarters. Here, indicators based on familiar core measures show forecast accuracy similar to our new indicators. Furthermore, a forecast breakdown test indicates that the out-of-sample forecast errors of the alternative candidate models do not deviate significantly from their in-sample fit, suggesting that the forecasts based on our indicators are more stable than the standard Phillips curve approach.

The results derived on the basis of German data are less uniform. Their heterogeneity is indicative for the role of circumstances: inflation sentiment indicator-based forecasts seem to be particularly powerful, if inflation is volatile as it has been the case in the 1985:1 to 1998:4 sample of West German data. The RMSFE is reduced up to an half. In the 1993:1 to 2007:4 period, tough, when inflation was quite stationary, the sentiment indicators do not outperform Phillips curve-based forecasts. However, also under these circumstances encompassing tests suggest that our inflation sentiment indicators add valuable information. For this German data the performance of core inflation-based indicators appears to be worse than that of the new indicators. Since the latter require less data – neither explicit weights nor subjective judgements concerning the products to be excluded from the analysis are needed – they offer a helpful and simple alternative to measures of core inflation.

References

Brachinger, H. W. (2006), Euro or "Teuro"?: The Euro-induced Perceived Inflation in Germany. Department of Quantitative Economics Working Paper No. 5, Freiburg/Fribourg.

Bryan, M. F. and S. G. Cechetti (1994), Measuring Core Inflation. In: *Monetary Policy*, edited by N. G. Mankiw: 195-215. Chicago: University of Chicago Press.

Burns, A. F. and W. C. Mitchell (1946), *Measuring Business Cycles*. New York.

Carstensen, K. (2007). Is Core Money Growth a Good and Stable Inflation Predictor in the Euro Area? Kiel Institute for the World Economy Working Paper No. 1318, Kiel.

Diebold, F. X. and R. S. Mariano (1995), Comparing Predictive Accuracy. *Journal of Business & Economic Statistics* **13** (3): 253-63.

Giacomini, R. and B. Rossi (2006), Detecting and Predicting Forecast Breakdowns. European Central Bank Working Paper No. 638, Frankfurt.

Gordon, R. J. (1975), The Impact of Aggregate Demand on Prices. *Brookings Papers* **3**: 613-70.

Granger, C.W. J. and P. Newbold (1973), Some Comments on the Evaluation of Economic Forecasts. *Applied Economics* **5**: 35-47.

Granger, C. W. J. and P. Newbold (1986), *Forecasting Economic Time Series*. San Diego.

Harvey, D. I., S. J. Leybourne and P. Newbold (1998), Tests for Forecast Encompassing. *Journal of Business & Economic Statistics* **16** (2): 254-59.

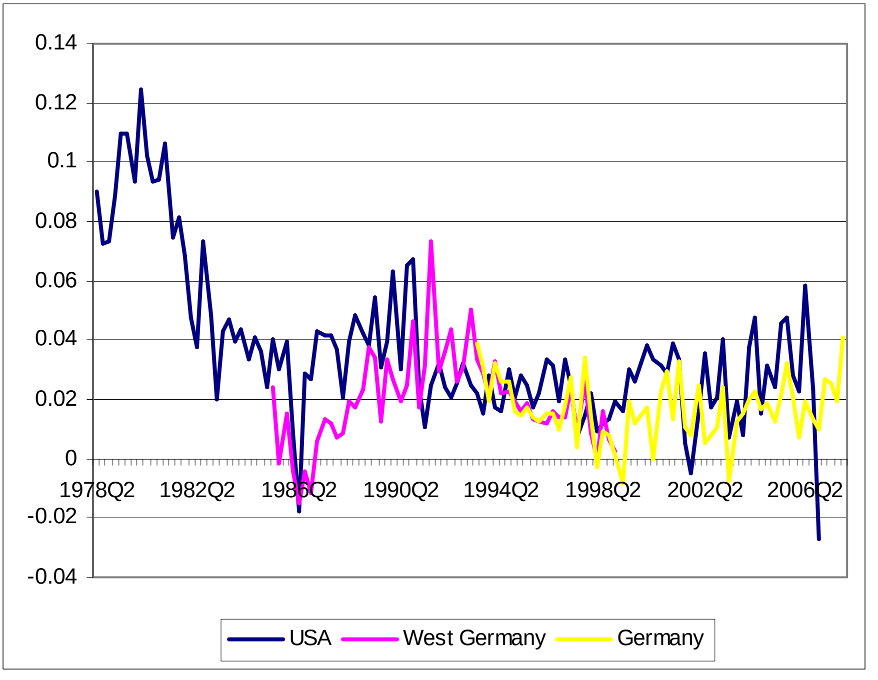
Reed, S. B. and K. J. Stewart (1999), *Consumer Price Index Research Series Using Current Methods, 1978-98*. *Monthly Labor Review* **122** (6): 29-38

