# MACROECONOMIC PREDICTIONS – THREE ESSAYS ON ANALYSTS' FORECAST QUALITY

Inauguraldissertation
zur
Erlangung des Doktorgrades
der
Wirtschafts- und Sozialwissenschaftlichen Fakultät
der
Universität zu Köln

2013

vorgelegt

von

Diplom-Kaufmann Sebastian Orbe

aus

Leverkusen

Referent: Univ.-Prof. Dr. Dieter Hess

Korreferent: Univ.-Prof. Dr. Alexander Kempf

Tag der Promotion: 30.04.2013

# Contents

Chapter 2 is based on Hess and Orbe (2013), Irrationality or Efficiency of Macroeconomic Survey Forecasts? Implications from the Anchoring Bias Test, Review of Finance, Advance Access published January 12, 2013, doi:10.1093/rof/rfs037. Chapter 3 is based on Hess and Orbe (2013), Macroeconomic Forecast Accuracy – Idiosyncratic Predictability and Analyst Characteristics, working paper. Chapter 4 is based on Orbe and Pucker (2013), Do Aggregate Company Outlooks have Macroeconomic Content?, working paper.

# List of Tables

# List of Figures

# Chapter 1 Introduction

Macroeconomic expectation data are of great interest to different agents due to their importance as central input factors in various applications. To name but a few, politicians, capital market participants, as well as academics, incorporate these forecast data into their decision processes. Consequently, a sound understanding of the quality properties of macroeconomic forecast data, their quality determinants, as well as potential ways to improve macroeconomic predictions is desirable.

This thesis consists of three essays on the quality of analysts' forecasts. The first essay deals with macroeconomic forecast quality on the consensus level, while the second one investigates individual analysts' predictions and their quality determinants. In the third essay a bottom-up approach is introduced to derive macroeconomic forecasts from analysts' predictions at the microeconomic level.

It is generally assumed that macroeconomic consensus forecasts provide a reasonable approximation of market participants' expectations regarding upcoming macroeconomic releases. Research areas in which these expectation data are a central input to isolate the unanticipated news component of a given announcement include studies analyzing the price impact of macroeconomic news in bond markets (e.g., Balduzzi et al., 2001; Gilbert et al., 2010), stock markets (e.g., Boyd et al., 2005; Cenesizoglu, 2011) as well as in foreign exchange markets (e.g., Andersen et al., 2003; Evans and Lyons, 2008). Furthermore, these forecast data are used to study market co-movement (e.g., Albuquerque and Vega, 2009), market volatility (e.g., Beber and Brandt, 2008; Brenner et al., 2009), changes in market liquidity (e.g., Brandt and Kavajecz, 2004; Pasquariello and Vega, 2007, 2009) as well as bond and equity risk premiums (e.g., Savor and Wilson, 2012; Dicke and Hess, 2012).

It appears reasonable to assume that macroeconomic consensus forecasts represent market participants' expectations properly. So far available studies on forecast rationality at the consensus level largely test for general quality properties.[1] They commonly find no evidence of systematic or persistent inefficiencies.[2] In contrast to these previous studies, Campbell and Sharpe (2009) test for a specific behavioral inefficiency, the anchoring bias, first documented by Tversky and Kahneman (1974) in psychological experiments. Transferred to the context of macroeconomic forecasts, anchoring means that analysts put too much importance on last months' data and therefore underweight meanwhile released relevant information. This behavior implies a false incorporation of all available information into their forecasts. Consequently, a correction, i.e., the efficient use of the entire available information set would yield forecast improvements.

Our analysis reveals a counter-intuitive result: We find strong statistical significance for anchoring in most macroeconomic forecast series, but applying a look-ahead bias free estimation and adjustment procedure leads to no systematic forecast improvements. Therefore, our results question the economical significance of the anchoring bias. To provide an explanation for the disconnection of statistical and economical significance, we decompose the anchoring bias test statistic and find that the test is biased itself. While the test assumes a univariate information environment, it neglects the possibility that analysts may provide superior forecasts by using a more comprehensive information set than just the univariate time series itself. Our empirical as well as our simulation results strongly support this explanation for a broad range of macroeconomic series.

[1] See e.g., Pesando (1975), Mullineaux (1978), Pearce and Roley (1985), Aggarwal et al. (1995), and Schirm (2003).

[2] The most recent study, Schirm (2003), only finds for a small number of investigated series some bias. However, his results partly contradict the findings of Aggarwal et al. (1995) obtained on a different sample.

Our analysis contributes to different strands of literature. First, our results directly add to the scarce literature analyzing the efficiency of macroeconomic survey forecasts by showing that informational advantages of analysts, i.e., the incorporation of related macroeconomic data, enable them to outperform mechanically generated time series forecasts. Furthermore, our results provide motivation for other research areas, such as studies analyzing equity analysts' outputs, to control for a larger information set, for instance by including earnings information of related companies or information about overall business conditions. Second, our findings strongly support the assumption that macroeconomic survey forecasts represent a reasonable proxy measure for the anticipated information component in macroeconomic releases and consequently justify their use in the above mentioned research areas. Furthermore, our results highlight the danger to test for cognitive biases in a time series context which were previously only tested in controlled experiments. Especially when experiments are conducted in a highly regulated informational setting, i.e., when information given to test participants has to be strictly controlled for, as in anchoring bias experiments, it is questionable whether a direct transfer in a time series setting is possible at all. Future studies analyzing cognitive biases in time series frameworks have to consider carefully whether informational constraints might drive the results and lead to false conclusions.

The first essay provides strong evidence for the quality of macroeconomic forecasts at the consensus level, the second essay deals with individual macroeconomic forecasts and analyzes why certain analysts provide better forecasts then others. In particular, we focus on the association between the idiosyncratic predictability of a given macroeconomic indicator and the relation between analyst characteristics and macroeconomic forecast accuracy.

Obviously, there might be quality differences on the individual analyst level, i.e., there are more and less precise macroeconomic analysts. Exploiting these quality differences is a desirable task, because academics would obtain better proxy measures for market participants' expectations, and for investors an information advantage should translate into higher profits. We argue that if an indicator's idiosyncratic predictability is low, i.e., the series is almost not predictable, for instance due to information constrains and very volatile processes, then analysts' forecast performance is rather random than systematic because skills cannot take effect. In contrast, if a macroeconomic indicator has a high idiosyncratic predictability, then analysts with certain characteristics benefit from their abilities and skills, and generate more precise forecasts than less skilled analysts. Accordingly, for the unpredictable indicators the relation between analyst characteristics and forecast accuracy should be less pronounced than for the predictable ones. Consequently, we hypothesize that the idiosyncratic predictability of a certain macroeconomic indicator has to be taken into account whenever the relation between analyst characteristics and forecast accuracy is analyzed.

So far there is only contradictory evidence concerning differences in individual forecast quality of macroeconomic analysts. While some studies provide evidence for different forecast quality among individual macroeconomic analysts (e.g. Zarnowitz, 1984; McNees, 1987; Zarnowitz and Braun, 1993; Kolb and Stekler, 1996; Brown et al., 2008) other articles come to the opposite conclusion (e.g. Stekler, 1987; Ashiya, 2006). Despite this disagreement, the relation between macroeconomic forecast accuracy differences and analyst characteristics has not been analyzed so far, although the extensive strand of literature analyzing the association of equity analyst characteristics and earnings per share forecast accuracy (e.g. Clement, 1999; Clement and Tse, 2005; Brown and Mohammad, 2010) provides a sound framework for an analysis.[3]

[3] Only Brown et al., 2008 investigate the relation between analysts' ability and forecast precision. However, their study design differs significantly with respect to the set of used variables as well as the employed model compared to recent research designs used to analyze the performance of equity analysts.

Most importantly, we find that model performance heavily depends on the idiosyncratic predictability of macroeconomic indicators. With decreasing idiosyncratic predictability the relevance of analyst characteristics for forecast accuracy diminishes for some characteristics and disappears for others. In terms of economic significance we find substantial differences between macroeconomic indicators with high and low idiosyncratic predictability. Consequently, our results show that the idiosyncratic predictability of a given forecast target has to be taken into account when the association between analyst characteristics and forecast accuracy is analyzed.

Our findings have implications for different research areas. Most importantly we directly add to the literature analyzing individual macroeconomic analysts' forecast performance. We provide evidence that the idiosyncratic predictability of an indicator has to be taken into account if the relation between analyst characteristics and forecast accuracy is analyzed. Differentiation among analysts is only very limited if the figure to be forecasted is virtually unpredictable, because analysts do not benefit from their abilities and experiences. Systematic forecast accuracy differences arise if the forecast target is predictable at all and more skilled analysts have the opportunity to differentiate themselves form less skilled ones based on superior skills. Since there are differences in the predictability of company earnings our framework is transferable. Analogous to our findings for macroeconomic analysts, we expect that idiosyncratic predictability plays an equally important role analyzing the association between equity analysts' characteristics and their earnings per share forecast performance, i.e., for company earnings with higher idiosyncratic predictability we expect higher heterogeneity in forecast accuracy which can be explained by analyst characteristics.

The first two essays provide evidence that macroeconomic predictions are in general of high quality as they incorporate rationally information from various sources. Besides the previously analyzed macroeconomic forecasts, agents such as politicians and employers, also heavily rely on other information, for example, on coincident and leading macroeconomic indicators. Determining the current state of the economy and obtaining sound projections about future overall macroeconomic developments plays an important role in their decision processes. Coincident and leading macroeconomic indicators incorporate a large set of macroeconomic variables as well as stock and bond market measures, e.g., returns and interest rate spreads. However, there is no evidence about how expectations at the microeconomic level relate to expectations at the macroeconomic level. Consequently, an aggregate of microeconomic expectation data, i.e., individual company expectations, are not included in coincident and leading macroeconomic indicators so far.

To overcome this shortcoming we introduce a bottom-up approach that aggregates individual company expectations to derive macroeconomic content. Since the development of the entire economy is closely related to the development of its individual parts, among them individual companies, aggregated company information must contain macroeconomic information. Unfortunately, there is no database containing managements' expectations, however, we use equity analysts' outputs as proxy measure. Equity analysts' information sets comprise public macroeconomic-, industry- and company-specific content as well as non-public company-specific information (Grossman and Stiglitz, 1980) and is therefore arguably the best available proxy for managements' expectations. Regarding the choice of the best analyst's output we use recommendation changes instead of earnings per share (EPS) changes, because recommendations comprise more information. Besides the one year earnings estimate, recommendations also contain a series of future earnings expectations as well as interest rate and risk premium expectations. We show that aggregated recommendation changes as proxy measure for changing company outlooks have predictive power for overall economic developments.

Our results provide evidence that aggregated recommendation changes, which approximate changing expectations about individual companies' economic prospects, have predictive power for future macroeconomic developments of about one year. Controlling for other well established macroeconomic predictors our results remain robust indicating that our measure contains additional independent information. Consequently, it seems promising to include our new predictor into the set of macroeconomic predictors in future applications. Additionally, we find that EPS changes have no predictive power lending support to our assumption that more forward looking information, as included in recommendation changes, is required if one attempts to forecast future macroeconomic developments. Furthermore, our findings provide the missing link between previous studies showing that aggregated analyst outputs have predictive power for overall stock market developments (Howe et al., 2009) and those showing that the stock market leads the real economy (Stock and Watson, 1998). Our results support the notion that changes in expectations about future company performance rationally determine asset values in advance of overall economic activity changes providing the explanation why stock markets lead the real economy.

Overall, the three essays in this thesis advance different strands of literature. We show that macroeconomic consensus forecasts are a reliable proxy measure for market participants' expectations. Furthermore, our results provide strong evidence that it is dangerous to transfer psychological experiments into time series frameworks without appropriately controlling the informational environment. Additionally, we show that the idiosyncratic predictability of a given forecast objective, i.e. whether a forecast task is satisfyingly feasible at all, has to be taken into account whenever the association between analyst characteristics and forecast accuracy is analyzed. Macroeconomic analysts do only benefit from their superior skills compared to their competitors if the macroeconomic series is idiosyncratically predictable. For unpredictable series, forecast accuracy is rather random than systematic, because superior skills do not systematically translate in better forecasts. Finally, we show that the aggregation of forecasts on the microeconomic level, i.e., company expectations, is a promising approach to extract macroeconomic information. Overall, we conclude that macroeconomic analysts are very efficient information processors and play an important role as intermediaries in financial markets.

# Chapter 2 Irrationality or Efficiency of Macroeconomic Survey Forecasts? Implications from the Anchoring Bias Test

## 2.1 Introduction

A large and growing body of literature provides evidence that overall economic conditions strongly influence financial markets. Surprising macroeconomic data, in particular, unanticipated information in scheduled macroeconomic reports, change market participants' perceptions of the fundamental value of assets. As a result, macroeconomic news is shown to explain price changes in bond markets (e.g., Balduzzi et al., 2001; Gilbert et al., 2010), stock markets (e.g., Boyd et al., 2005; Cenesizoglu, 2011) as well as foreign exchange markets (e.g., Andersen et al., 2003; Evans and Lyons, 2008).[4] Macroeconomic news helps to explain co-movement of markets (e.g., Albuquerque and Vega, 2009), market volatility (e.g., Beber and Brandt, 2008; Brenner et al., 2009), changes in market liquidity (e.g., Brandt and Kavajecz, 2004; Pasquariello and Vega, 2007, 2009) as well as bond and equity risk premiums (e.g., Savor and Wilson, 2012; Dicke and Hess, 2012). As a common theme, all of these studies need to isolate the unanticipated information component in macroeconomic releases from already expected information. Therefore they rely heavily on macroeconomic survey forecasts collected, for example, by Bloomberg or Money Market Services (MMS)[5],

[4] Stock prices, for example, are largely affected by macroeconomic risk through three channels (e.g., Boyd et al., 2005), i.e., risk-free rates (e.g., Balduzzi et al., 2001), risk premiums (e.g., De Goeij et al., 2009; Bestelmeyer et al., 2012) and earnings expectations (e.g., Agarwal and Hess, 2012).

[5] Most frequently MMS survey data are used, for example, by Urich and Wachtel (1984), McQueen and Roley (1993), Almeida et al. (1998), Elton (1999), Balduzzi et al. (2001), Flannery and Protopapadakis (2002), Andersen et al. (2003), Green (2004), Bernanke and Kuttner (2005), Gürkaynak et al. (2005), Hautsch and Hess (2007), Evans and Lyons (2008), and Hautsch et al. (2012).

assuming that survey forecasts provide a reasonable approximation of market participants' expectations.

In contrast, Campbell and Sharpe (2009) recently suggest that there is a substantial "anchoring bias" in macroeconomic survey forecasts. But then these survey forecasts do not adequately approximate market participants' expectations as they can be easily improved. However, we cannot find improvements in forecast quality once we apply a look-ahead bias free test and adjustment procedure. Compared to the extensive research that has been conducted in the area of macroeconomic information processing in financial markets (e.g., Urich and Wachtel, 1984; McQueen and Roley, 1993; Andersen et al., 2003; Beber and Brandt, 2010; Gilbert, 2011) comparatively little analysis is available concerning the properties of macroeconomic survey forecasts. This fact is somewhat surprising, but possibly due to the high quality of survey forecasts. The few thus far available forecast rationality studies largely test for general quality properties derived from the rational expectations hypothesis of Muth (1961).[6] As a common outcome, general forecast rationality studies provide no evidence of systematic or persistent inefficiencies.[7] In contrast, Campbell and Sharpe (2009) test for a specific behavioral inefficiency, the anchoring bias, first documented by Tversky and Kahneman (1974) in psychological experiments. Anchoring implies that too much weight is attached to a certain prior available piece of information, the anchor, in the decision process. In the context of macroeconomic forecasts it means that analysts put too much importance on last months' data and therefore underweight meanwhile released important information. Thus the entire information set available at the survey date is not

[6] See e.g., Pesando (1975), Mullineaux (1978), Pearce and Roley (1985), Aggarwal et al. (1995), and Schirm (2003).

[7] The most recent study, Schirm (2003), only finds for a small number of investigated series some bias. However, his results partly contradict the findings of Aggarwal et al. (1995) obtained on a different sample.

efficiently incorporated into their forecasts. But then, correcting the analysts' mistake and utilizing the entire available information appropriately must yield improved forecasts.

However, we cannot reach this conclusion. In contrast, our analysis reveals a counter-intuitive result: despite a seemingly strong and statistically significant anchoring bias in most macroeconomic survey series, adjusting forecasts for the seemingly apparent bias leads to no systematic forecast improvements. Decomposing the anchoring bias test statistic provides an explanation for this puzzling result: the test itself is biased. Testing solely against a limited information set the anchoring bias test neglects the possibility that analysts may provide superior forecasts by using a more comprehensive information set than just the univariate time series itself. Our empirical results strongly support this explanation for a broad range of macroeconomic series.

Our analysis points out a universal risk inherent in behavioral tests using time series data: a test may detect a bias because rational agents' forecasts are based on a more comprehensive information set than the bias test controls for. In psychological studies,[8] the origin of the anchoring hypothesis, the available information set plays a central role. Participants usually have limited prior information about the figure they have to estimate, e.g., when students are asked to give an estimate for the price of a bottle of wine without knowing its quality. In contrast, professional forecasters are not confronted with a black box. Most likely, they have a much better understanding of the content of macroeconomic indicators and thus provide much more sound projections. Discrepancies between the information environment in which the forecast was generated and the information the bias test controls for may lead to false conclusions.

[8] See for instance Tversky and Kahneman (1974) and Ariely et al. (2003).

Our analysis proceeds in five steps. First, we replicate the anchoring bias test of Campbell and Sharpe (2009). However, we use a larger set of macroeconomic indicators, allowing for a more comprehensive analysis. More importantly, we use a much longer sample period in order to facilitate out-of-sample tests. This "dynamic" analysis, i.e., testing on a rolling-window and correspondingly adjusting forecasts out-of-sample, enables us to model the exact information flow as we consider only information available to market participants at a given point in time. Hence, our procedure avoids a look-ahead bias. This distinction is important because Croushore and Stark (2001), for example, show that different vintages of data can lead to different forecasting results. Even more important, in our analysis the available information set is of particular importance when we adjust the forecasts for the anchoring bias, because only a real-time proceeding ensures a realistic comparison of unadjusted and adjusted data. In contrast, Campbell and Sharpe's analysis (which we call "static") is based on a single in-sample regression. Therefore, a corresponding adjustment incorporates a potentially severe look-ahead bias.

If the highly significant anchoring coefficients would stem from a systematic cognitive bias, then adjusting the original survey forecasts must yield substantial improvements in forecast quality. Therefore, in a second step, we analyze whether bias adjustments actually improve survey forecasts. Surprisingly, despite highly significant anchoring coefficients, we can hardly find any significant improvements in forecast quality when adjusting for this seemingly apparent bias. Only when we allow for a look-ahead bias, i.e., when we apply the static estimation and adjustment, we find some modest improvements. In contrast, if we do not allow for a look-ahead bias, we can hardly find any statistically significant improvements. Overall, we have to conclude that nothing is gained by adjusting the original survey forecasts for anchoring.

In order to explain this puzzling result and to provide a basis for our subsequent empirical analysis we inspect the mechanics of the anchoring bias test analytically in an intermediate step. Most importantly, the anchoring bias test implicitly assumes a univariate time series framework. This creates a substantial problem since the test neglects other information which most likely alters rational forecasts. In particular, we show that the overall test statistic can be decomposed into two components: The first component captures inefficient processing of univariate time series information, possibly due to anchoring. The second component, however, captures superior information processing abilities of analysts, supposedly due to using a richer information set. Hence, large and significant anchoring coefficients cannot only arise when analysts face a cognitive bias but also when they correctly incorporate additional information into their predictions and therefore outperform time series forecasts. This finding provides a first indication that neglecting other information may be responsible for the misleading anchoring bias test results.

Outperforming optimal univariate time series forecasts implies that analysts have to use some additional information while generating their forecasts. In fact, in the third step of our empirical analysis, we provide evidence supporting the view that macroeconomic analysts utilize a more comprehensive information set. In particular, we find that a substantial part of the forecast improvements analysts achieve over time series models can be explained by other macroeconomic data. This finding suggests that analysts draw on several other macroeconomic indicators, in particular, on those macroeconomic figures that are usually found to have the strongest impact on prices of financial assets, because they are released early and have substantial information content (e.g., Gilbert et al., 2010).

In a fourth step we perform a simulation to better understand whether the inclusion of additional information may affect the anchoring test. The main idea is that there are several macroeconomic reports which provide correlated information, but are released successively.

Consequently, analysts can draw on more than just univariate time series information when they generate their forecasts. As expected, we obtain insignificant anchoring test results for the simulated forecast when these are generated utilizing only univariate time series information. In contrast, we get highly significant anchoring bias test results for the forecasts based on a more comprehensive information set. This clearly indicates that the anchoring test is very sensitive to additional information, for example, picked up by analysts through observing related releases. Our simulation is specific as it assumes a certain process environment, in particular, we assume that macroeconomic data series follow ARIMA processes. While this is a commonly used assumption, we cannot rule out that the test may work better in other environments. Nevertheless, our simulation clearly shows that the test can produce biased results, in particular, if we allow (mechanically generated) forecasts to draw on additional information besides the historical univariate time series.

Finally, in the fifth step, we quantify the relative contributions of the "inefficiency" and the "additional information" component to the overall anchoring bias test coefficient. We find that for the majority of seemingly biased forecast series, the "additional information" component accounts for more than half of the size of the overall anchoring bias test coefficient. In line with the findings in the previous steps, this again indicates that the additional information component is responsible for creating a substantial bias in the anchoring bias test statistic.

Overall, our analysis clearly shows that the highly significant anchoring bias results are due to deficiencies of the test. Focusing exclusively on a limited information set, the anchoring test can produce strongly misleading results if analysts process additional information beyond the univariate time series data. More generally, our analysis suggests that testing for a specific bias such as anchoring by focusing exclusively on univariate time series properties is dangerous since it neglects the ability of agents to aggregate additional information. In summary, our findings show that anchoring does not constitute any problem for studies building on survey forecasts, most importantly, because bias adjustments do not result in superior forecasts. Nevertheless, rejecting the validity of the anchoring test does not provide evidence that analysts are processing available information fully rationally.

Our study contributes to different strands of literature. First, our results directly add to the scarce literature analyzing the (in)efficiency of macroeconomic survey forecasts by showing that and explaining why analysts easily outperform mechanically generated time series forecasts. We find that the superiority of survey forecasts is largely attributable to the incorporation of related macroeconomic data. Second, our findings have important implications for a broad range of studies analyzing the impact of macroeconomic data as our results strongly support their assumption that macroeconomic survey forecasts provide unbiased estimates of the anticipated information components in macroeconomic releases (e.g., Albuquerque and Vega, 2009; Beber and Brandt, 2010; Gilbert, 2011). Furthermore, since the anchoring bias adjusted forecast is basically a weighted combination of the survey forecast and an autoregressive model we contribute to the area of forecast combination in which currently no results concerning monthly macroeconomic survey forecasts are available.

The remainder of the study is organized as follows. Section 2.2 briefly delineates the anchoring bias test and introduces our framework for the evaluation of analysts' forecasts. Section 2.3 describes the data and their properties. Section 2.4 provides the empirical results and section 2.5 concludes.

## 2.2 Methodology

The basic assumption of the anchoring bias test suggested by Campbell and Sharpe (2009) is that a survey forecast $F_{t-\tau}^{t}$ for the actual released figure $A_t$ in period *t* generated in *t*–*τ* is a linear combination of the two components $E\left[A_t \middle| I_{t-\tau}\right]$ and $A_{\bar{h}}$. $E\left[A_t \middle| I_{t-\tau}\right]$ denotes the unbiased forecast conditional on the information set $I_{t-\tau}$ which is available on the day the forecast was generated. $A_{\bar{h}}$ denotes an average of already released values for the $\bar{h}$ previous months. Hence, the survey forecast can be written as:

$$F_{t-\tau}^{t} = \lambda \cdot E\left[A_t \middle| I_{t-\tau}\right] + \left(1-\lambda\right) \cdot A_{\bar{h}}, \tag{2.1}$$

where $\lambda$ denotes the weight given to the unbiased forecast which already incorporates all available information efficiently. The inclusion of additional past information is redundant and therefore $\lambda$ should be equal to one. A value of $\lambda$ significantly smaller than one would suggest anchoring, i.e., that forecasters put too much weight on previously released values in comparison to an unbiased estimator.

Since the unbiased estimator is unobservable a direct estimation of Equation (2.1) is not feasible. Nevertheless, an indirect estimation of $\lambda$ is possible by means of Equation (2.2) (see Appendix 2 A for a derivation):

$$S_t = \gamma \cdot \left(F_{t-\tau}^{t} - A_{\bar{h}}\right) + \eta_t \,, \quad \text{with } \gamma \equiv \frac{\left(1-\lambda\right)}{\lambda}. \tag{2.2}$$

$S_t$ denotes the unanticipated news component defined as actual released figure minus forecast and $\eta_t$ the error term.[9] $\gamma > 0$ would indicate anchoring, as it implies that $\lambda < 1$. We use Equation (2.2) to test for anchoring in the first step of our analysis. Equation (2.1) suggests that the unbiased forecast $E\left[A_t \middle| I_{t-\tau}\right]$ is compounded of the survey based forecast and the anchor. This can be rewritten as:

$$E\left[A_t \middle| I_{t-\tau}\right] = \frac{F_{t-\tau}^t}{\lambda} - \frac{\left(1-\lambda\right)}{\lambda} \cdot A_{\bar{h}} = \left(1+\gamma\right) \cdot F_{t-\tau}^t - \gamma \cdot A_{\bar{h}}. \tag{2.3}$$

By means of Equation (2.3) and correspondingly estimated $\gamma$ the original forecast data can be adjusted for the bias induced by anchoring. These adjusted forecasts serve as input for our forecast quality comparison tests in step 2 of our empirical analysis. We apply the Diebold and Mariano (1995) test on differences in mean squared forecast errors for two different bias adjusted forecasts. First, to evaluate the in-sample impact of the anchoring bias we estimate Equation (2.2) over the entire sample period and then adjust the forecasts retrospectively. Second, to avoid an in-sample look-ahead bias, we perform a dynamic adjustment by means of a rolling estimation of Equation (2.2), i.e., to adjust the next forecast for *t* we take the estimated anchoring bias coefficient using information up to *t–1*. This dynamic out-of-sample adjustment represents an implementable strategy which market participants could adopt in real time, and thus, is preferable over the static in-sample adjustment.

Astonishingly, we find virtually no forecast improvements despite highly significant anchoring coefficients. To better understand this result we decompose the anchoring test coefficient analytically. The data generating process can well be described by a linear model

[9] Although Equation (2.2) does not include a constant we always include one in the estimation.

and therefore we use an ARIMA model to generate time series forecasts.[10] Accordingly, we assume that $A_t$ follows some ARIMA(p,d,q) process. Moreover, we suppose that analysts also use a corresponding ARIMA(p,d,q) model to generate their forecasts, but do not restrict themselves to looking at the historical time series of a single indicator. Instead we suppose that they possess some additional information $Z_t$ which may be useful to predict the innovation εt in $A_t$, e.g., from inspecting other macroeconomic announcements or simply from reading the daily press. Based on these considerations we show that $\hat{\gamma}$ in Equation (2.2) can be written as (see Appendix 2 B):

$$\hat{\gamma} = \frac{Cov\left(y'_{t-1}, x'_{t-1}\right) + Cov\left(\varepsilon_t, Z_t\right) - Var\left(Z_t\right)}{Var\left(x'_{t-1}\right) + Var\left(Z_t\right)}. \tag{2.4}$$

To separate the part of $\hat{\gamma}$ driven by the additional information set measured by $Z_t$ we decompose $\hat{\gamma}$ into two parts (see Appendix 2 B):

$$\hat{\gamma}_1 = \frac{Cov\left(y'_{t-1}, x'_{t-1}\right)}{Var\left(x'_{t-1}\right) + Var\left(Z_t\right)} \tag{2.5}$$

[10] Though previous studies detect nonlinearities in macroeconomic time series, there is strong evidence that simple linear forecasting models have a superior forecasting performance compared to sophisticated nonlinear models, such as self-exciting threshold autoregressive models (SETAR), markov-switching autoregressive models, or artificial neural networks (ANN) (e.g., Stock and Watson, 1999; Swanson and White, 1997). Furthermore, based on a Monte Carlo study Clements and Krolzig (1998) show that even if the data generating process is nonlinear, linear models provide robust forecasts. In fact, most recent studies (e.g., Marcellino, 2008) provide strong evidence that linear prediction models can hardly be beaten by more sophisticated nonlinear ones. Consequently, we follow Stock and Watson (1999): “If, however, a macroeconomic forecaster is restricted to using a single method, then, for the family of loss functions considered here [including MSE], he or she would be well advised to use an autoregression with a unit root pretest and data-dependent lag length selection.”

and

$$\hat{\gamma}_2 = \frac{Cov\left(\varepsilon_t, Z_t\right) - Var\left(Z_t\right)}{Var\left(x'_{t-1}\right) + Var\left(Z_t\right)} \tag{2.6}$$

with[11]

$$\begin{aligned} y'_{t-1} &\equiv \sum_{j=1}^{\infty}\left(\beta_j - \hat{\beta}_j\right)\cdot \varepsilon_{t-j} \\ x'_{t-1} &\equiv \sum_{j=1}^{\infty}\hat{\beta}_j \cdot \varepsilon_{t-j} + \frac{1}{h}\sum_{i=1}^{h}\sum_{j=1}^{h-i}\left(\varepsilon_{t-j} + \sum_{k=1}^{\infty}\beta_k \cdot \varepsilon_{t-j-k}\right). \end{aligned}$$

The first component $\hat{\gamma}_1$ captures inefficiencies in analysts' forecasts while the second component $\hat{\gamma}_2$ captures advantages from processing additional information. Clearly, $\hat{\gamma}_1 = 0$ if $Cov\left(y'_{t-1}, x'_{t-1}\right) = 0$, i.e., if analysts' estimates $\hat{\beta}_j$ are unbiased and $\hat{\beta}_j = \beta_j \,\forall j$ holds. Then $y'_{t-1} = 0$, i.e., analysts use unbiased coefficient estimates.

On the other hand, if analysts have superior forecasting abilities compared to the optimal time series model, $Z_t$ and $\varepsilon_t$ should be positively correlated. Consequently, $Cov\left(\varepsilon_t, Z_t\right) > Var\left(Z_t\right)$ suggests that some part of the anchoring coefficient $\hat{\gamma}$ is driven by the additional information utilized by analysts.

In the third step of our analysis we evaluate whether our assumption that analysts use a more comprehensive information set is realistic. Since additional information, $Z_t$, is not directly observable we have to use a proxy measure for our empirical analysis. The basic idea is to produce an optimal univariate time series forecast and to compare this to our MMS survey

---

[11] Note that the definition of $x'_{t-1}$ and $y'_{t-1}$ differ if $A_t$ does not follow an integrated process (see Appendix 2 B for details).

forecast. The difference of these two forecasts must stem from the use of additional information. To estimate an optimal ARIMA model for the actual announcement $A_t$, we first run a Phillips-Perron test to determine whether the considered series is stationary and difference the series if necessary. Then we estimate the corresponding ARMA models for all combinations of p = 0, …, 6 and q = 0, 1 using a rolling window:

$$A_t = \sum_{i=0}^{p} \alpha_i \cdot A_{t-i} - \sum_{j=0}^{q} \beta_j \cdot \varepsilon_{t-j} + \varepsilon_t. \tag{2.7}$$

Among these we select every period the best fitting model according to Bayes' information criterion (BIC), excluding models with serially correlated residuals. Based on this selection procedure, we obtain a dynamic "optimal" time series forecast for $A_t$. The estimated residuals of this model serve as proxy measure for the innovations $\varepsilon_t$, i.e., the component in $A_t$ which is not predictable from historical univariate time series information. To approximate $Z_t$ we apply a corresponding distributed lag model to the survey forecasts $F_t$, i.e., we regress the (differenced) forecasts $F_t$ on p lags of (differenced) $A_t$ and q lags of $\hat{\varepsilon}_t$:[12]

$$F_t = \sum_{i=0}^{p} \delta_i \cdot A_{t-i} - \sum_{j=0}^{q} \theta_j \cdot \hat{\varepsilon}_{t-j} + Z_t. \tag{2.8}$$

The residuals of this estimation, $Z_t$, i.e., the component in survey forecasts $F_t$ which cannot be traced back to past observed actuals, serve as approximation for the additional information component. To rule out the possibility that our proxy for additional information just picks up noise and to answer the question where analysts' outperformance stems from, we analyze how

[12] Although Equation (2.8) suggests the existence of a generated regressor problem ($\hat{\varepsilon}_{t-j}$) we do not control for it since we are only interested in the values for $Z_t$ which would not be altered using an adjusted inference.

$Z_t$ is related to information available at the time when analysts produce their forecasts. For this purpose we estimate the following model:

$$Z_t = \alpha + \beta M_t + \varphi_t, \tag{2.9}$$

where $Z_t$ denotes the approximated additional information component in survey forecasts. $M_t$ is a vector which contains the available macroeconomic information set for the 23 considered indicators seven days prior to an announcement and $\varphi_t$ denotes the error term. We use a stepwise regression approach to determine whether $Z_t$ is related to other macroeconomic news.[13] Our results suggest that the additional information component is highly correlated with previously released macroeconomic indicators. This supports our hypothesis that analysts use more than just univariate time series information.

In step four, we perform a simulation to obtain further evidence whether additional information can influence the anchoring bias test results. We exclude parameter instability, model uncertainty and look-ahead advantages. Most importantly, our simulation accounts for the fact that macroeconomic data for a given reporting month are released successively while the released data are contemporaneously correlated (see Appendix 2 C for technical details). First, we simulate correlated normally distributed random variables for $\varepsilon_t$ and for $Z_t$. With a higher correlation of $\varepsilon_t$ and $Z_t$, $Z_t$ should become more useful to predict $\varepsilon_t$. Using $\varepsilon_t$ we generate a time series of actuals, $A_t$. Once we have the simulated time series $A_t$ and $Z_t$, we estimate corresponding models on a rolling estimation window and produce one-step-ahead

[13] Since $Z_t$ is based on estimates and used as dependent variable in the second step regression conventional inference is invalid. To control for the resulting generated regressand problem we follow the approach of Dumont et al. (2005) and adjust the coefficient covariance matrix, respectively.

out-of-sample predictions.[14] We produce two types of forecasts: a "simple" one, i.e., one exploiting only information in past $A_t$, and a "sophisticated" one, i.e., one exploiting $Z_t$ in addition to $A_t$. Finally, we perform anchoring tests on both forecasts. Since both types of forecasts are mechanically generated the anchoring test should not be able to detect deficiencies. We find that the anchoring test produces insignificant results for the simple forecasts but significant results for the sophisticated ones. Since the simple and sophisticated forecasts differ only with respect to the information sets they utilize, our simulation clearly shows that the test can produce misleading results if analysts exploit additional information.