Rich, R. and C. Steindel (2005), A Review of Core Inflation and an Evaluation of its Measures. Federal Reserve Bank of New York Staff Report no. 236, New York.

Smith, J. K. (2004), Weighted Median Inflation: Is This Core Inflation? *Journal of Money, Credit and Banking* **36** (2): 253-63.

Stock, J. H. and M. W. Watson (1999), Forecasting inflation. *Journal of Monetary Economics* **44** (2): 293-335.

Figure 1: Quarterly inflation, USA and Germany



Notes: The quarterly inflation rates are averages of seasonally adjusted monthly data.

Table 1
Structure of the German CPI 2000=100

COICOP-Code	Product category	Share in CPI basket (‰)	Number of products[1]
01	Food and non-alcoholic beverages	103.65	11
02	Alcoholic beverages, tobacco and narcotics	36.73	5
03	Clothing and footwear	55.09	6
04	Housing, water, electricity, gas and other fuels	302.66	12
05	Furnishings, household equipment and routine household maintenance	68.54	12
06	Health	35.46	7
07	Transport	138.65	13
08	Communication	25.21	3
09	Recreation and culture	110.85	20
10	Education	6.66	3
11	Restaurants and hotels	46.57	3
12	Miscellaneous goods and services	70.23	12
01-12	Individual consumption expenditure of households	1000.00	109

Source: Destatis. [1]on 4-digit COICOP-level.

Table 2
Correlation of inflation sentiment indicators and selected core measures

USA (1978:1-2006:4)					
	s^{Med}	s^{Diff}	s^{Mom}	s^{WMed}	s^{tr20}
s^{Med}	1.00	0.97	0.09	0.85	0.95
s^{Diff}		1.00	0.09	0.82	0.93
s^{Mom}			1.00	-0.03	0.02
s^{WMed}				1.00	0.91
s^{tr20}					1.00
West Germany (1985:1 – 1998:4)					
	s^{Med}	s^{Diff}	s^{Mom}	s^{WMed}	s^{tr20}
s^{Med}	1.00	0.89	-0.29	0.84	0.82
s^{Diff}		1.00	-0.17	0.79	0.74
s^{Mom}			1.00	-0.11	-0.14
s^{WMed}				1.00	0.94
s^{tr20}					1.00
Germany (1993:1 – 2007:4)					
	s^{Med}	s^{Diff}	s^{Mom}	s^{WMed}	s^{tr20}
s^{Med}	1.00	0.97	-0.26	0.88	0.94
s^{Diff}		1.00	-0.25	0.87	0.91
s^{Mom}			1.00	-0.38	-0.29
s^{WMed}				1.00	0.94
s^{tr20}					1.00

Table 3
Forecast accuracy of alternative Phillips curves, USA, 1978:1 – 2006:4

		x = GDP					x = unemployment rate				
h		s^{Med}	s^{Diff}	s^{Mom}	s^{WMed}	s^{tr20}	s^{Med}	s^{Diff}	s^{Mom}	s^{WMed}	s^{tr20}
1	RMSFE	**0.91**	**0.92**	1.08	**0.93**	**0.94**	**0.92**	**0.93**	1.11	**0.94**	**0.96**
	DM	-1.80**	-1.53*	1.17	-1.17	-1.28	-1.55*	-1.30*	1.66**	-0.99	-0.88
	MDM	-1.79**	-1.52*	1.16	-1.16	-1.27	-1.54*	-1.29	1.65*	-0.98	-0.87
2	RMSFE	**0.87**	**0.86**	1.18	**0.92**	**0.85**	**0.85**	**0.85**	1.18	**0.91**	**0.84**
	DM	-1.47*	-1.51*	2.17**	-1.24	-1.51*	-1.74**	-1.78**	2.39***	-1.18	-1.73**
	MDM	-1.44*	-1.48*	2.13**	-1.22	-1.48*	-1.71**	-1.75**	2.35**	-1.16	-1.70**
3	RMSFE	**0.85**	**0.85**	1.16	**0.93**	**0.84**	**0.83**	**0.83**	1.15	**0.91**	**0.85**
	DM	-1.67**	-1.69**	2.56***	-1.09	-1.82**	-1.99**	-1.98**	2.47***	-1.37*	-2.31**
	MDM	-1.62*	-1.64*	2.49***	-1.06	-1.77**	-1.93**	-1.92**	2.40***	-1.33*	-2.24**
4	RMSFE	**0.83**	**0.84**	1.18	**0.86**	**0.82**	**0.86**	**0.86**	1.24	**0.89**	**0.85**
	DM	-2.01**	-2.07**	3.35***	-1.67**	-2.02**	-1.61*	-1.66**	2.95***	-1.35*	-1.62*
	MDM	-1.93**	-1.98**	3.21***	-1.60*	-1.94**	-1.54*	-1.59*	2.83***	-1.29	-1.55*
5	RMSFE	**0.82**	**0.82**	1.18	**0.82**	**0.78**	**0.81**	**0.81**	1.19	**0.81**	**0.77**
	DM	-1.99**	-2.25**	2.63***	-2.02**	-2.18**	-2.25**	-2.53***	2.65***	-2.36***	-2.41***
	MDM	-1.88**	-2.13**	2.49***	-1.91**	-2.06**	-2.13**	-2.39***	2.51***	-2.23**	-2.28**
6	RMSFE	**0.84**	**0.83**	1.18	**0.83**	**0.79**	**0.81**	**0.80**	1.18	**0.79**	**0.74**
	DM	-2.37***	-2.85***	3.11***	-1.89**	-2.32**	-2.20**	-2.46***	2.33***	-2.16**	-2.40***
	MDM	-2.21**	-2.66***	2.90***	-1.76**	-2.17**	-2.05**	-2.30**	2.18**	-2.02**	-2.24**
7	RMSFE	**0.83**	**0.81**	1.17	**0.81**	**0.77**	**0.76**	**0.75**	1.12	**0.73**	**0.69**
	DM	-2.66***	-3.83***	3.46***	-1.92**	-2.21**	-2.49***	-2.73***	1.97**	-2.42***	-2.61***
	MDM	-2.45***	-3.53***	3.19***	-1.77**	-2.03**	-2.29**	-2.51***	1.81**	-2.23**	-2.40***
8	RMSFE	**0.80**	**0.79**	1.13	**0.76**	**0.75**	**0.74**	**0.72**	1.09	**0.68**	**0.65**
	DM	-2.56***	-2.95***	2.10**	-2.18**	-1.95**	-2.61***	-2.75***	1.24	-2.92***	-2.76***
	MDM	-2.32**	-2.68***	1.91**	-1.98**	-1.77**	-2.37**	-2.50***	1.13	-2.65***	-2.50***

Notes: The Table reports the RMSFE relative to the Phillips curve benchmark model. Relative RMFSEs smaller than 1 are documented in bold print. The lag length is set to minimize the Schwarz-information criteria. (M)DM indicates the (modified) Diebold-Mariano test statistic. *** (**) (*) denotes significance at the 0.99 (0.95) (0.90) level.