Finally, in step five, we quantify the contribution of the "additional information" and the "inefficiency" component to the overall anchoring bias coefficient. Based on our theoretical considerations including Equation (2.5) and (2.6) a partition is feasible and we can conclude whether irrationality or information efficiency drives the anchoring bias test results. For most series the additional information component accounts for a substantial proportion of the overall anchoring bias test coefficient which clearly contradicts the purpose of the anchoring bias test.

## 2.3 Data Description

We use a comprehensive data set comprising 23 well known macroeconomic indicators. Table 2.1 lists the series along with the abbreviations used in the following sections, their availability during the sample periods and the respective reporting unit. Medians of analysts' forecasts for these macroeconomic data are obtained from MMS and Action Economists.[15]

[14] Note that we abstract from model uncertainty since we use the same p, d and q as for the simulated series.

[15] Each Friday, MMS polls analysts' forecasts of macroeconomic figures to be released during the following week. Survey responses are received over a three- to four-hour period every Friday morning via fax or phone. The results of the survey are published at around 1:30 PM EST. In September 2003 MMS was acquired by

Whenever available, we use ALFRED vintage data to measure actual announced values.[16] Otherwise announced values provided by the survey agencies are used.

Table 2.2 shows sample means $(\mu)$ and standard deviations $(\sigma)$ for the 23 considered indicators (actuals, forecasts, and surprises). Sample means of the surprises are close to zero for most indicators implying that the forecasts are unbiased if not conditioned on a specific information set.

---

Informa. However, the original MMS survey was conducted until mid of December 2003. For the time after December 2003 we use forecasts provided by Action Economics (AE) because many of the former MMS employees responsible for the survey went to AE after MMS was taken over.

[16] The Federal Reserve Bank of St. Louis provides access to a broad set of US macroeconomic data in their online database called ArchivaL Federal Reserve Economic Data (ALFRED).

Table 2.1 Indicator Overview

This table reports the used indicators, the abbreviations used throughout the manuscript, the sampling period, and the corresponding unit.

| | | Sample Period | | |
|---|---|---|---|---|
| Indicator | Abbreviation | Start | End | Unit |
| Consumer Confidence | CC | 07/1991 | 12/2009 | Level |
| ISM (formerly NAPM) | ISM | 01/1990 | 11/2009 | Level |
| Nonfarm Payrolls | NF | 01/1985 | 11/2009 | Change (Thousands) |
| Civilian Unemployment Rate | UN | 01/1980 | 11/2009 | Level |
| Hourly Earnings | HE | 10/1989 | 11/2009 | % change |
| Producer Price Index | PPI | 01/1980 | 11/2009 | % change |
| Producer Price Index ex Food & Energy | PPI ex | 07/1989 | 11/2009 | % change |
| Retail Sales | RS | 01/1980 | 11/2009 | % change |
| Retail Sales ex autos | RS ex | 07/1989 | 11/2009 | % change |
| Consumer Price Index | CPI | 01/1980 | 11/2009 | % change |
| Consumer Price Index ex Food & Energy | CPI ex | 07/1989 | 11/2009 | % change |
| Industrial Production | IP | 01/1980 | 11/2009 | % change |
| Capacity Utilization | CU | 03/1988 | 11/2009 | Level |
| Housing Starts | HS | 01/1980 | 11/2009 | Level (Millions of Units) |
| Durable Goods Orders | DGO | 01/1980 | 11/2009 | % change |
| New Home Sales | NHS | 02/1988 | 11/2009 | Level (Thousands of Units) |
| Personal Income | PI | 01/1980 | 11/2009 | % change |
| Personal Consumption Expenditures | PCE | 06/1985 | 11/2009 | % change |
| Index of Leading Indicators | LI | 01/1980 | 11/2009 | % change |
| Construction Spending | CS | 02/1988 | 10/2009 | % change |
| Factory Orders | FO | 02/1988 | 10/2009 | % change |
| Business Inventories | BI | 02/1988 | 10/2009 | % change |
| Goods and Service Trade Balance | TRD | 01/1980 | 10/2009 | Level ($ Billions) |

Table 2.2 Summary Statistics

This table reports the means ( $\mu$ ) and standard deviations ( $\sigma$ ) of the actual announced value (Actual), the MMS forecast (Forecast), and the resulting surprise calculated as the difference of Actual and Forecast.

| | | Actual | | Forecast | | Surprise | |
|---|---|---|---|---|---|---|---|
| Indicator | N | $\mu$ | $\sigma$ | $\mu$ | $\sigma$ | $\mu$ | $\sigma$ |
| CC | 222 | 95.595 | 27.533 | 95.476 | 26.779 | 0.119 | 5.212 |
| ISM | 239 | 51.579 | 5.622 | 51.630 | 5.418 | -0.051 | 2.022 |
| NF | 299 | 106.080 | 200.982 | 115.080 | 157.507 | -8.783 | 109.935 |
| UN | 359 | 6.150 | 1.514 | 6.188 | 1.518 | -0.019 | 0.164 |
| HE | 240 | 0.273 | 0.205 | 0.261 | 0.064 | 0.011 | 0.195 |
| PPI | 359 | 0.219 | 0.653 | 0.258 | 0.382 | -0.039 | 0.399 |
| PPI ex | 245 | 0.135 | 0.282 | 0.161 | 0.091 | -0.026 | 0.265 |
| RS | 359 | 0.305 | 1.136 | 0.324 | 0.749 | -0.032 | 0.734 |
| RS ex | 245 | 0.286 | 0.592 | 0.319 | 0.295 | -0.033 | 0.445 |
| CPI | 358 | 0.291 | 0.319 | 0.299 | 0.258 | -0.010 | 0.151 |
| CPI ex | 244 | 0.225 | 0.133 | 0.223 | 0.062 | 0.001 | 0.116 |
| IP | 359 | 0.118 | 0.687 | 0.129 | 0.495 | -0.010 | 0.331 |
| CU | 261 | 80.203 | 3.586 | 80.190 | 3.579 | 0.006 | 0.370 |
| HS | 359 | 1.473 | 0.349 | 1.461 | 0.338 | 0.012 | 0.098 |
| DGO | 357 | 0.211 | 3.592 | 0.191 | 1.367 | 0.083 | 2.979 |
| NHS | 261 | 799.123 | 244.938 | 792.236 | 234.279 | 6.887 | 61.263 |
| PI | 358 | 0.456 | 0.443 | 0.407 | 0.304 | 0.051 | 0.304 |
| PCE | 292 | 0.413 | 0.499 | 0.377 | 0.385 | 0.033 | 0.227 |
| LI | 359 | 0.150 | 0.764 | 0.141 | 0.602 | 0.009 | 0.321 |
| CS | 260 | 0.218 | 1.072 | 0.113 | 0.568 | 0.105 | 1.003 |
| FO | 261 | 0.254 | 2.194 | 0.225 | 1.935 | 0.030 | 0.767 |
| BI | 261 | 0.221 | 0.458 | 0.189 | 0.342 | 0.032 | 0.239 |
| TRD | 358 | -21.695 | 19.241 | -21.524 | 19.291 | -0.170 | 2.272 |

## 2.4 Empirical Results

Our empirical analysis proceeds in five steps. First, we perform both in- and out-of-sample anchoring tests for the 23 macroeconomic series. Given the bias estimates, we analyze in a second step whether analysts' forecasts can be improved by adjustments for anchoring. Third, we evaluate analysts' forecasting abilities in comparison to optimally selected univariate time series models and analyze which additional information is contained in analysts' forecasts. Fourth, we present corresponding simulation results. Finally, we decompose estimated anchoring coefficients into an "inefficiency" and an "additional information" component and evaluate their relative contributions.

### 2.4.1 Anchoring Test Results

We start with a "static" or in-sample test design and estimate Equation (2.2) on the full sample for three different specifications of $\bar{h}$, where $\bar{h}$=1 corresponds to anchoring on the last month's actual only and $\bar{h}$=2 or 3 to anchoring on the mean of the two or three previously announced actual values, respectively. Since the static test involves a serious look-ahead-bias we perform a "dynamic" analysis in addition, estimating the anchoring coefficients on a rolling window with a fixed length of 10 years.

Table 2.3 reports results for the static as well as for the dynamic test. Regarding the static tests, we report the optimal $\bar{h}$, i.e., which regression specification performed best according to the Bayes' information criterion (BIC), along with the corresponding anchoring bias coefficient $\hat{\gamma}$. These results suggest that in about two thirds of the cases analysts use an average and not a single value as anchor. According to the test results survey forecasts for 18 out of the 23 macroeconomic series are substantially and significantly biased. However, for two of these series we obtain significantly negative coefficients which can hardly be explained by anchoring. Moreover, the large variation in the estimated coefficients suggests substantially different degrees of anchoring. For factory orders (FO), for example, $\gamma = 0.04$ implies that analysts put about 4% weight on last month's release and about 96% on the

Table 2.3 Anchoring Bias Test Results

This table reports results of anchoring bias estimates according to

$$S_t = \gamma \cdot \left( F_t - A_{\bar{h}} \right) + \eta_t,$$

where $S_t$ denotes surprises, i.e., actual value ($A_t$) minus MMS forecast ($F_t$), and $A_{\bar{h}}$ is the $\bar{h}$ month anchor (i.e., the mean of the $\bar{h}$ previously released actuals). The first two columns report the optimal $\bar{h}$ and estimated $\hat{\gamma}$ for a test performed on the full sample. Columns (3) to (5) report the results for rolling window regressions with a fixed length of 10 years. Inference is based on White Standard Errors. $^{*}$, $^{**}$, $^{***}$ indicates significance at the 10%, 5% and 1% level, respectively.

| | Static estimates | | Dynamic estimates | | |
|---|---|---|---|---|---|
| | | | most frequent | mean | std.dev. |
| Indicator | $\bar{h}$ | $\hat{\gamma}$ | $\bar{h}$ | $\hat{\gamma}$ | $\hat{\gamma}$ |
| CC | 1 | 0.940$^{***}$ | 1 | 0.922 | 0.365 |
| ISM | 1 | 0.297$^{**}$ | 2 | 0.225 | 0.284 |
| NF | 2 | 0.070 | 3 | 0.137 | 0.271 |
| UN | 2 | 0.054 | 1 | -0.187 | 0.284 |
| HE | 2 | 0.516$^{***}$ | 2 | 0.439 | 0.264 |
| PPI | 3 | 0.315$^{***}$ | 3 | 0.303 | 0.146 |
| PPI ex | 1 | 0.205$^{**}$ | 2 | 0.176 | 0.262 |
| RS | 1 | 0.166$^{***}$ | 1 | 0.183 | 0.114 |
| RS ex | 1 | 0.275$^{***}$ | 1 | 0.382 | 0.217 |
| CPI | 3 | 0.150$^{***}$ | 3 | 0.130 | 0.127 |
| CPI ex | 1 | -0.214$^{**}$ | 1 | -0.149 | 0.175 |
| IP | 3 | 0.256$^{***}$ | 2 | 0.204 | 0.173 |
| CU | 1 | 0.319$^{***}$ | 1 | 0.268 | 0.122 |
| HS | 1 | 0.281$^{**}$ | 1 | 0.339 | 0.153 |
| DGO | 3 | 0.398$^{***}$ | 2 | 0.350 | 0.103 |
| NHS | 3 | -0.104 | 1 | 0.557 | 0.228 |
| PI | 2 | 0.094 | 3 | 0.153 | 0.115 |
| PCE | 2 | 0.189$^{***}$ | 2 | 0.250 | 0.098 |
| LI | 3 | 0.174$^{***}$ | 3 | 0.150 | 0.107 |
| CS | 3 | -0.222$^{***}$ | 1 | -0.247 | 0.087 |
| FO | 3 | 0.040$^{**}$ | 1 | 0.036 | 0.033 |
| BI | 3 | 0.197$^{***}$ | 2 | 0.260 | 0.108 |
| TRD | 3 | -0.014 | 1 | 0.267 | 0.210 |

unbiased forecast. In contrast, for consumer confidence (CC) it seems that the unbiased estimator and the previously released actual enter the MMS forecast with approximately equal weights.

Results of the dynamic anchoring tests are given in Table 2.3 as well. For simplicity we only report the most frequently observed optimal $\overline{h}$ along with means and the standard deviations of the $\hat{\gamma}$s estimated on rolling windows of 10 years length. For most macroeconomic series the mean dynamic $\hat{\gamma}$s are largely comparable to their static $\hat{\gamma}$ counterparts, in particular, for the series which exhibit a significant static $\hat{\gamma}$. Note that the standard deviations of the dynamic $\gamma$ estimates are rather large and indicate a substantial variation over the sample. For instance, for CC we obtain a mean of 0.922 and a standard deviation of 0.365, stemming from a range of dynamic $\hat{\gamma}$s of -0.306 to 1.434 (unreported).

Although the dynamic test results appear to be slightly weaker, overall they are akin to the static test outcomes. For both, static and dynamic tests, we get sizable $\hat{\gamma}$ coefficients for most of the macroeconomic forecast series indicating substantial anchoring. At first glance this suggests partly predictable surprises and portends a poor quality of the frequently used survey forecasts. Consequently, this finding questions their appropriateness as proxy measures for market participants' expectations.

### 2.4.2 Can Anchoring Adjustments Improve Analysts‘ Forecasts?

Given the highly significant and sizable anchoring coefficients we would expect that analysts' forecasts can be substantially improved by adjusting them according to Equation (2.3). Results are given in Table 2.4. First, we compute in-sample adjustments applying the estimated static $\hat{\gamma}$ coefficients. Then, to evaluate the economic impact, we apply the dynamic adjustment. For both static and dynamic adjustments we report the change in root mean squared forecast errors (Δ RMSFE) resulting from these adjustments. Negative values indicate that the RMSFE of the adjusted MMS forecast is smaller than the unadjusted one, i.e., that the anchoring bias adjustment improves forecasts. To test whether these

Table 2.4 Impact of Anchoring Adjustments on Forecast Quality

This table reports adjusted survey forecasts according to the estimated anchoring bias

$$E\left[A_t \middle| I_{t-\tau}\right] = F_t^{adj} = \left(1+\hat{\gamma}\right)\cdot F_t - \hat{\gamma}\cdot A_{\bar{h}},$$

where $F_t$ denotes MMS forecast ($F_t$) and $A_{\bar{h}}$ is the $\bar{h}$ months anchor (i.e., the mean of the $\bar{h}$ previously released actuals). For convenience, column (1) redisplays static estimates of $\hat{\gamma}$. Columns (2)-(5) report the results of a Diebold-Mariano (DM) test with small sample adjustment for the equality of mean squared errors (MSE). $H_0$: MSE of $F_t^{\text{adj}}$ = MSE of $F_t$. Inference of $\hat{\gamma}$ is based on White standard errors. $^{*}$, $^{**}$, $^{***}$ indicates significance at the 10%, 5% and 1% level, respectively.

| | | Static adjustment | | Dynamic adjustment | |
|---|---|---|---|---|---|
| Indicator | Static $\hat{\gamma}$ | Δ RMSFE | DM | Δ RMSFE | DM |
| CC | 0.940$^{***}$ | -8.38% | 3.0611$^{***}$ | -7.94% | 2.3913$^{**}$ |
| ISM | 0.297$^{**}$ | -1.14% | 1.1600 | 3.36% | -1.4817 |
| NF | 0.070 | -0.15% | 0.3318 | 3.47% | -2.0611$^{**}$ |
| UN | 0.054 | -0.36% | 1.0481 | 1.06% | -1.1930 |
| HE | 0.516$^{***}$ | -6.09% | 1.9677$^{*}$ | -2.11% | 0.3788 |
| PPI | 0.315$^{***}$ | -6.81% | 2.2901$^{**}$ | -4.63% | 1.5963 |
| PPI ex | 0.205$^{**}$ | -2.44% | 1.0270 | -1.98% | 0.7873 |
| RS | 0.166$^{***}$ | -5.73% | 2.2316$^{**}$ | 15.41% | -0.9479 |
| RS ex | 0.275$^{***}$ | -6.39% | 2.2695$^{**}$ | -7.87% | 1.5492 |
| CPI | 0.150$^{***}$ | -2.58% | 1.7648$^{*}$ | -2.21% | 1.0900 |
| CPI ex | -0.214$^{**}$ | -2.42% | 1.5437 | -0.12% | 0.0846 |
| IP | 0.256$^{***}$ | -6.01% | 1.7234$^{*}$ | -4.82% | 1.0857 |
| CU | 0.319$^{***}$ | -3.65% | 1.7317$^{*}$ | -3.24% | 1.3350 |
| HS | 0.281$^{**}$ | -1.02% | 0.9885 | 1.22% | -1.0694 |
| DGO | 0.398$^{***}$ | -4.56% | 3.2828$^{***}$ | -5.71% | 2.5341$^{**}$ |
| NHS | -0.104 | -0.33% | 0.9153 | 1.36% | -0.8149 |
| PI | 0.094 | -0.53% | 0.9634 | 0.56% | -0.7182 |
| PCE | 0.189$^{***}$ | -6.95% | 1.8066$^{*}$ | -1.09% | 0.2461 |
| LI | 0.174$^{***}$ | -6.75% | 1.0762 | -1.56% | 0.8598 |
| CS | -0.222$^{***}$ | -1.90% | 2.0148$^{**}$ | -2.21% | 1.3186 |
| FO | 0.040$^{**}$ | -0.70% | 0.9962 | 0.12% | -0.0589 |
| BI | 0.197$^{***}$ | -1.40% | 1.2159 | -1.87% | 1.2158 |
| TRD | -0.014 | 0.00% | 0.1045 | 3.91% | -1.9588$^{*}$ |

improvements are significant, we run Diebold and Mariano (1995) tests on differences in mean squared errors (MSE).[17] Since macroeconomic analysts, in contrast to stock market analysts, have no incentives to issue systematic overoptimistic or pessimistic forecasts the assumption of a quadratic loss function implied by the MSE is uncritical.

By construction, the static adjustments cannot yield a larger RMSFE for the adjusted series. Nevertheless, the improvements are rather small. We observe a reduction of 8.38% at best. Moreover, the Diebold-Mariano tests find that only about 60% of the significantly biased forecast series can be improved. This is somewhat surprising since the static anchoring tests make use of forward looking information. Naturally, one would expect significant forecast changes whenever we get a significant anchoring test coefficient, at least for the static case.

The results of dynamic adjustments are much worse. When we adjust forecasts dynamically we obtain almost no improvements. There are only two exceptions, CC and DGO for which forecast improvements are significant according to the Diebold-Mariano test. These correspond to a reduction in RMSFEs of nearly 8% for CC and less than 6% for DGO. On the other hand, we observe also two cases with significantly worsened forecast errors, i.e., NF and TRD. For all other series, changes in forecast errors are insignificant though large in some cases. For example, we observe the largest though insignificant forecast error change for RS, worsening the series' RMSFE by around 15%. Since the dynamic adjustment best represents market participants' approach to correct for the cognitive bias our results provide strong evidence against the economic significance of the anchoring bias.

Moreover, note that the size of the anchoring coefficient is at best loosely related to the improvements. For instance the DGO bias coefficient is 0.398 and results in an RMFSE improvement of about 4.6%. In contrast the PCE anchoring bias coefficient is only 0.189 and

[17] The test we apply includes the small sample adjustment of Harvey et al. (1997).

leads to a considerable larger RMFSE reduction of about 7%. This odd pattern already suggests that the anchoring bias test results might be misleading, i.e., a sizable $\hat{\gamma}$ does not necessarily lead to large forecast improvements.

### 2.4.3 Analysts' Information Set

Our theoretical analysis provides a possible – though disturbing – explanation for the disconnection of forecast improvements and $\gamma$ coefficients. Equation (2.6) suggests that we may find a significant anchoring bias simply because analysts provide sophisticated forecasts by incorporating additional information beyond the univariate time series information. This is definitely not unreasonable. For example, just by reading the current newspapers, analysts can process other contemporaneous business news. Technically speaking, $\gamma$ coefficients may just reflect that analysts can forecast part of the innovation in the data generating process by drawing on a richer information set. This implies that survey forecasts are efficient – not inefficient as indicated by the anchoring bias test results.

To analyze whether survey forecasts actually outperform time series forecasts, we compare analysts' median forecasts $F_t$ for a given month $t$ with a mechanically generated "optimal" univariate time series forecast ($F_t^{TS}$) for the actual macroeconomic release $A_t$ obtained from the rolling estimation (10 year window) and prediction using Equation (2.7). Table 2.5 reports the results. The first column shows the parameters p, d, q for the most frequently best fitting ARIMA model. For example, for ISM a specification with p=1, d=0 and q=0, i.e., a simple AR(1) model, turns out to provide the best fit in most cases. Similarly, for the majority of the other series the optimal model is rather simple. Columns (2) and (3) of Table 2.5 provide a comparison of forecast errors (RMSFE) of our out-of-sample time series forecasts and analysts' predictions, respectively. Column (5) reports the relative difference. For every single macroeconomic series the RMSFE of analysts' forecasts is smaller, for most series by more than 20% implying economically significant better forecasts. To evaluate the statistical significance of these forecast improvements, we again use a Diebold-Mariano test with small sample adjustment. For 20 out of the 23 series we find significant differences in MSE. For the remaining three series, i.e., core PPI, HS, and CS, analysts' forecasts have a smaller error as well, though the differences are statistically insignificant. Hence, in line with previous

Table 2.5 Best Performing Time Series Model

In column (1) this table reports the most frequent ARIMA specification from the rolling estimation procedure. Column (2) and (3) report the root mean squared forecast errors (RMSFE) of the time series forecasts and the original MMS data. Column (4) shows the percentage difference of the RMSFE, where negative values indicate superiority of the MMS data. Column (5) contains the results of a modified Diebold-Mariano test (DM) for MSE equality ($H_0$: $MSE^{\text{time series forecast}} = MSE^{MMS}$). $^{*}$, $^{**}$, $^{***}$ indicates significance at the 10%, 5% and 1% level, respectively.

| Indicator | Most frequent ARIMA(p,d,q) | RMSFE ARIMA | RMSFE MMS | Δ RMSFE | DM |
|---|---|---|---|---|---|
| CC | 0,1,0 | 6.48 | 5.20 | -19.7% | -5.46$^{***}$ |
| ISM | 1,0,0 | 2.38 | 2.06 | -13.5% | -3.71$^{***}$ |
| NF | 1,0,1 | 136.71 | 109.97 | -19.6% | -4.17$^{***}$ |
| UN | 0,1,0 | 0.21 | 0.17 | -15.3% | -2.84$^{***}$ |
| HE | 3,0,0 | 0.15 | 0.13 | -10.8% | -1.87$^{*}$ |
| PPI | 0,0,1 | 0.72 | 0.43 | -39.6% | -2.88$^{***}$ |
| PPI ex | 0,0,0 | 2.10 | 0.29 | -86.2% | -1.04 |
| RS | 0,0,1 | 1.15 | 0.73 | -36.8% | -3.14$^{***}$ |
| RS ex | 1,0,0 | 0.96 | 0.53 | -45.0% | -2.25$^{**}$ |
| CPI | 0,0,0 | 0.26 | 0.13 | -49.3% | -3.33$^{***}$ |
| CPI ex | 1,0,0 | 0.31 | 0.10 | -68.3% | -1.89$^{*}$ |
| IP | 1,0,0 | 0.63 | 0.33 | -47.2% | -4.77$^{***}$ |
| CU | 0,1,0 | 0.56 | 0.40 | -28.4% | -4.18$^{***}$ |
| HS | 1,1,0 | 1.71 | 0.08 | -95.2% | -1.54 |
| DGO | 0,0,1 | 3.12 | 2.81 | -9.8% | -2.50$^{**}$ |
| NHS | 0,1,1 | 74.11 | 67.98 | -8.3% | -1.87$^{*}$ |
| PI | 0,0,0 | 0.46 | 0.31 | -32.7% | -3.30$^{***}$ |
| PCE | 0,0,1 | 0.43 | 0.20 | -52.7% | -2.78$^{***}$ |
| LI | 0,0,0 | 0.46 | 0.19 | -59.3% | -4.57$^{***}$ |
| CS | 0,0,0 | 1.22 | 0.91 | -25.4% | -1.52 |
| FO | 2,0,0 | 2.29 | 0.74 | -67.5% | -2.97$^{***}$ |
| BI | 1,0,1 | 0.37 | 0.25 | -32.7% | -3.75$^{***}$ |
| TRD | 0,1,1 | 2.93 | 2.46 | -16.2% | -3.01$^{***}$ |

research on equity analysts' earnings forecast performance (see, e.g., Brown et al., 1987), our estimation results clearly show that analysts provide superior forecasts in comparison to optimally selected univariate time series models.

Outperforming a model which optimally exploits univariate time series information can only stem from using a richer information set. To extract the forecast component which is unrelated to historical announcements (i.e., $Z_t$) we use Equation (2.8) and decompose $F_t$ into a component explained by historical time series information and a residual $\hat{Z}_t$. This residual could just represent noise picked up by analysts when producing their forecasts. In this case $\hat{Z}_t$ would not help to predict $A_t$, or more precisely, would be uncorrelated with our estimate of the innovation in $A_t$, i.e., $\hat{\varepsilon}_t$.

Correlations of $\hat{Z}_t$ and $\hat{\varepsilon}_t$ are reported in Table 2.6. Most importantly, we find solely positive and highly significant correlations of $\hat{Z}_t$ and $\hat{\varepsilon}_t$. This strongly suggests that $\hat{Z}_t$ represents not just noise being picked up somehow by analysts. In contrast, the additional information component in analysts' forecasts is able to predict some part of the innovation in announcements and therefore provides evidence that macroeconomic analysts possess certain predictive abilities. Since our approximated innovation $\hat{\varepsilon}_t$ constitutes the unpredictable part in an announcement after optimally employing univariate time series information, the high correlation of $\hat{Z}_t$ and $\hat{\varepsilon}_t$ also suggests that analysts' superior forecasting abilities stem from the incorporation of valuable additional information. Again, this finding is in line with studies analyzing stock analysts' forecast performance. For instance, Fried and Givoly (1982) document that stock analysts' outperformance over time series models is based on autonomous, i.e., additional information.[18]

Table 2.6 Residual Correlations

This table reports the variances of the innovation in announcements $\hat{\varepsilon}_t$ and the approximated additional information component in survey forecasts $\hat{Z}_t$, which we retrieved from optimally fitted distributed lag models as described in section 1. In addition, the correlation of $\hat{\varepsilon}_t$ and $\hat{Z}_t$ is provided. $^{***}$, $^{**}$, and $^{*}$ denotes significance of these correlations at the 1%, 5%, and 10% level, respectively.

| Indicator | $Var(\hat{\varepsilon}_t)$ | $Var(\hat{Z}_t)$ | Correlation$(\hat{Z}_t, \hat{\varepsilon}_t)$ |
|---|---|---|---|
| CC | 41.73 | 5.06 | $0.68^{***}$ |
| ISM | 5.31 | 0.92 | $0.50^{***}$ |
| NF | 19000.00 | 5182.06 | $0.60^{***}$ |
| UN | 0.04 | 0.01 | $0.50^{***}$ |
| HE | 0.04 | 0.00 | $0.35^{***}$ |
| PPI | 0.40 | 0.12 | $0.82^{***}$ |
| PPI ex | 0.08 | 0.01 | $0.35^{***}$ |
| RS | 1.16 | 0.55 | $0.77^{***}$ |
| RS ex | 0.35 | 0.09 | $0.69^{***}$ |
| CPI | 0.07 | 0.04 | $0.82^{***}$ |
| CPI ex | 0.01 | 0.00 | $0.24^{***}$ |
| IP | 0.39 | 0.17 | $0.86^{***}$ |
| CU | 0.28 | 0.08 | $0.74^{***}$ |
| HS | 0.01 | 0.00 | $0.57^{***}$ |
| DGO | 10.79 | 1.40 | $0.50^{***}$ |
| NHS | 4396.60 | 447.31 | $0.41^{***}$ |
| PI | 0.18 | 0.08 | $0.71^{***}$ |
| PCE | 0.22 | 0.14 | $0.90^{***}$ |
| LI | 0.50 | 0.29 | $0.90^{***}$ |
| CS | 1.10 | 0.32 | $0.37^{***}$ |
| FO | 4.16 | 3.15 | $0.93^{***}$ |
| BI | 0.12 | 0.04 | $0.77^{***}$ |
| TRD | 6.69 | 1.30 | $0.53^{***}$ |

[18] For example, it has long been argued that financial analysts provide more accurate earnings forecasts than univariate time series models because analysts use a broader information set than just the univariate time series of historical earnings. For earnings forecasts this enlarged information set presumably includes, among other things, macroeconomic information (e.g., Brown, 1993; Brown et al., 1987).

One potential source of valuable additional information is other macroeconomic news. Due to interrelations between macroeconomic indicators it is quite plausible that analysts utilize these releases in their forecast generation process. Therefore, other macroeconomic news should be able to, at least partly, explain the additional information approximated by $\hat{Z}_t$. Especially indicators released early in the cycle and those with large information content about the state of the economy should be useful (Gilbert et. al., 2010). We regress $\hat{Z}_t$ on all macroeconomic information available seven days prior to the announcements using a stepwise regression approach to identify the most influential indicators as described in Equation (2.9). Since $\hat{Z}_t$ is the result of a first step estimation, inference in this second step might be inefficient due to the uncertainty inherent in the generated regressand. To control for a possible influence on the estimators' covariance matrix we apply the adjustment proposed by Dumont et al. (2005). Table 2.7 shows the regression results for selected indicators. For the Producer Price Index excluding Food and Energy (PPI ex), for instance, the best fitting model comprises five other macroeconomic indicators, namely UN, CPI ex, CU, CS, and TRD, which contribute to the explanation of $\hat{Z}_t$.

Table 2.8 provides an overview for all indicators. The first column shows how many other indicators contribute to the explanation of $\hat{Z}_t$, the second column reports the associated $R^2$ and the last column shows how often the indicator is useful to explain $\hat{Z}_t$ of other macroeconomic series. The very early released CC, for example, helps to explain $\hat{Z}_t$ of 9 other macroeconomic series, i.e., it is contained in the best model for 9 other indicators. Overall, the results provide strong evidence that additional macroeconomic information can explain analysts' outperformance over mechanically generated forecasts. Depending on the indicator, between about 7.0% (CPI) and almost 81% (PCE) of the variation in $\hat{Z}_t$ is explained by other macroeconomic information. On average, R-squares amount to 36%.

Table 2.7 Additional Information Content for Selected Indicators

This table reports the regression results of the following model:

$$Z_t = \alpha + \beta M_t + \varphi_t,$$

where $Z_t$ denotes the approximated additional information component in survey forecasts and $M_t$ a vector containing the available macroeconomic information set for the 23 considered indicators seven days prior to an announcement. A stepwise regression approach is used to obtain the models. *, **, *** indicates significance at the 10%, 5%, and 1% level, respectively. To account for the generated regressand problem we report adjusted p-values according to the adjustment procedure introduced in Dumont et al. (2005).

| Indicator | CC | NF | PPI ex | RS ex | IP | DGO |
|---|---|---|---|---|---|---|
| CC | | 0.792*** | | 0.001* | -0.004*** | |
| ISM | | 2.592*** | | 0.013*** | 0.013*** | 0.064*** |
| NF | | | | 0.000*** | 0.001*** | 0.001* |
| UN | 0.749*** | 28.197*** | 0.042*** | | -0.095** | 0.201*** |
| HE | | | | | 0.389*** | |
| PPI | | | | | -0.098** | 0.163** |
| PPI ex | | | | -0.122** | | |
| RS | | | | | 0.064* | 0.132** |
| RS ex | | | | | | |
| CPI | | | | 0.197*** | 0.312*** | |
| CPI ex | | 81.980** | 0.195*** | | | 0.823* |
| IP | | | | | | 0.649*** |
| CU | | | 0.017*** | | -0.047*** | |
| HS | 1.124** | | | | | |
| DGO | | | | -0.035*** | | |
| NHS | | 0.066*** | | | 0.000* | |
| PI | | | | | | -0.223* |
| PCE | | 38.721*** | | | -0.183** | 0.273* |
| LI | | | | | | |
| CS | | 9.868*** | 0.010** | | 0.038** | 0.128** |
| FO | | | | 0.073*** | | -0.171*** |
| BI | | | | -0.067* | | |
| TRD | | | -0.001*** | | | |

Furthermore, column (3) reveals that the most influential indicators are those which are released early and which are commonly viewed to be good indicators of current or future economic activity. Consequently, we find CC, ISM, NFP, UN, IP, CU and NHS to be the most important components of the additional macroeconomic information set. ISM and NF for instance contribute in 12 out of 23 cases to the explanation of $\hat{Z}_t$. Overall we find strong evidence that analysts use a comprehensive information set which supports our hypothesis that additional information is the driving force of the anchoring bias coefficients.

Table 2.8 Additional Information Content, R-squared, and Indicator Frequency

This table reports the number of explanatory variables in the vector of available macroeconomic information $M_t$ in the regression

$$Z_t = \alpha + \beta M_t + \varphi_t,$$

the associated R-squared and the frequency of each indicator in $M_t$, i.e., in how many cases the respective indicator contributes to the explanation of $Z_t$.

| Indicator | # of variables in $M_t$ | $R^2$ | Frequency of indicator in $M_t$ |
|---|---|---|---|
| CC | 2 | 0.121 | 9 |
| ISM | 9 | 0.199 | 12 |
| NF | 7 | 0.314 | 12 |
| UN | 4 | 0.128 | 14 |
| HE | 6 | 0.401 | 5 |
| PPI | 3 | 0.102 | 8 |
| PPI ex | 5 | 0.395 | 6 |
| RS | 6 | 0.248 | 7 |
| RS ex | 8 | 0.439 | 3 |
| CPI | 2 | 0.071 | 6 |
| CPI ex | 5 | 0.239 | 4 |
| IP | 12 | 0.414 | 9 |
| CU | 10 | 0.302 | 10 |
| HS | 9 | 0.388 | 4 |
| DGO | 11 | 0.535 | 5 |
| NHS | 7 | 0.318 | 11 |
| PI | 7 | 0.433 | 4 |
| PCE | 7 | 0.808 | 7 |
| LI | 6 | 0.201 | 2 |
| CS | 6 | 0.572 | 7 |
| FO | 8 | 0.768 | 6 |
| BI | 11 | 0.392 | 6 |
| TRD | 7 | 0.509 | 6 |
| Min | | 0.071 | |
| Max | | 0.808 | |
| Mean | | 0.361 | |
| Median | | 0.388 | |

### 2.4.4 Simulation Results

Our results so far indicate that the significant anchoring test results are largely due to the use of a more comprehensive information set than the test assumes. Additional evidence from our simulation is provided in Table 2.9. We report the fraction of significant anchoring parameter tests (in percentage points) requiring a significance level of 1%. Hence an entry above 1.0 percent indicates that the test too often rejects the null hypothesis of no-anchoring for a given set of process parameters.