Table 4
Encompassing test, USA, 1978:1-2006:4

		x = GDP					x = unemployment rate				
h		s^{Med}	s^{Diff}	s^{Mom}	s^{WMed}	s^{tr20}	s^{Med}	s^{Diff}	s^{Mom}	s^{WMed}	s^{tr20}
1	DM	2.38***	2.24**	0.61	2.27**	2.01**	2.36***	2.21**	0.29	2.28**	1.87**
	MDM	2.37**	2.23**	0.61	2.26**	2.00**	2.35**	2.20**	0.29	2.27**	1.86**
2	DM	1.91**	1.95**	0.13	2.43***	1.95**	2.16**	2.18**	-0.01	2.33***	2.15**
	MDM	1.88**	1.92**	0.13	2.39***	1.92**	2.12**	2.14**	-0.01	2.29**	2.11**
3	DM	2.11**	2.25**	0.09	2.66***	2.44***	2.42***	2.46***	0.22	2.37***	3.15***
	MDM	2.05**	2.18**	0.09	2.58***	2.37**	2.35**	2.39***	0.21	2.30**	3.06***
4	DM	2.45***	2.66***	-0.07	2.87***	2.71***	2.17**	2.31**	-0.86	2.29**	2.31**
	MDM	2.35**	2.55***	-0.07	2.75***	2.60***	2.08**	2.21**	-0.82	2.20**	2.21**
5	DM	2.39***	2.75***	0.16	3.09***	2.68***	2.63***	2.94***	0.18	3.00***	2.92***
	MDM	2.26**	2.60***	0.15	2.92***	2.54***	2.49***	2.78***	0.17	2.84***	2.76***
6	DM	2.81***	3.39***	0.22	3.71***	3.21***	2.51***	2.78***	0.39	2.80***	2.67***
	MDM	2.62***	3.17***	0.21	3.46***	3.00***	2.34**	2.60***	0.36	2.61***	2.49***
7	DM	3.12***	4.09***	0.08	3.72***	3.24***	2.64***	2.85***	1.02	3.02***	2.73***
	MDM	2.87***	3.77***	0.07	3.43***	2.98***	2.43***	2.62***	0.94	2.78***	2.51***
8	DM	2.75***	3.21***	0.52	4.05***	2.96***	2.78***	2.92***	1.05	3.67***	2.94***
	MDM	2.50***	2.91***	0.47	3.67***	2.69***	2.52***	2.65***	0.95	3.33***	2.67***

Notes: The lag length is set to minimize the Schwarz-information criteria. (M)DM indicates the (modified) Diebold-Mariano test statistic. *** (**) (*) denotes significance at the 0.99 (0.95) (0.90) level.

Table 5
Results of forecast breakdown tests, USA, 1978:1-2006:4

	x = GDP					x = unemployment rate				
h	s^{Med}	s^{Diff}	s^{Mom}	s^{WMed}	s^{tr20}	s^{Med}	s^{Diff}	s^{Mom}	s^{WMed}	s^{tr20}
1	1.00	1.50*	1.20*	1.40*	1.50*	1.20	1.40*	1.20	1.50*	1.70**
2	0.48	0.99	1.20*	1.00	1.10	0.77	0.95	1.30*	1.10	1.10
3	0.13	0.40	0.74	0.64	0.65	0.58	0.54	0.83	1.00	1.00
4	0.59	0.69	1.10	0.76	0.95	1.10	0.84	0.91	1.10	1.10
5	0.39	0.33	0.84	0.59	0.66	1.30	0.99	0.97	1.20	1.20
6	0.56	0.60	0.82	0.48	0.93	1.20	1.10	0.97	0.98	1.20
7	0.44	0.71	0.93	0.24	0.88	1.20	0.83	0.85	0.85	1.30
8	0.63	0.81	0.81	0.36	0.95	0.98	0.33	0.89	0.57	1.20

Notes: The Table reports the FB-test statistics. *** (**) (*) denotes significance at the 0.99 (0.95) (0.90) level. Throughout the estimation the lag length is fixed to 4.

Table 6
Forecast accuracy of alternative Phillips curves, Germany

		West Germany, 1985:1 – 1998:4									
		x = GDP					x = unemployment rate				
h		s^{Med}	s^{Diff}	s^{Mom}	s^{WMed}	s^{tr20}	s^{Med}	s^{Diff}	s^{Mom}	s^{WMed}	s^{tr20}
1	RMSFE	**0.85**	**0.85**	**0.94**	**0.94**	**0.97**	**0.99**	1.00	1.01	1.05	1.06
	DM	-0.55	-0.55	-0.25	-0.21	-0.10	-0.05	-0.02	0.05	0.22	0.29
	MDM	-0.54	-0.54	-0.25	-0.21	-0.10	-0.05	-0.02	0.05	0.22	0.28
2	RMSFE	**0.77**	**0.70**	**0.94**	**0.98**	**0.95**	**0.96**	**0.89**	1.00	1.01	1.05
	DM	-1.20	-1.10	-0.35	-0.11	-0.23	-0.20	-0.36	0.02	0.02	0.17
	MDM	-1.13	-1.04	-0.33	-0.10	-0.22	-0.19	-0.34	0.02	0.02	0.16
3	RMSFE	**0.77**	**0.70**	**0.86**	**0.85**	**0.82**	**0.90**	**0.88**	**0.89**	**0.90**	**0.93**
	DM	-1.88**	-1.75**	-2.98***	-3.17***	-1.69**	-0.51	-0.55	-0.83	-0.59	-0.32
	MDM	-1.70*	-1.58*	-2.69***	-2.86***	-1.53*	-0.46	-0.50	-0.75	-0.53	-0.29
4	RMSFE	**0.52**	**0.50**	**0.85**	**0.82**	**0.63**	**0.98**	**0.75**	**0.78**	**0.77**	**0.84**
	DM	-2.23**	-1.78**	-0.83	-1.23	-2.23**	NA	-1.50*	-1.58*	-1.79**	-1.40*
	MDM	-1.92**	-1.53*	-0.71	-1.06	-1.92**	0.00	-1.29	-1.36*	-1.54*	-1.20

		Germany, 1993:1 – 2007:4									
		x = GDP					x = unemployment rate				
h		s^{Med}	s^{Diff}	s^{Mom}	s^{WMed}	s^{tr20}	s^{Med}	s^{Diff}	s^{Mom}	s^{WMed}	s^{tr20}
1	RMSFE	**0.95**	1.01	1.18	1.02	1.02	1.06	1.19	1.24	1.16	1.09
	DM	-0.65	0.05	1.89**	0.22	0.25	0.73	1.93**	2.60***	1.63*	0.96
	MDM	-0.64	0.05	1.86**	0.22	0.25	0.72	1.90**	2.56***	1.61*	0.95
2	RMSFE	**0.86**	1.05	1.21	1.01	1.01	1.06	1.16	1.27	1.26	1.14
	DM	-1.05	0.51	1.37*	0.05	0.08	0.53	2.97***	1.87**	3.05***	1.20
	MDM	-1.00	0.49	1.31*	0.05	0.08	0.51	2.84***	1.79**	2.92***	1.15
3	RMSFE	**0.85**	1.04	1.20	**0.96**	1.00	1.06	1.21	1.18	1.22	1.13
	DM	-0.85	0.28	1.70**	-0.24	0.02	0.34	1.45*	1.28	1.26	0.74
	MDM	-0.79	0.26	1.57*	-0.22	0.02	0.31	1.34*	1.19	1.17	0.69
4	RMSFE	**0.84**	**0.97**	1.23	**0.95**	1.04	1.12	1.24	1.28	1.29	1.18
	DM	-0.80	-0.15	2.60***	-0.25	0.28	0.63	1.47*	2.39***	1.50*	1.17
	MDM	-0.72	-0.13	2.32**	-0.22	0.25	0.56	1.31*	2.14**	1.34*	1.05

Notes: The Table reports the RMSFE relative to the Phillips curve benchmark model. Relative RMSFEs smaller than 1 are documented in bold print. The lag length is set to minimize the Schwarz-information criteria. (M)DM indicates the (modified) Diebold-Mariano test statistic. *** (**) (*) denotes significance at the 0.99 (0.95) (0.90) level.