Most importantly, the anchoring test performs well for simple forecasts (i.e., forecasts exploiting exclusively univariate time series information) but not for sophisticated ones (i.e., forecasts exploiting a more comprehensive information set). As expected, in around one percent of the simulation runs the anchoring test indicates that simple forecasts are significantly biased (at the 1% significance level), irrespective of the specific ARIMA specification. In contrast, the anchoring test strongly suggests that sophisticated forecasts are substantially anchoring biased. For some parameter constellations, we obtain in nearly 100% of the simulation runs highly significant anchoring test coefficients. In general, the anchoring test performs more poorly if we assume a higher correlation of the simulated $\widetilde{Z}_t$ and $\widetilde{\varepsilon}_t$. Clearly, with a higher correlation the additional information becomes more valuable and the deviations of the sophisticated forecasts from the simple forecasts become more pronounced.

Since the mechanically generated simple and sophisticated forecasts differ only with respect to the utilized information set, the simulation results clearly show that the anchoring test is strongly misled by additional information picked up by the sophisticated forecasts. The high correlations of $\hat{Z}_t$ and $\hat{\varepsilon}_t$ reported in Table 2.6 strongly support the notion that analysts are able to extract such additional information, most likely from previously released reports.

Table 2.9 Simulation Results

This table reports simulation results for the anchoring test assuming that announcements of a given macroeconomic indicator series follow an AR(1) process (panel a) or an ARMA(1,1) process (panel b) or an ARIMA(1,1,1) process (panel c). For each process we allow for different parameter constellations, in particular, AR-parameters of 0.2, 0.4, 0.6, and 0.8 in combination with an MA-parameter of 0.2 and correlations between $\tilde{\varepsilon}_t$ and $\tilde{Z}_t$ of 0.2, 0.4, 0.6, and 0.8 (see Appendix 2 C for more details). For these parameter combinations, we report the percentage of cases in which the anchoring test indicates that a series of one-step-ahead forecast is significantly anchoring biased at the 1% significance level despite the fact that the simulated forecast series is bias free. Left hand side tables report results for simulated “simple” forecast series (i.e., forecasts utilizing exclusively univariate time series information). Right hand side tables report test results for simulated “sophisticated” forecasts series (i.e., forecasts utilizing in addition the information contained in $\tilde{Z}_t$, and thus, a more comprehensive information set).

Panel a: ARIMA(1,0,0) processes

Simple Forecasts

| | | Correlation ($\tilde{\varepsilon}_t, \tilde{Z}_t$) | | | |
|---|---|---|---|---|---|
| | | 0.8 | 0.6 | 0.4 | 0.2 |
| AR-par. | 0.8 | 1.0% | 1.2% | 0.9% | 1.5% |
| | 0.6 | 1.4% | 1.2% | 1.2% | 1.0% |
| | 0.4 | 1.4% | 1.2% | 0.9% | 0.8% |
| | 0.2 | 1.2% | 1.1% | 0.8% | 0.9% |

Sophisticated Forecasts

| | | Correlation ($\tilde{\varepsilon}_t, \tilde{Z}_t$) | | | |
|---|---|---|---|---|---|
| | | 0.8 | 0.6 | 0.4 | 0.2 |
| AR-par. | 0.8 | 65.4% | 56.7% | 16.4% | 2.2% |
| | 0.6 | 80.4% | 33.7% | 8.2% | 1.6% |
| | 0.4 | 44.3% | 17.5% | 5.8% | 1.0% |
| | 0.2 | 37.4% | 13.7% | 4.1% | 1.3% |

Panel b: ARIMA(1,0,1) processes

Simple Forecasts

| | | Correlation ($\tilde{\varepsilon}_t, \tilde{Z}_t$) | | | |
|---|---|---|---|---|---|
| | | 0.8 | 0.6 | 0.4 | 0.2 |
| AR-par. | 0.8 | 1.7% | 1.0% | 1.1% | 1.2% |
| | 0.6 | 1.1% | 1.0% | 1.1% | 1.2% |
| | 0.4 | 1.5% | 1.0% | 0.6% | 1.1% |
| | 0.2 | 1.1% | 1.2% | 1.3% | 1.3% |

Sophisticated Forecasts

| | | Correlation ($\tilde{\varepsilon}_t, \tilde{Z}_t$) | | | |
|---|---|---|---|---|---|
| | | 0.8 | 0.6 | 0.4 | 0.2 |
| AR-par. | 0.8 | 84.4% | 31.8% | 6.5% | 1.4% |
| | 0.6 | 63.5% | 33.5% | 10.2% | 2.2% |
| | 0.4 | 60.7% | 26.2% | 6.2% | 1.6% |
| | 0.2 | 44.9% | 16.2% | 5.6% | 1.5% |

Panel c: ARIMA(1,1,1) processes

Simple Forecasts

| | | Correlation ($\tilde{\varepsilon}_t, \tilde{Z}_t$) | | | |
|---|---|---|---|---|---|
| | | 0.8 | 0.6 | 0.4 | 0.2 |
| AR-par. | 0.8 | 1.2% | 1.0% | 0.9% | 1.1% |
| | 0.6 | 1.2% | 1.2% | 1.1% | 1.2% |
| | 0.4 | 1.6% | 1.0% | 1.1% | 1.3% |
| | 0.2 | 1.4% | 1.5% | 1.2% | 1.2% |

Sophisticated Forecasts

| | | Correlation ($\tilde{\varepsilon}_t, \tilde{Z}_t$) | | | |
|---|---|---|---|---|---|
| | | 0.8 | 0.6 | 0.4 | 0.2 |
| AR-par. | 0.8 | 67.3% | 17.5% | 3.6% | 1.2% |
| | 0.6 | 96.7% | 50.1% | 7.3% | 1.3% |
| | 0.4 | 99.8% | 72.6% | 12.5% | 1.3% |
| | 0.2 | 99.6% | 63.4% | 11.7% | 1.7% |

Hence, the additional information picked up by analysts may well explain the puzzling result of highly significant anchoring coefficients but insignificant gains from adjusting the forecasts.

### 2.4.5 Empirical Decomposition of Anchoring Coefficients

According to Equations (2.4) to (2.6) we can decompose the anchoring coefficient $\hat{\gamma}$ into an "inefficiency" component $\hat{\gamma}_1$ and an "additional information" component $\hat{\gamma}_2$. Table 2.10 provides statistics on $\hat{\gamma}_1$ and $\hat{\gamma}_2$. For comparison, static as well as dynamic $\gamma$ estimates are displayed in columns (1) and (2), respectively. Column (3) shows the approximated $\hat{\gamma}$ calculated on the basis of Equation (2.4). In addition, columns (4) and (5) show the two components of $\hat{\gamma}$, i.e., the "inefficiency" component $\hat{\gamma}_1$ and the "additional information" component $\hat{\gamma}_2$.[19] Most importantly, we find that $\hat{\gamma}_2$ is quite large for most macroeconomic series. Considering only the 16 macroeconomic series with a significantly positive anchoring bias coefficient, we find that in 10 out of these 16 cases $\hat{\gamma}_2$ accounts for more than 50% of the overall $\hat{\gamma}$ and in another three cases for more than 25%. Hence, in line with the previously described results, this analysis again tells us that the use of additional information is largely responsible for the misleading test results.

[19] Table 2.10 supports the notion that the assumptions underlying our decomposition are realistic. In general, we find a high conformance between the original regression based coefficients and the approximated ones, for instance for ISM a value of 0.297 versus 0.315.

Table 2.10 Gamma Decomposition

This table reports results of anchoring bias estimations:

$$S_t = \gamma \cdot \left(F_t - A_{\bar{h}}\right) + \eta_t,$$

where $S_t$ denotes surprise, i.e., actual value ($A_t$) minus MMS forecast ($F_t$), and $A_{\bar{h}}$ is the $\bar{h}$ months anchor (i.e., the mean of the $\bar{h}$ previously released actuals). Column (1) contains the coefficients from the static test setting, column (2) reports the mean coefficients from the rolling estimation. Column (3) to (5) show the corresponding approximations of $\hat{\gamma}$ and its decomposition into an "inefficiency" ($\hat{\gamma}_1$) and an "additional information" ($\hat{\gamma}_2$) component:

$$\hat{\gamma}_1 = \frac{Cov\left(y'_{t-1}, x'_{t-1}\right)}{Var\left(x'_{t-1}\right) + Var\left(Z_t\right)}$$

$$\hat{\gamma}_2 = \frac{Cov\left(\varepsilon_t, Z_t\right) - Var\left(Z_t\right)}{Var\left(x'_{t-1}\right) + Var\left(Z_t\right)}$$

| | Test results | | Model based approximation | | |
|---|---|---|---|---|---|
| Indicator | Static estimates $\hat{\gamma}$ | Dynamic estimates mean $\hat{\gamma}$ | Total $\hat{\gamma}$ | Inefficiency component $\hat{\gamma}_1$ | Add. information component $\hat{\gamma}_2$ |
| CC[1] | 0.940*** | 0.922 | 0.102 | 0.000 | 0.102 |
| ISM | 0.297** | 0.225 | 0.315 | 0.136 | 0.179 |
| NF | 0.070 | 0.137 | 0.099 | 0.011 | 0.088 |
| UN[1] | 0.054 | -0.187 | 0.010 | -0.010 | 0.020 |
| HE | 0.516*** | 0.439 | 0.298 | 0.285 | 0.013 |
| PPI | 0.315*** | 0.303 | 0.239 | 0.001 | 0.238 |
| PPI ex | 0.205** | 0.176 | 0.206 | 0.201 | 0.005 |
| RS | 0.166*** | 0.183 | 0.150 | 0.118 | 0.031 |
| RS ex | 0.275*** | 0.382 | 0.075 | 0.000 | 0.075 |
| CPI | 0.150*** | 0.130 | 0.067 | -0.026 | 0.092 |
| CPI ex | -0.214** | -0.149 | -0.214 | -0.158 | -0.056 |
| IP | 0.256*** | 0.204 | 0.189 | 0.026 | 0.163 |
| CU[1] | 0.319*** | 0.268 | 0.072 | -0.024 | 0.096 |
| HS[1] | 0.281** | 0.339 | 0.068 | 0.039 | 0.029 |
| DGO | 0.398*** | 0.350 | 0.328 | 0.229 | 0.099 |
| NHS[1] | -0.104 | 0.557 | 0.188 | 0.133 | 0.055 |
| PI | 0.094 | 0.153 | 0.101 | 0.060 | 0.041 |
| PCE | 0.189*** | 0.250 | 0.128 | 0.065 | 0.062 |
| LI | 0.174*** | 0.150 | 0.106 | -0.007 | 0.113 |
| CS | -0.222*** | -0.247 | -0.210 | -0.082 | -0.128 |
| FO | 0.040** | 0.036 | 0.040 | -0.002 | 0.042 |
| BI | 0.197*** | 0.260 | 0.197 | -0.036 | 0.233 |
| TRD[1] | -0.014 | 0.267 | 0.052 | -0.007 | 0.059 |

[1] The deviation between column (1) and (3) is largely due to neglecting non-stationarity. For details see Table 2.11.

Table 2.11 Gamma Decomposition for Non-Stationary Macroeconomic Series

This table reports results of anchoring bias estimations:

$$S_t = \gamma \cdot \left(F_t - A_{\overline{h}}\right) + \eta_t,$$

where $S_t$ denotes the surprise, i.e., actual value ($A_t$) minus MMS forecast ($F_t$), and $A_{\overline{h}}$ is the $\overline{h}$ months anchor (i.e., the mean of the $\overline{h}$ previously released actuals) considering that the actual time series are non-stationary and therefore are regressed in changes. Column (1) contains the coefficients from the static test setting. Column (2) to (4) show the corresponding approximations of $\hat{\gamma}$ and its decomposition into an "inefficiency" ($\hat{\gamma}_1$) and an "additional information" ($\hat{\gamma}_2$) component for integrated series (see Appendix 2 B):

$$\hat{\gamma}_1 = \frac{Cov\left(y'_{t-1}, x'_{t-1}\right)}{Var\left(x'_{t-1}\right) + Var\left(Z_t\right)}$$

$$\hat{\gamma}_2 = \frac{Cov\left(\varepsilon_t, Z_t\right) - Var\left(Z_t\right)}{Var\left(x'_{t-1}\right) + Var\left(Z_t\right)}$$

| Indicator | Static estimates $\hat{\gamma}$ | Model based approximate total $\hat{\gamma}$ | Inefficiency component $\hat{\gamma}_1$ | Add. information component $\hat{\gamma}_2$ |
|---|---|---|---|---|
| CC | 0.066 | 0.102 | 0.000 | 0.102 |
| UN | -0.034 | 0.010 | -0.010 | 0.020 |
| CU | 0.072 | 0.072 | -0.024 | 0.096 |
| HS | 0.071* | 0.068 | 0.039 | 0.029 |
| NHS | 0.069 | 0.188 | 0.133 | 0.055 |
| TRD | 0.071* | 0.052 | -0.007 | 0.059 |

## 2.5 Conclusion

We find no support for the hypothesis that survey forecasts of macroeconomic data are anchoring biased. Despite highly significant anchoring coefficients in most forecast series, macroeconomic survey forecasts can hardly be improved. We attribute the misleading test results to the fact that the anchoring test itself is biased. While macroeconomic analysts easily outperform mechanically generated forecasts by incorporating additional information contained in other macroeconomic reports, the anchoring test neglects the possibility that analysts use a comprehensive information set. This focus on univariate time series information is largely responsible for misleading results of the anchoring test.

Our findings have important implications for a broad range of empirical studies relying on macroeconomic survey forecast, as we show that the anchoring test results are invalid. There is no need to replicate earlier studies with anchoring adjusted survey forecasts and further studies would not be well advised to adjust for this seemingly apparent bias. More generally, our results show that it is dangerous to test for a cognitive bias in a time series context which was previously only tested in controlled experiments. Especially when the prior information given to test participants has to be strictly controlled for, as in anchoring bias experiments, it is questionable whether a direct transfer in a time series setting is possible at all. Future research has to consider case-by-case whether a non-experimental setting is comparable to an experimental one and whether (implicit) informational constraints drive the results. Moreover, our results suggest that the analysis of analysts' forecast models provides a fruitful approach for further research. An intriguing question is whether and to what extent other forecasts, such as equity analysts' forecasts, can be better explained using a more comprehensive information set, for example, by including earnings information of related companies or information about overall business conditions.

# Chapter 3 Macroeconomic Forecast Accuracy — Idiosyncratic Predictability and Analyst Characteristics

## 3.1 Introduction

Differentiating between more and less precise macroeconomic forecasters based on their characteristics is a desirable task. From an academic perspective this would allow to calculate better consensus measures representing a superior proxy for market participants' expectations. For investors, obtaining more precise information directly translates into potentially higher profits.

While there are approaches to differentiate equity analysts (e.g. Clement, 1999; Clement and Tse, 2005; Brown and Mohammad, 2010) and macroeconomic analysts (e.g. Brown et al., 2008) based on their characteristics we argue that analysts can only benefit from superior skills or experience if the idiosyncratic predictability of a given indicator is taken into account. While some macroeconomic indicators are better predictable, for instance, due to earlier released information or a generally smoother process, other indicators are unpredictable. Consequently, analysts' individual forecast performance for the idiosyncratically unpredictable indicators is rather random than systematic, i.e., does not depend on their characteristics. In contrast, if the indicator is better predictable, then certain skills and experience should translate into superb forecast accuracy and heterogeneity among analysts. Our results support this notion and provide evidence that the idiosyncratic predictability is a major factor for the performance of models differentiating accuracy based on analyst characteristics. We find that indicator specific experience, general ability, the number of covered indicators, as well as the forecast horizon contribute positively to forecast accuracy, i.e., reduce absolute forecast errors. Most importantly, we find that model performance heavily depends on the idiosyncratic predictability of macroeconomic indicators. The relevance of analyst characteristics for forecast accuracy diminishes for general ability and disappears for the remaining characteristics with decreasing idiosyncratic predictability. Economic significance differs substantially between macroeconomic indicators with high and low idiosyncratic predictability. Consequently, we have to conclude that predictability has to be taken into account and that analyst characteristics are nearly immaterial for indicators which are hard to predict.

Although general macroeconomic forecast quality, i.e., the absence of systematic or persistence inefficiencies, on the consensus level is largely agreed on (e.g., Pesando, 1975; Mullineaux, 1978; Pearce and Roley, 1985; Aggarwal et al., 1995; Schirm, 2003; Hess and Orbe, 2013) evidence concerning differences in individual forecast quality of macroeconomic analysts is mixed. Some studies provide evidence for different forecast quality among individual macroeconomic analysts (e.g. Zarnowitz, 1984; McNees, 1987; Zarnowitz and Braun, 1993; Kolb and Stekler, 1996; Brown et al., 2008). However, other articles, come to opposite conclusions (e.g. Stekler, 1987; Ashiya, 2006). Despite this disagreement, the relation between macroeconomic forecast accuracy differences and analyst characteristics is so far an untouched research area.[20] In order to analyze the relation between idiosyncratic predictability, analyst characteristics, and macroeconomic forecast accuracy it is desirable to employ a model which showed its ability to detect relations between forecast accuracy and analyst characteristics. Surprisingly, despite the extensive strand of literature analyzing the association of equity analyst characteristics and earnings per share forecast accuracy (e.g. Clement, 1999; Clement and Tse, 2005; Brown and Mohammad, 2010) the established

[20] Only Brown et al., 2008 investigate the relation between analysts' ability and forecast precision. However, their study design differs significantly with respect to the set of used variables as well as the employed model compared to recent research designs used to analyze the performance of equity analysts.

characteristics framework used in these studies has not been transferred to macroeconomic analysts although Zarnowitz (1984) already noticed that "… a further study of the characteristics, methods and results of the forecasters with the best records will be needed". Moreover, none of the studies analyzes the association between analyst characteristics, accuracy, and idiosyncratic predictability differences.

We find that the association between analyst characteristics and forecast accuracy heavily depends on idiosyncratic predictability of the respective macroeconomic indicator. While our results for the predictable indicators imply sizable benefits from differentiating between more and less skilled analysts, results for the unpredictable indicators are much weaker. We find a decreasing influence on forecast accuracy for some characteristics and for other characteristics the impact even disappears with diminishing idiosyncratic predictability. Consequently, we have to conclude that the idiosyncratic predictability is an important determinant which has to be taken into account whenever the relationship between characteristics and forecast accuracy is analyzed.

Our findings have implications for different research areas. Most importantly we show that there is a relation between analyst characteristics and forecast accuracy, but the idiosyncratic predictability of an indicator has to be taken into account when employing an analyst characteristics model. There is only very limited differentiation among analysts' accuracy if the figure to be forecasted is virtually unpredictable because analysts do not benefit from their abilities and experiences. In contrast, if the forecast target is predictable more skilled analysts have the opportunity to differentiate themselves form less skilled ones and consequently heterogeneity in characteristics translates into systematic forecast accuracy differences. A similar argumentation holds for equity analysts' forecasts. While it is reasonable to assume that analysts obtain earnings guidance from the respective management and that they receive information in the course of the year idiosyncratic predictability is given for all companies.

However, more skilled analyst should outperform less skilled ones if the forecast task gets more complex. Consequently, we expect the association between equity analyst characteristics and accuracy to be more pronounced for companies which earnings are more difficult to predict than for companies with easy to forecast earnings.

The remainder of this paper is organized as follows. Section 3.2 states the hypotheses. Section 3.3 describes our research methodology and the data set. Section 3.4 provides the results and section 3.5 concludes.

## 3.2 Hypotheses

The idiosyncratic predictability across macroeconomic indicators differs significantly. While it is easy to predict certain indicators, it is harder to predict others. There are various reasons why some indicators are harder to predict. The employment report, for instance, is released very early within the monthly release cycle. Until its release there is only a very limited amount of information available to market participants. Consequently, it might be harder to predict nonfarm payrolls than for instance the consumer price index which is released later in the month. A report such as the employment report might therefore be idiosyncratically unpredictable. If it is unpredictable then analyst characteristics, e.g. whether an analyst is more or less experienced, do not matter and analysts' forecast performance is more random than systematic. In contrast, if a series is idiosyncratically predictable, for instance because prior to the release other information concerning the reference month are already released, analyst characteristics matter. Idiosyncratic predictability therefore is a necessary condition for systematic forecast quality differences. Therefore we expect that only for these indicators analysts benefit from their skills and experience. Consequently, we hypothesize that analyst characteristics explain accuracy differences better for macroeconomic indicators which are idiosyncratically predictable and worse for less predictable ones.

To evaluate forecast accuracy and the relation to analyst characteristics depending on idiosyncratic predictability we rely on well established characteristics used in previous research analyzing the accuracy of equity research analysts.[21] While some of these characteristics are directly observable, such as experience, there are others, e.g. ability, which are only indirectly observable. We classify characteristics in two categories, inherent characteristics such as experience and ability which the analyst possesses. Furthermore, there are exogenous characteristics on which the analyst explicitly decides on, i.e., at which point in time he submits his forecast or how many macroeconomic indicators he covers.

Obviously, a more experienced analyst should perform his assigned task better than a less experienced one. Over time he gains a better understanding of the underlying driving forces and increases his forecast performance. Experience has two dimensions, general and indicator specific experience. While it is beneficial to have some general experience in performing forecasts, every single indicator has its own specific characteristics. To detect other indicators with predictive power for the series to be forecasted and to isolate relevant information which is not contained in databases, i.e., from the daily press, a certain degree of indicator specific experience is required to generate precise forecasts. Consequently, we hypothesize that analysts with more general or indicator specific experience in forecasting macroeconomic variables generate forecasts with smaller deviations from the actual released figures, i.e., are more precise.

By default analysts gain experience with every forecast they submit. However, experience does not take into account the quality of these forecasts. It is possible that an analyst submits forecasts for years, however his precision might be inferior over the entire period. Consequently, he might be considered experienced, but this does not imply that he is a precise analyst. Therefore, we need a measure that takes an analyst's ability to generate high quality

[21] See for instance Clement (1999), Clement and Tse (2005), and Brown and Mohammad (2010).

forecasts into account. Only if the analyst is able to set up a model properly, understands the interrelations between different macroeconomic indicators and can interpret the available information set correctly, he is able to generate precise forecasts. In contrast to experience we cannot directly observe ability, however, we can use an established measure (Brown and Mohammad, 2010) which proxies unobservable skills of the analyst, hereafter called general ability. As with experience we expect forecast accuracy to increase with higher general ability.

A further forecast quality determinant is portfolio complexity, i.e., how many indicators an analyst has to follow. In contrast to the equity analyst literature, (e.g. Clement, 1999) which associates an increasing number of covered companies and industries with decreasing forecast precision, the situation is different in a macroeconomic forecasting context. Since interrelations between macroeconomic indicators are very distinctive, it is unfavorable for analysts to cover only a subset of the important indicators. The more indicators the analyst covers the better his understanding of the entire economy and the higher the likelihood that he identifies interrelations correctly and uses these information in his forecast generation process. Accordingly, we hypothesize that the more indicators an analyst covers, the better is his understanding of overall macroeconomic developments and consequently his forecast quality.

Prior research of equity analysts' accuracy determinants shows that the forecast horizon has a large negative influence on performance, i.e., analysts submitting their forecasts early tend to provide substantially less accurate forecasts than those who submit them late. There may be several weeks or even months between the forecasts of equity analysts. In contrast, macroeconomic analysts provide their forecasts within a very narrow window, usually within one week. However, even this short window might be enough to gather relevant information which might contribute to forecast quality (e.g. Hess and Orbe, 2013). The shorter the forecast horizon, the longer the time span in which analysts can collect relevant data and incorporate this information advantage into their forecast. Therefore, we expect forecast accuracy to increase with decreasing forecast horizons.

Since the literature analyzing the performance of equity analysts (e.g. Sinha et al., 1997; Clement and Tse, 2005) finds positive short-term persistence in forecast accuracy we include past forecast performance in our analysis. We refrain from including broker size as proxy for recourses available to the analyst, as common in the equity analyst literature, for several reasons. In contrast to equity analysts, macroeconomic analysts are not dependent on relationships to the covered companies which maintenance is associated with costs, e.g. participation on facility visits and travel expenses to analyst conferences.[22] The macroeconomic analysts must not maintain relationships with agencies, because the leakage of information from the reporting agencies is highly unlikely. Furthermore, costs for the collection of information are very low for macroeconomic analysts, because macroeconomic data are freely available from the reporting agencies and in databases.[23] Moreover, it is common that one analyst covers all indicators for one broker leading to virtually no differentiation based on the number of employed analysts as a proxy measure for broker size as usually used in the equity analyst literature.

[22] Although the company might invite the analyst on their costs, it is advisable for the broker to pay for the trip of the analyst to stay objective (see e.g. Code of Conduct for Financial Analysts from the Chartered Financial Analyst Institute).

[23] The Federal Reserve Bank of St. Louis provides free access to a broad set of US macroeconomic data in their online database called ArchivaL Federal Reserve Economic Data (ALFRED).

## 3.3 Methodology and Data Description

Previous research analyzing the quality of macroeconomic forecasts on an individual level used the ASA-NBER Quarterly Economic Outlook Survey (e.g. Zarnowitz, 1984). However, it is desirable to have data on a higher frequency to have sufficient observations on the individual indicator level. This enables us to conduct our analysis on subsamples dependent on the idiosyncratic predictability of certain macroeconomic indicators. Consequently, we use monthly data for 20 macroeconomic indicators obtained from Bloomberg covering the report periods from July 1998 to July 2012 on a monthly basis.

Let *j* denote analyst j, *i* the covered macroeconomic indicator and *t* the respective report period. Correspondingly, $F_{j,i,t}$ denotes the forecast of analyst j for indicator i in period t and $A_{i,t}$ the actual announced figure for indicator i in period t. Our measure of forecast accuracy is the absolute forecast error ($AFErr$) scaled by the corresponding indicator and time specific standard deviation to make the data comparable across the range of considered indicators in this study:

$$AFErr_{j,i,t} = \frac{\left| A_{i,t} - F_{j,i,t} \right|}{Std\left( AFErr_{i,t} \right)}. \tag{3.1}$$

$L.AFErr$ denotes the absolute forecast error from last period and is included in the model to control for short term persistence.

In order to measure idiosyncratic predictability we define the reciprocal of the median scaled absolute forecast error of a given indicator as its predictability:

$$Predictability_i = \frac{1}{Median\left( AFErr_{j,i,t} \right)}. \tag{3.2}$$

Concerning the analyst characteristics we distinguish between indicator specific ($SpExp$) and general experience ($GenExp$):

$SpExp_{j,i,t}$ = number of previously reported forecasts of analyst j for indicator i until report period t.

$GenExp_{j,t}$ = number of previously reported forecasts of analyst j until report period t.

To measure analysts' ability we follow Brown and Mohammad (2010) and define general ability ($GAbil$) as:

$$\begin{aligned} F\text{-}Ability_{j,i,t} &= 1 - \left( \frac{Rank\left( AFErr_{j,i,t} \right) - 1}{Analyst\,Coverage_{i,t} - 1} \right) \\ GAbil_{j,i,t} &= \frac{1}{K-1} \sum_{\substack{k=1 \\ k \neq i}}^{K} F\text{-}Ability_{j,k,t}, \end{aligned} \tag{3.3}$$

where $Analyst\ Coverage_{i,t}$ denotes the number of analysts submitting forecasts for indicator i in period t. *F-Ability* is a standardized ranking of analysts based on their indicator specific forecast performance. The most precise analyst receives a value of 1, the least accurate one a value of 0. To generate a broad measure of analysts' ability, *GAbil* averages the indicator specific ability scores for a given report period for all other indicators, i.e., excluding the score of the indicator which is analyzed.

Our proxy measure of portfolio complexity is defined as the number of covered indicators:

$CovInd_{j,t}$ = measure for the number of covered indicators and defined as the number of indicators analyst j follows in report period t.

Since our dataset allows a daily identification of submitted forecasts, we are able to precisely calculate the time span between the forecast and the announcement date:

$Horizon_{j,i,t}$ = number of days between the announcement and the date analyst j submitted his forecast for indicator i for report period t.

Our dataset comprises many different indicators over a sample period characterized by turbulent periods. It is reasonable to assume that it was more difficult to predict certain indicators in certain months. For instance, the median of the absolute forecast error averaged across all analysts in a given month for nonfarm payrolls is about 60 thousand. After the last recession and the beginning of the recovery in the labor market in February/March 2010, macroeconomic analysts made forecast errors of 102 thousand for the April and 116 thousand for the May figures. This fact indicates that they were surprised by the turning point and that it was more difficult to predict those numbers than for instance in a period of stable growth. To control for this indicator-month effect we estimate the following equation using an indicator-month fixed effect estimator[24]:

$$\begin{aligned} AFErr_{j,i,t} = {} & \alpha + \beta_1 \cdot L.AFErr_{j,i,t} + \beta_2 \cdot SpExp_{j,i,t} + \beta_3 \cdot GenExp_{j,t} \\ & + \beta_4 \cdot GAbil_{j,i,t} + \beta_5 \cdot CovInd_{j,t} + \beta_6 \cdot Horizon_{j,i,t} + \varepsilon_{j,i,t}, \end{aligned} \tag{3.4}$$

where all variables are defined as described above and $\varepsilon$ denotes the error term.

Besides the analysts' individual estimates and estimation dates, the dataset comprises the name as well as the broker the analyst works for. Our initial sample consists of 164,030 observations for 20 indicators covering the report periods from July 1998 until July 2012.

[24] Note that this estimation procedure is equivalent to an OLS estimation with indicator-month specific dummy variables.

Table 3.1 reports the distribution of submitted forecasts over time as well as the number of analysts participating and the number of covered indicators.

While the majority of indicators considered in this study are included in the sample right from the beginning (18 out of 20), the number of participating analysts is low in the first years. Starting with only 6 analysts in 1998, the number of survey participants steadily increases to about 100 in 2005 and stabilizes on this level. Along with the increasing number of analysts, the number of submitted estimates increases rapidly. Excluding the first five years in our analysis to reduce the effect of left censoring and to achieve a sufficient cross sectional variation in characteristics such as specific or general experience, we still retain 85.6% of the initial sample.

Table 3.1 Yearly Summary Statistics

This table reports the total number of submitted forecasts to Bloomberg across all indicators on a yearly basis (Number of Estimates), as well as the percentage each year's forecasts contribute to the overall sample (Number of Estimates as % of total sample). Furthermore, the number of participating analysts in each year (Number of distinctive Analysts) and the number of covered indicators (Number of distinctive Indicators) are shown.

| Year | Number of Estimates | Number of Estimates as % of total sample | Number of distinctive Analysts | Number of distinctive Indicators |
|---|---|---|---|---|
| 1998 | 113 | 0.07 | 6 | 18 |
| 1999 | 3829 | 2.33 | 44 | 18 |
| 2000 | 4697 | 2.86 | 47 | 18 |
| 2001 | 5600 | 3.41 | 52 | 20 |
| 2002 | 9344 | 5.70 | 73 | 20 |
| 2003 | 12379 | 7.55 | 83 | 20 |
| 2004 | 13391 | 8.16 | 90 | 20 |
| 2005 | 13540 | 8.25 | 98 | 20 |
| 2006 | 13726 | 8.37 | 90 | 20 |
| 2007 | 15248 | 9.30 | 99 | 20 |
| 2008 | 14939 | 9.11 | 97 | 20 |
| 2009 | 14944 | 9.11 | 98 | 20 |
| 2010 | 15615 | 9.52 | 98 | 20 |
| 2011 | 16214 | 9.88 | 98 | 20 |
| 2012 | 10451 | 6.37 | 99 | 20 |

While the number of analysts remains about constant from 2005 onward, the number of submitted forecasts increases over time.[25] Consequently, the average number of submitted forecasts per analyst increases. This finding lends support to the view that the demand for macroeconomic forecasts increases and that broker spend more resources on the generation of these forecasts, highlighting the importance of a thorough quality analysis on the individual analyst level.

Table 3.2 Indicator Specific Summary Statistics

This table reports indicator names as well as the associated abbreviations (Abbr.), the month and year of the first observation (Start), and how many surveys are conducted for the respective indicator (Surveys), as well as the minimum, maximum and average number of participating analysts (Min, Max and Mean Analysts).

| | | | | Analysts | | |
|---|---|---|---|---|---|---|
| Indicator | Abbr. | Start | Surveys | Min | Max | Mean |
| Consumer Confidence | CC | 08/98 | 166 | 3 | 79 | 62.1 |
| ISM Index | ISM | 08/98 | 167 | 4 | 85 | 67.8 |
| Nonfarm Payrolls | NFP | 08/98 | 167 | 4 | 94 | 74.7 |
| Unemployment Rate | UN | 08/98 | 168 | 4 | 89 | 72.1 |
| Retail Sales | RS | 07/01 | 133 | 41 | 85 | 73.2 |
| Retail Sales ex Auto | RSex | 06/01 | 134 | 39 | 78 | 68.5 |
| Producer Price Index ex Food & Energy | PPIex | 09/98 | 166 | 3 | 78 | 65.6 |
| Producer Price Index | PPI | 09/98 | 165 | 3 | 81 | 68.1 |
| Consumer Price Index | CPI | 08/98 | 165 | 3 | 86 | 70.8 |
| Consumer Price Index ex Food & Energy | CPIex | 08/98 | 166 | 3 | 84 | 69.6 |
| Capacity Utilization | CU | 09/98 | 163 | 3 | 73 | 60.9 |
| Industrial Production | IP | 09/98 | 166 | 3 | 85 | 69.2 |
| Housing Starts | HS | 10/98 | 166 | 5 | 82 | 65.9 |
| Leading Economic Indicator | LEI | 07/98 | 169 | 4 | 63 | 51.8 |
| Durable Goods Orders | DGO | 07/98 | 167 | 4 | 83 | 67.5 |
| New Home Sales | NHS | 07/98 | 165 | 4 | 79 | 63.5 |
| Personal Consumption Expenditures | PCE | 07/98 | 167 | 4 | 83 | 65.8 |
| Personal Income | PI | 07/98 | 169 | 4 | 78 | 62.6 |
| Factory Orders | FO | 08/98 | 166 | 4 | 71 | 59.1 |
| Business Inventories | BI | 08/98 | 167 | 3 | 61 | 49.2 |

[25] Note that a linear extrapolation of the 2012 figures leads to 17,916 submitted forecasts.