Table 7
Encompassing tests, Germany

West Germany, 1985:1 – 1998:4											
		x = GDP					x = unemployment rate				
h		s^{Med}	s^{Diff}	s^{Mom}	s^{WMed}	s^{tr20}	s^{Med}	s^{Diff}	s^{Mom}	s^{WMed}	s^{tr20}
1	DM	1.38*	1.45*	1.06	1.51*	1.34*	1.30*	1.29*	1.08	1.34*	1.27
	MDM	1.36*	1.42*	1.04	1.48*	1.32*	1.28	1.27	1.06	1.32*	1.25
2	DM	1.78**	1.92**	1.17	2.08**	2.00**	1.34*	1.51*	1.08	1.59*	1.53*
	MDM	1.68*	1.81**	1.10	1.96**	1.89**	1.27	1.43*	1.02	1.50*	1.44*
3	DM	2.86***	3.29***	2.43***	3.60***	3.12***	2.23**	2.72***	2.68***	3.17***	2.68***
	MDM	2.58***	2.97***	2.20**	3.25***	2.82***	2.02**	2.46**	2.42**	2.86***	2.42**
4	DM	2.03**	2.20**	1.18	2.43***	2.10**	1.47*	1.67**	1.44*	1.76**	1.65**
	MDM	1.75**	1.89**	1.01	2.09**	1.81**	1.26	1.44*	1.24	1.51*	1.42*
Germany, 1993:1 – 1998:4											
		x = GDP					x = unemployment rate				
h		s^{Med}	s^{Diff}	s^{Mom}	s^{WMed}	s^{tr20}	s^{Med}	s^{Diff}	s^{Mom}	s^{WMed}	s^{tr20}
1	DM	2.77***	1.90**	-0.20	0.95	2.01**	1.18	0.00	-0.73	-0.01	1.18
	MDM	2.73***	1.87**	-0.20	0.94	1.98**	1.16	0.00	-0.72	-0.01	1.16
2	DM	2.73***	2.35***	0.16	1.26	2.41***	2.15**	1.37*	0.00	-0.85	1.63*
	MDM	2.61***	2.25**	0.15	1.21	2.31**	2.06**	1.31*	0.00	-0.81	1.56*
3	DM	2.07**	1.77**	0.18	1.03	2.01**	1.47*	0.52	0.66	0.09	1.29
	MDM	1.92**	1.64**	0.17	0.95	1.86**	1.36*	0.48	0.61	0.08	1.19
4	DM	2.05**	1.43*	-0.08	0.86	2.22**	1.32*	1.05	-0.13	0.07	1.37*
	MDM	1.83**	1.28	-0.07	0.77	1.98**	1.18	0.94	-0.12	0.06	1.22

Notes: The lag length is set to minimize the Schwarz-information criteria. (M)DM indicates the (modified) Diebold-Mariano test statistic. *** (**) (*) denotes significance at the 0.99 (0.95) (0.90) level.

Table 8
Forecast breakdown tests, Germany

West Germany, 1985:1 – 1998:4										
	x = GDP					x = unemployment rate				
h	s^{Med}	s^{Diff}	s^{Mom}	s^{WMed}	s^{tr20}	s^{Med}	s^{Diff}	s^{Mom}	s^{WMed}	s^{tr20}
1	-2.12	-2.09	-1.70	-1.63	-1.38	-1.51	-1.52	-1.43	-0.82	-0.76
2	-2.54	-2.96	-2.31	-1.25	-1.39	-1.80	-2.11	-2.27	-0.45	-0.89
3	-2.48	-3.12	-2.54	-1.45	-1.62	-1.47	-1.96	-2.44	-0.43	-1.03
4	-3.48	-3.25	-1.71	-1.99	-2.16	-2.24	-2.11	-1.93	-0.80	-1.42
Germany, 1993:1 – 2007:4										
	x = GDP					x = unemployment rate				
h	s^{Med}	s^{Diff}	s^{Mom}	s^{WMed}	s^{tr20}	s^{Med}	s^{Diff}	s^{Mom}	s^{WMed}	s^{tr20}
1	1.6*	2**	0.5	1.60*	2.00**	2.2**	2**	0.91	1.60*	2.10**
2	1.2	1.9**	0.67	1.50*	2.30**	2.1**	1.7**	0.69	1.30*	2.30**
3	1.2	0.82	0.6	0.78	2.00**	2.1**	0.86	0.35	0.79	2.20**
4	1.8**	1	0.54	0.91	2.80***	3***	1.4*	0.28	0.99	2.90***

Notes: The Table reports the FB-test statistics. *** (**) (*) denotes significance at the 0.99 (0.95) (0.90) level. Throughout the estimation the lag length is fixed to 2.