Table 3.2 reports indicator specific sample statistics. Besides the date of the first survey (Start) it contains the number of surveys, as well as the minimum, maximum and average number of participating analysts. For all indicators our sample covers more than 10 years and the number of participating analysts varies from 3 to 94 across all indicators. Note that there are only 12 months in which the number of participating analysts takes the minimum value of 3. The first decile already includes 48 participating analysts and the first quartile 60. Accordingly, we have a sufficient number of analysts in every month.

## 3.4 Results

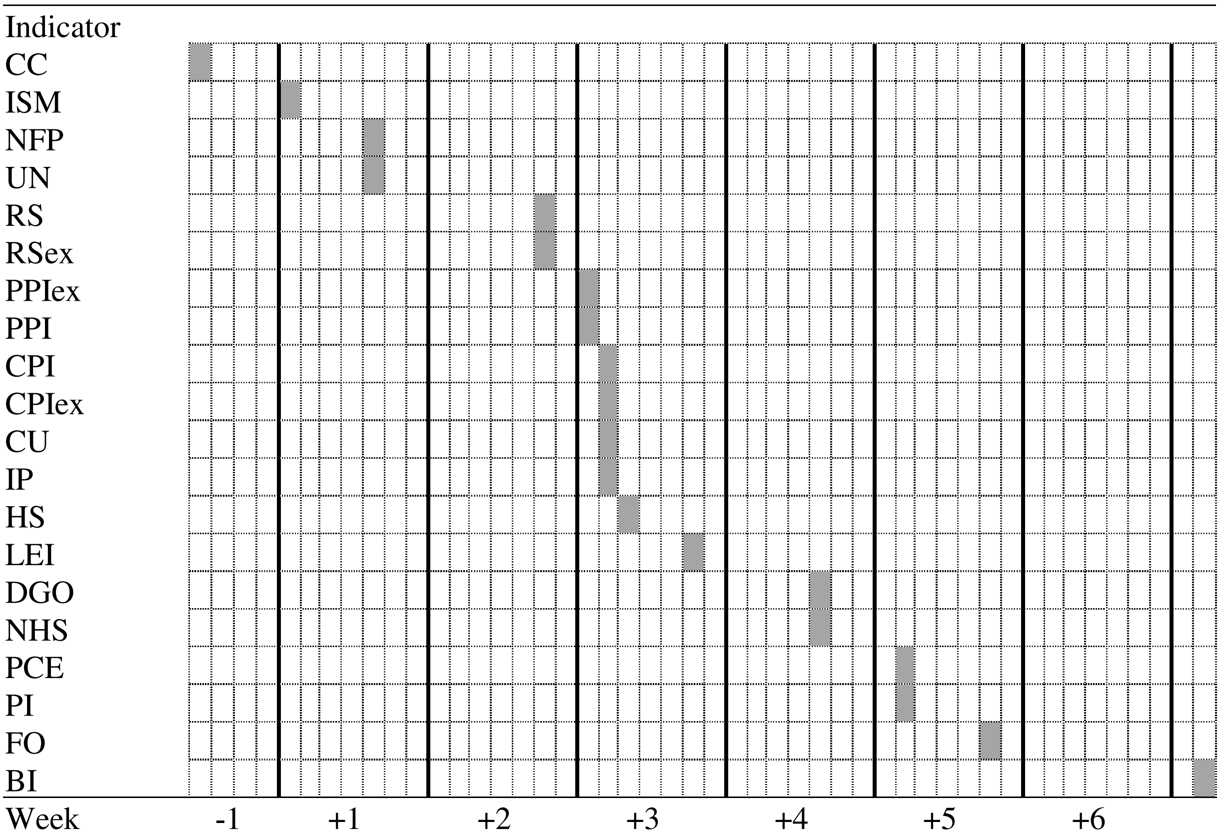
As stated in section 3.2 we expect differences in model performance because it is more difficult to predict certain indicators. Accordingly, analyst characteristics are not equally important for forecast accuracy across all indicators. To provide evidence for our hypothesis that an indicator's predictability has to be taken into account when dealing with accuracy we define our proxy measures for predictability in section 3.3 as the reciprocal of the median scaled absolute forecast error, see equation (3.2). Larger absolute forecast errors imply that a certain macroeconomic series is harder to predict, i.e., idiosyncratic predictability is low, and that the analyst can be very experienced and able, however, if the series is not or only hardly to predict these characteristics do not help him at all. For the following analysis we group indicators into quartiles according to their predictability. The classification along with indicator specific predictability is reported in Table 3.3.

One reason for low predictability might be information constrains. Indicators which are released early in the release cycle and therefore contain much new information concerning the reference month should be less predictable. Later released indicators are likely more predictable due to the earlier released information.

Figure 3.1 illustrates the monthly release cycle and classifies the median number of days between the announcement date and the end of the reference month into a weekly scheme (Timeliness). Indicators which are released in a week relatively early after the end of the reference months are therefore considered as timelier than those released many weeks after the end of the reference month.  For instance the employment report, which includes the unemployment rate and nonfarm payrolls, is released three calendar days (median) after the last day of the reference period and consequently considered a very timely macroeconomic release.

Figure 3.1 Monthly Release Cycle

For each indicator the median number of calendar days between the announcement date and the last day of the respective reference month is illustrated.




The results in Table 3.3 show a large dispersion of idiosyncratic predictability ranging from a high predictability of 0.84 for personal consumption expenditures (PCE) to low predictability of 0.39 for consumer confidence (CC). This finding provides evidence for our hypothesis that

Table 3.3 Indicator Specific Predictability and Timeliness

Column (1) reports indicators' predictability, column (2) the corresponding predictability quartile and column (3) the median number of weeks the announcement date lies behind the last day of the reference month.

| Indicator | Predictability | Quartile | Timeliness |
|---|---|---|---|
| PCE | 0.84 | 1 | 5 |
| LEI | 0.79 | 1 | 3 |
| BI | 0.79 | 1 | 7 |
| CPI | 0.76 | 1 | 3 |
| CU | 0.74 | 1 | 3 |
| PI | 0.73 | 2 | 5 |
| FO | 0.69 | 2 | 5 |
| PPI | 0.64 | 2 | 3 |
| RS | 0.61 | 2 | 2 |
| RSex | 0.60 | 2 | 2 |
| IP | 0.59 | 3 | 3 |
| DGO | 0.56 | 3 | 4 |
| PPIex | 0.55 | 3 | 3 |
| UN | 0.52 | 3 | 1 |
| CPIex | 0.52 | 3 | 3 |
| ISM | 0.50 | 4 | 1 |
| NFP | 0.50 | 4 | 1 |
| NHS | 0.47 | 4 | 4 |
| HS | 0.44 | 4 | 3 |
| CC | 0.39 | 4 | -1 |

there really is a large variation in idiosyncratic predictability. Note that the early released indicators, CC, NFP, UN and ISM, i.e., those for which basically no prior information about the reference month is available, range among the indicators with the lowest predictability. The correlation between predictability and timeliness is 0.66 and emphasizes that there is a relation between idiosyncratic predictability and the amount of information that is already released before the respective indicator is announced. However, low predictability, i.e., high median absolute forecast errors, might simply be due to missing efforts of analysts to generate precise forecasts, because they consider these indicators as unimportant. Nevertheless, the fact that such important indicators as NFP and CC are in the bottom predictability quartile rules this argument out because macroeconomic analysts have a keen interest to forecast those indicators.

Table 3.4 reports correlations as well as summary statistics of our variable set. Note that the shown correlations are for indicator and report period demeaned variables since the fixed effect estimator uses these transformed variables. We find solely significant correlations between our dependent and the independent variables except for Horizon. Furthermore, all of these correlations have the predicted sign. We find a positive correlation between AFErr and L.AFErr implying short-term persistence, i.e., small absolute forecast errors are followed by small ones on the individual analyst level. For the remaining significant variables we find negative correlations.

Table 3.4 Correlation Coefficients and Summary Statistics

Panel A reports Pearson correlation coefficients for the demeaned dependent as well as the independent variables and Panel B the corresponding summary statistics. AFErr denotes the scaled absolute forecast error, L.AFErr the absolute forecast error from the previous period, SpExp specific experience, GenExp general experience, GAbil the general ability measure, CovInd the number of covered indicators, and Horizon the forecast horizon. $^{*}$, $^{**}$, $^{***}$ indicates significance at the 10%, 5%, and 1% level, respectively.

Panel A: Correlation Coefficients

| | AFErr | L.AFErr | SpExp | GenExp | GAbil | CovInd | Horizon |
|---|---|---|---|---|---|---|---|
| AFErr | 1.00 | | | | | | |
| L.AFErr | 0.03$^{***}$ | 1.00 | | | | | |
| SpExp | -0.01$^{***}$ | -0.01$^{***}$ | 1.00 | | | | |
| GenExp | -0.01$^{***}$ | -0.01$^{***}$ | 0.96$^{***}$ | 1.00 | | | |
| GAbil | -0.09$^{***}$ | -0.05$^{***}$ | 0.05$^{***}$ | 0.05$^{***}$ | 1.00 | | |
| CovInd | -0.01$^{**}$ | -0.02$^{***}$ | 0.30$^{***}$ | 0.25$^{***}$ | 0.01$^{***}$ | 1.00 | |
| Horizon | 0.00 | 0.00 | 0.02$^{***}$ | 0.02$^{***}$ | 0.00 | 0.02$^{***}$ | 1.00 |

Panel B: Summary Statistics

| | AFErr | L.AFErr | SpExp | GenExp | GAbil | CovInd | Horizon |
|---|---|---|---|---|---|---|---|
| Mean | 1.96 | 1.69 | 43.60 | 56.00 | 0.62 | 18.37 | 7.34 |
| Median | 1.71 | 1.53 | 35.00 | 48.00 | 0.62 | 20.00 | 5.00 |
| 10$^{th}$ Percentile | 0.00 | 0.00 | 5.00 | 9.00 | 0.50 | 12.00 | 1.00 |
| 90$^{th}$ Percentile | 3.99 | 3.45 | 97.00 | 117.00 | 0.72 | 22.00 | 11.00 |

Overall correlations are low, however the results are in line with those in the equity analyst literature (e.g. Clement, 1999), except for the high correlation between specific and general experience. The high correlation of 0.96 indicates two problems. First, we face multicollinearity and the model will not be able to attribute the importance to each individual variable correctly. Second, the high correlation indicates that the vast majority of analysts predict all indicators. If this is the case, then specific and general experience increase in lockstep and there is little additional information of knowing both variables. Since we include a measure of general ability in our model we abandon the inclusion of general experience in the further analysis and keep specific experience. Consequently, our model reads:

$$\begin{aligned} AFErr_{j,i,t} &= \alpha + \beta_1 \cdot L.AFErr_{j,i,t} + \beta_2 \cdot SpExp_{j,i,t} + \beta_3 \cdot GAbil_{j,i,t} \\ &+\beta_4 \cdot CovInd_{j,t} + \beta_5 \cdot Horizon_{j,i,t} + \varepsilon_{j,i,t}. \end{aligned} \tag{3.5}$$

Before we perform the analysis on the quartile level to investigate the impact of idiosyncratic predictability we analyze the influence of analyst characteristics on the full sample to make sure that the framework is transferable to macroeconomic predictions. Table 3.5 reports results of the fixed effect estimation of Equation (3.5) for three different sample periods. Column (1) shows results for the entire sample period after an initialization period of five years as described in section 3.3. Column (2) and (3) show sub-sample results for an equally divided sample to check for robustness over time.

Results for the entire sample are predominately in line with the hypothesis stated in section 3.2. We find a positive significant relationship between absolute forecast errors in report period t and t-1 indicating short term persistence in forecast errors. Furthermore, coefficients for indicator specific experience (SpExp), general ability (GAbil), i.e., the analyst inherent characteristics are significantly negative, meaning that more specific experience and general ability lead to smaller forecast errors. Surprisingly, the exogenous characteristics, i.e., the number of covered indicators (CovInd) and the forecast horizon are insignificant. This finding

Table 3.5 Fixed Effect Regression Results

This table reports results of the fixed effects estimation according to:

$$AFErr_{j,i,t} = \alpha + \beta_1 \cdot L.AFErr_{j,i,t} + \beta_2 \cdot SpExp_{j,i,t} + \beta_3 \cdot GAbil_{j,i,t}$$
$$+\beta_4 \cdot CovInd_{j,t} + \beta_5 \cdot Horizon_{j,i,t} + \varepsilon_{j,i,t},$$

where AFErr denotes the scaled absolute forecast error of analyst j for indicator i in report month t, L.AFErr the absolute forecast error from the previous period, SpExp specific experience, GAbil the general ability measure, CovInd the number of covered indicators, and Horizon the forecast horizon. Column (1) reports results for the entire sample, i.e., 01/2003-07/2012, column (2) reports results for the sub-sample 01/2003-10/2007, and column (3) for the sub-sample 11/2007-07/2012. $R^2$ denotes the within R-squared of the regression. $^{*}$, $^{**}$, $^{***}$ indicates significance at the 10%, 5% and 1% level, respectively.

| | 01/2003-07/2012 | 01/2003-10/2007 | 11/2007-07/2012 |
|---|---|---|---|
| | AFErr | AFErr | AFErr |
| L.AFErr | 0.0219$^{***}$ | 0.0171$^{***}$ | 0.0257$^{***}$ |
| SpExp | -0.0002$^{**}$ | 0.0004$^{**}$ | -0.0004$^{***}$ |
| GAbil | -0.9902$^{***}$ | -0.9325$^{***}$ | -1.0198$^{***}$ |
| CovInd | -0.0011 | -0.0012 | -0.0030$^{***}$ |
| Horizon | 0.0000 | 0.0000 | 0.0080$^{***}$ |
| Obs | 139,470 | 66,144 | 73,326 |
| $R^2$ | 0.0080 | 0.0070 | 0.0100 |

is in sharp contrast to findings in the equity analyst literature (e.g. Clement, 1999; Clement and Tse, 2005) where especially the forecast horizon plays an important role.

While the low $R^2$ suggests that the economical significance is low, there are substantial gains applying our model. For instance, consider the comparison of two analysts' general ability and holding all other factors constant. One ranges in the bottom decile (GAbil=0.5) the other one in the top decile (GAbil=0.72).[26] Our results imply that the more able analyst has an about 0.22 lower (scaled) absolute forecast error than the less able analyst. Taking the average (scaled) absolute forecast error from Table 3.4, which is 1.96 this numerical example implies that the able analyst is on average about 11% more precise than the less able analyst and therefore the economic significance is sizable.

[26] See Table 3.4.

Results for the first sub-sample are largely in line to the overall sample results. However, we observe a positive coefficient for specific experience, which is in contrast to our hypotheses and economic intuition. For the second sub-sample all coefficients have the predicted sign and are highly significant. This finding suggests that the used variables need more than five years as initialization period to get enough cross sectional variation. Most notably, the relationship between accuracy and general ability as well as past accuracy is very robust across all periods.

To analyze the relevance of idiosyncratic predictability we perform an analysis on the quartile level according to the classification in Table 3.3. Table 3.6 reports results for the fixed effect estimation.

Table 3.6 Fixed Effect Regression Results Quartile Predictability Ranking

This table reports results of the fixed effects estimation according to:

$$AFErr_{j,i,t} = \alpha + \beta_1 \cdot L.AFErr_{j,i,t} + \beta_2 \cdot SpExp_{j,i,t} + \beta_3 \cdot GAbil_{j,i,t} + \beta_4 \cdot CovInd_{j,t} + \beta_5 \cdot Horizon_{j,i,t} + \varepsilon_{j,i,t},$$

where AFErr denotes the scaled absolute forecast error of analyst j for indicator i in report month t, L.AFErr the absolute forecast error from the previous period, SpExp specific experience, GAbil the general ability measure, CovInd the number of covered indicators, and Horizon the forecast horizon. Column (1) reports results for the indicators in the first predictability quartile according to Table 3.3 and column (2) to (4) for the second, third and fourth quartile. $R^2$ denotes the within R-squared of the regression. $^{*}$, $^{**}$, $^{***}$ indicates significance at the 10%, 5% and 1% level, respectively.

| | Quartile | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| L.AFErr | 0.0471$^{***}$ | 0.0088$^{*}$ | 0.0215$^{***}$ | 0.0071 |
| SpExp | -0.0005$^{***}$ | -0.0004$^{**}$ | 0.0000 | 0.0001 |
| GAbil | -1.3354$^{***}$ | -1.3373$^{***}$ | -0.8989$^{***}$ | -0.4969$^{***}$ |
| CovInd | -0.0053$^{***}$ | 0.0004 | -0.0010 | 0.0007 |
| Horizon | 0.0001 | 0.0000 | 0.0000 | 0.0003 |
| Obs | 31,996 | 34,929 | 36,798 | 35,747 |
| $R^2$ | 0.017 | 0.013 | 0.007 | 0.002 |

Notably, general ability appears to be the most important variable. For all quartiles we obtain highly significant coefficients. Most importantly, the impact of general ability on accuracy decreases with diminishing predictability. For the indicators which are best predictable we observe a regression coefficient of -1.33 compared to about -0.5 for the least predictable indicators. This finding implies that analysts' general ability is more important for indicators which are idiosyncratically predictable because for these indicators the forecast task is feasible at all. For the generally unpredictable series analysts even do not benefit from their general ability simply because the forecast task is almost impossible and individual forecast performance is unsystematic. Combining the regression coefficients with the respective bottom and top decile general ability figures[27] translate into about 20% more precise predictions for the most able analysts compared to the less able ones for the indicators included in the first predictability quartile compared to approximately 5% for those indicators in the fourth quartile. The huge difference in economic significance highlights the importance of a proper differentiation based on idiosyncratic predictability. Accordingly, this finding provides strong evidence for our hypothesis that model performance heavily depends on idiosyncratic predictability.