Appendix A: Unit root tests

Table A1
Unit root tests, USA (1978:1-2006:4)

Series	Augmented Dickey- Fuller		KPSS	
	Levels	First differences	Levels	First differences
GDP	-3.07	-4.07***	0.05	0.04
U	-3.74**	-4.68***	0.06	0.07
π^1	-2.51	-12.41***	0.19**	0.07
π^2	-3.95**	-2.63*	0.20**	0.05
π^3	-4.04**	-4.09***	0.20**	0.08
π^4	-3.45*	-3.30**	0.20**	0.13
π^5	-3.78**	-4.18***	0.20**	0.14
π^6	-4.36***	-2.48	0.20**	0.20
π^7	-3.57**	-3.23**	0.20**	0.25
π^8	-2.88	-3.70***	0.20**	0.31
s^{Med}	-4.79***	-5.64***	0.11	0.17
s^{Diff}	-7.56***	-5.59***	0.11	0.13
s^{Mom}	-4.06***	-10.49***	0.06	0.13
s^{WMed}	-2.83	-5.37***	0.09	0.08
s^{tr20}	-4.30***	-5.33***	0.14*	0.03

Notes: For the ADF tests, the lag length is automatically selected based on AIC. The only exogenous variable in the first-differenced model is a constant, while the test in levels also includes a time trend. The null hypothesis states that the time series has a unit root. The test statistics are compared with MacKinnon (1996) one-sided p-values. *** (**) (*) denotes rejection of the null hypothesis at the 0.99 (0.95) (0.90) level.
For the KPSS tests, the bandwidth is automatically selected based on a Newey-West using a Bartlett kernel. The only exogenous variable in the first-differenced model is a constant, while the test in levels also includes a time trend. The null hypothesis states that the time series is (trend-) stationary. The test statistics are compared with the critical values from Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1). *** (**) (*) denotes rejection of the null hypothesis at the 0.99 (0.95) (0.90) level.

Table A2
Augmented Dickey-Fuller unit root tests, Germany

Series	West Germany, 1985:1-1998:4		Germany, 1993:1-2004:4	
	Levels	First differences	Levels	First differences
GDP	-1.49	-2.35	-1.92	-7.30***
U	-2.35	-2.44	-1.68	-3.60***
π^1	-1.05	-10.74***	-3.57**	-11.06***
π^2	-1.50	-4.54***	-2.28	-7.43***
π^3	-0.87	-4.80***	-1.72	-2.60*
π^4	-1.91	-1.27	-2.44	-3.42**
s^{Med}	-1.51	-11.05***	-6.81***	-10.79***
s^{Diff}	-1.17	-9.83***	-3.88**	-10.72***
s^{Mom}	-7.26***	-2.51	-10.43***	-6.15***
s^{WMed}	-5.94***	-9.05***	-7.05***	-11.99***
s^{tr20}	-5.82***	-8.63***	-7.26**	-11.08***

Notes: The lag length is automatically selected based on AIC. The only exogenous variable in the first-differenced model is a constant, while the test in levels also includes a time trend. The null hypothesis states that the time series has a unit root. The test statistics are compared with MacKinnon (1996) one-sided p-values. *** (**) (*) denotes rejection of the null hypothesis at the 0.99 (0.95) (0.90) level.

Table A3:
KPSS unit root tests, Germany

Series	West Germany, 1985:1-1998:4		Germany, 1993:1-2004:4	
	Levels	First differences	Levels	First difference
GDP	0.19**	0.24	0.12*	0.11
U	0.20**	0.25	0.16**	0.43*
π^1	0.23***	0.12	0.25***	0.11
π^2	0.23***	0.17	0.25***	0.42*
π^3	0.23***	0.29	0.25***	0.53**
π^4	0.23***	0.29	0.22***	0.51**
s^{Med}	0.18**	0.09	0.17**	0.06
s^{Diff}	0.14*	0.12	0.10	0.04
s^{Mom}	0.10	0.05	0.03	0.11
s^{WMed}	0.15**	0.06	0.19**	0.08
s^{tr20}	0.16**	0.05	0.19**	0.06

Notes: The bandwidth is automatically selected based on a Newey-West using a Bartlett kernel. The only exogenous variable in the first-differenced model is a constant, while the test in levels also includes a time trend. The null hypothesis states that the time series is (trend-) stationary. The test statistics are compared with the critical values from Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1). *** (**) (*) denotes rejection of the null hypothesis at the 0.99 (0.95) (0.90) level.