Furthermore, we find that forecast quality does only benefit from indicator specific experience if the indicator ranges in the two upper quartiles of predictability. For both quartiles with the lowest predictability, indicator specific experience has no effect on forecast accuracy. This finding further supports the notion that predictability has to be taken into account. If an indicator is not predictable than indicator specific experience does not help to improve forecast quality, because accuracy is not systematically related to this characteristic. For the fourth quartile even the previous month's accuracy loses its significance, i.e., there is no short term persistence of accuracy. This again highlights that it is very difficult to systematically

[27] For details see Appendix 3 A.

generate superior forecasts for these macroeconomic indicators. Lastly, only for the first quartile our measure for portfolio complexity, the number of covered indicators, is significant. A larger set of covered indicators consequently implies lower forecast errors for only those indicators which are classified as highly predictable lending support to the notion that it pays off for analysts to have a sound understanding of the entire economy and the interrelations between different macroeconomic indicators.

There is not only a pattern in model performance in terms of the impact of certain characteristics on accuracy, but also in terms of explanatory power. Although R-squared is low compared to results in the equity analyst research literature, there are large differences depending on predictability. While R-squared for the entire sample is 0.8% (see Table 3.5) it varies between 0.2% and 1.7% from the fourth to the first predictability quartile. Consequently, also on a low level, the explanatory power varies considerably. However, despite low level of R-squared economic significance, i.e., sizable forecast accuracy improvements exist as shown before.

If an indicator is idiosyncratically not predictable, even the best analysts do not perform better than an average analyst. Consequently, there is no differentiation between good and bad analysts and the characteristics model, i.e., Equation (3.5) cannot detect differences. If on the other hand predictability is high, there is competition among macroeconomic analysts and above average analysts generate superior forecasts.

## 3.5 Conclusion

Evidence concerning accuracy differences among macroeconomic analysts is mixed. Furthermore, the association between accuracy differences and characteristics of macroeconomic analysts has not been analyzed so far. Moreover, neither the studies analyzing macroeconomic forecast accuracy differences, nor those analyzing equity analysts' earnings per share forecasts take the idiosyncratic predictability of the forecast objective into account.

We analyze the relation between an indicator's idiosyncratic predictability, macroeconomic analyst characteristics and forecast accuracy differences. The main idea is that analysts' forecast performance is rather random than systematic if the respective macroeconomic indicator is idiosyncratically unpredictable, because more skilled analysts do not benefit from their superior skills. For those series which are predictable, superior skills translate into more accurate forecasts.

Our results provide evidence that there is some degree of heterogeneity among macroeconomic analysts and that their forecast performance differs. On the one hand, we find only a limited advantage of following certain analysts for macroeconomic series with low idiosyncratic predictability. The characteristics model has little explanatory power for these indicators and analyst characteristics are more or less immaterial. Only if the forecast objective is idiosyncratically predictable analyst characteristics, especially the inherent ones, contribute to forecast accuracy and the economic significance of our results is sizable. Consequently, for those indicators it is worth to choose an analyst based on the characteristics and method used in our analysis.

Our results have implications for researchers and users of macroeconomic as well as equity analysts' forecasts. Macroeconomic forecast quality differs across analysts, however only for series idiosyncratically predictable these differences are detectable and worth the effort to choose the corresponding analysts. The question arises whether the idiosyncratic predictability of company earnings plays a comparable role for model performance analyzing equity analysts' accuracy. Only if company earnings have a certain degree of idiosyncratic predictability skilled analysts are able to outperform less skilled ones. Consequently, it is a fruitful approach for future research to analyze whether analysts' accuracy differences are more pronounced for those companies whose earnings are idiosyncratically more predictable.

# Chapter 4 Do Aggregate Company Outlooks have Macroeconomic Content?

## 4.1 Introduction

The state determination of the economy and the prediction of future macroeconomic developments are essential for many purposes, e.g., to guide monetary policy decisions or decisions about federal tax cuts to stimulate economic growth. In this context coincident and leading macroeconomic indicators play an important role. Economic agents rely on them to gauge how the economy is doing and how it will likely do in the future. To measure economic activity these indicators incorporate a wide information set of macroeconomic variables, such as industrial production and capital market related measures, such as interest rate spreads and stock returns. Company expectations as a valuable source of information to forecast overall economic activity, however, are not directly included. To a substantial extent the economy consists of firms listed on the stock market. About one-third of all employees in the U.S. private business sector work at publicly traded companies (Davis, 2006). Therefore, the development of the overall economy is significantly related to the development of these firms as a whole. In this paper we show that aggregated company outlooks predict overall economic activity, i.e., combined forecasts of individual parts of the economy predict the entire economy. As a proxy for individual company level outlooks we use analysts' stock recommendations, since analysts have direct access to company specific information and therefore their information set is arguably the best available proxy for company specific forecasts.

At the firm level, company managers must make various decisions that are of great importance for the company such as investment decisions, financing decisions, employment decisions etc. Moreover, managers plan their actions ahead by setting specific goals. For example, the decision of managers to lay off a significant amount of the company's workforce in the next couple of months to cut expenses is a crucial decision for the future development of the company. Nevertheless, decisions at the firm-level are ceteris paribus not important for the overall economy. However, aggregated projected layoffs over all companies, which in sum constitute a significant amount of the entire economy, are basically a crucial input to forecast future employment figures. The same argumentation holds for other company specific information such as orders, production, and capacity utilization. Every company outlook includes growth perspectives and therefore information whether the management assumes a more optimistic or pessimistic development of the company over the next months and years. Consequently, aggregated expectations over all companies contain predictive power for future macroeconomic developments.

Observing and predicting the entire economy is similar to observing all companies in the economy simultaneously and aggregating these observations. Unfortunately, we cannot observe managers' expectations directly and a survey of all companies is unfeasible for several reasons (e.g., cost, timing etc.). However, stock analysts collect and process all kinds of information at the firm level for publicly traded companies. Besides public macroeconomic-, industry- and company-specific content their information sets should also comprise non-public company-specific information (Grossman and Stiglitz, 1980). By attending analysts' conferences, management meetings, and telephone conferences, stock analysts develop a sound understanding of the company's business and its future performance. Due to their extensive company knowledge their information set is arguably the best available proxy measure for company specific outlooks. In addition to public company information it presumably contains company-specific information that is not public knowledge. Analysts issue stock recommendations that indicate a general direction of the development of the underlying company: A positive recommendation suggests in general a relatively positive development, while a negative recommendation suggests in general a relatively negative development. We show that aggregating these signals over all companies listed on the stock market allows conclusions about the overall economy.

The existing literature provides support for our new approach to combine macroeconomic developments with stock analysts' outputs. Previous literature indicates that analyst recommendations have predictive power for stock returns on a firm-level, i.e., analysts possess a unique company specific information set. Womack (1996) shows that recommendations lead to a significant post-announcement drift up to 6 months. Barber et al. (2001) demonstrate that a trading strategy based on recommendations yields significant excess returns before transaction costs. Howe et al. (2009) find that aggregated recommendations predict future excess returns on a market level. Stock returns in turn have predictive power for macroeconomic activity. Stock and Watson (1998) provide evidence that the stock market leads the real economy. Consequently, it is reasonable to analyze the predictive power of analysts' recommendations for macroeconomic developments. If analysts' outputs can forecast stock market developments and the stock market anticipates macroeconomic developments, then analysts' recommendations should also contain predictive power for the real economy.

To the best of our knowledge, we provide the first study to show this link. To do so we use monthly changes in aggregated analysts' recommendations as a proxy for changes in aggregated company outlooks. We evaluate the predictive power of company outlook changes for future macroeconomic developments by using a regression framework in which we analyze the relation of recommendation changes to changes in a broad measure of economic activity. Our results provide evidence that aggregated company outlooks have predictive power for future macroeconomic developments of about one year. Our results remain valid when we control for other well-known macroeconomic predictors, such as the term-spread and the dividend yield. We find that the predictive power of aggregated company outlooks is not included in the leading economic indicator (LEI) – which specifically has been developed to forecast economic activity – and other well-known macroeconomic predictors.

Overall, we show that aggregated company outlooks are an important macroeconomic variable that has been overlooked so far when forecasting economic developments. This result might be especially of interest to agencies which release economic indicators as well as to investors who might incorporate this information to form better expectations of upcoming overall economic activity. Our results also suggest that analysts implicitly or explicitly anticipate macroeconomic developments when processing information in their stock recommendations. The change in the aggregated monthly consensus has predictive power for the upcoming economic development up to about one year. Indicating that analysts have first-hand information about the individual firms they cover, our findings therefore also support the theoretical notion that analysts play an important role as financial intermediaries (Grossman and Stiglitz, 1980). Moreover, our results indicate a link between previous studies: Stock and Watson (1998) show that the stock market leads the real economy, while Howe et al. (2009) provide evidence that aggregated analysts' recommendations predict the overall stock market. Our results fill the missing link between both studies, since we find that aggregated analysts' recommendations also predict the real economy. We provide evidence that changes in expectations about future firm performance rationally (i.e., correct on average) determine asset values before overall economic activity changes. Therefore, our results provide a potential explanation, why the stock market leads the real economy.

The remainder of the paper is organized as follows. In Section 4.2, we introduce the research design. Section 4.3 describes the data. Section 4.4 discusses the empirical results. Finally, section 4.5 concludes.

## 4.2 Research Design

We consider the following general model to forecast overall macroeconomic activity:

$$Economic\,Activity_{t,t+\tau} = \alpha + \beta \cdot Aggregated\,Company\,Outlooks_t + \varepsilon_t. \qquad (4.1)$$

Unfortunately, company outlooks cannot be observed directly. We argue that a good proxy for company outlooks are assessments of analysts who cover the companies and therefore process company expectations. Sell-side analysts aggregate a vast amount of information on the company level in their role as information intermediaries on financial markets. Since analysts possess a first-hand information set about the covered companies, it is straightforward to approximate companies' expectations with analysts' outputs. Our approach is supported by the theoretical notion that analysts' outputs should contain more than just public knowledge (Grossmann and Stiglitz, 1980).

However, since analysts issue two key outputs, namely earnings per share (EPS) forecasts and stock recommendations, the proxy measure's choice should be carefully considered. We decide to focus on stock recommendations[28] for the following reasons. First, revisions of earnings forecasts might be based on stale information, while revisions of stock recommendations are not (Jegadeesh and Kim, 2010). Second, recommendations contain more information than earnings forecasts. A recommendation reflects the upcoming yearly earnings, as well as all future earnings, the payout ratios and discount rates. Therefore, to determine the value of a company the analyst needs interest rates, risk premiums and growth expectations as inputs in addition to EPS forecasts. These are basically the same information the company manager has to consider when making investment decisions for the firm.

[28] As an alternative measure, we analyze the predictive content of changes in aggregate earnings forecasts. The results are discussed in section 4.4.

Consequently, solely using EPS forecasts is not sufficient because a substantial proportion of the information would be neglected. Third, a measure with "natural" boundaries is preferred to identify significant expectation changes.[29] While small changes of EPS forecasts might be due to minor changes in expectations, a recommendation revision is the result of major changes and therefore it is a clear interpretable signal that the companies' prospects changed significantly. Fourth, it is desirable to have a proxy measure undistorted by the impact of accounting policies and accounting standard changes. A company is easily able to influence, to a certain degree, their earnings per share by applying the accounting standards in different ways ("earnings management"). Therefore, analysts have to adjust their earnings forecasts even without a change in the underlying fundamentals. Moreover, our model would receive an incorrect signal if accounting standards are changed and as a consequence thereof all EPS forecasts change without fundamental reason, e.g., when stock option compensation is treated differently from an accounting perspective. Fifth, we avoid a "forecast horizon problem", since recommendations are always valid for a certain future time period (usually up to one year if not revised) independent of the point in time of the fiscal year they are issued. In contrast, the forecast horizon of EPS forecasts is determined by the end of a firm's fiscal year. For example, an earnings forecast issued on October, 1st, for a firm with a fiscal year end at December, 31st, has a forecast horizon of three months. An earnings forecast for the same firm issued one month later would have a forecast horizon of two months. Recommendations are easier to compare and to aggregate, since they are based on expectations for stable forecast periods. Following these considerations our main proxy measure for company outlooks are analysts' recommendations.

[29] In I/B/E/S all recommendations are coded as integers between 1 ("Strong Buy") to 5 ("Strong Sell").

We are aware of the fact that analysts' recommendations might be biased, for instance that they are overoptimistic.[30] However, assuming that their incentive-driven overoptimism is constant over time this does not constitute a problem. Since we use recommendation changes, the positive bias cancels out.

Our dependent variable in Equation (4.1), the measure for economic activity, is the 3-month moving average of the Chicago Fed National Activity Index (CFNAI) and the 3-month average of its four components.[31] Based on the methodology of Stock and Watson (1999b) the index is a weighted average of 85 macroeconomic series from four categories: 23 series from production and income, 24 series from employment, unemployment, and hours, 15 series from personal consumption and housing, and 23 series from sales, orders, and inventories. The four categories' contributions are for (1) Production and Income (PI) about 33%, (2) Employment, Unemployment, and Hours (EUH) 32%, (3) Personal Consumption and Housing (CH) 14% and (4) Sales, Orders and Inventories (SOI) 21%. The CFNAI is a "single summary measure of a common factor" (Federal Reserve Bank of Chicago, 2012) in these series which is released monthly. It is constructed to have an average value of zero and a standard deviation of one. A positive value of the CFNAI implies economic growth above the historical trend, while a negative value corresponds to growth below the trend. The three month moving average of the CFNAI has desirable properties concerning the state determination of the economy. Compared to other measures, such as the GDP, the timely availability of the CFNAI provides valuable information about the state of the economy almost in real time on a monthly basis.[32]

---

[30] See for instance Jegadeesh et al. (2004).

[31] For information on the CFNAI see www.chicagofed.org/webpages/research/data/cfnai/current_data.cfm.

[32] For a quality evaluation of the CFNAI see for instance Evans et al. (2002).

Our analysis proceeds in 4 steps. First, for each month we calculate the consensus of stock recommendations issued in the respective month. In contrast to Howe et al.(2009), we solely use recommendations issued in the respective month and not all recommendations outstanding in the respective month to remove stale recommendations and to focus solely on the most recent information. In a second step, to measure the changes of managers' expectations we calculate the monthly change of the consensus ($rec\_consensus_t$) in month t $\Delta rec_t$. We define the monthly consensus change in month t as:

$$\Delta rec_t = \frac{rec\_consensus_t - rec\_consensus_{t-1}}{rec\_consensus_{t-1}} \tag{4.2}$$

The monthly consensus change in t is the percentage change in the consensus of all stock recommendations issued in month t compared to month t-1. In the third step we quantify the in month t unknown upcoming change of the economic indicator ($\Delta CFNAI$) between t and $t + \tau\left(\tau = 1,...,T\right)$ for different future time periods. Finally, in the fourth step, we perform a regression of the change of the economic indicator on the recommendation change, several lags of the dependent variable and various controls as specified in Equation (4.3):

$$\Delta CFNAI_{t,t+\tau} = \alpha + \beta_1 \cdot \Delta rec_t + \sum_{n=2}^{N} \beta_n \cdot control\,variable_t^n + \varepsilon_t,\ \forall\,\tau = 1,...,T \tag{4.3}$$

where all variables are specified as previously described and $\varepsilon_t$ denotes the error term.

We estimate the model in different settings. First, we estimate the predictive power of a recommendation change for the future CFNAI development alone. In a second setting, to control for other common variables which possess macroeconomic prediction power, we specifically control for changes in the leading economic indicator ($\Delta lei$), market excess returns (mex), dividend yield (divy), default spread (defs), term spread (tes), 3-month T-Bill rate (3mtb), and three lags of the dependent variable. Third, we analyze the predictive power of the recommendation change for the four CFNAI components (IP, EUH, CH, and SOI) separately. Fourth, we only use stock recommendations issued for S&P 500 firms. Fifth, as an alternative measure for company outlooks, we evaluate the predictive power of changes in aggregate earnings forecasts.

## 4.3 Data

Table 4.1 Summary Statistics

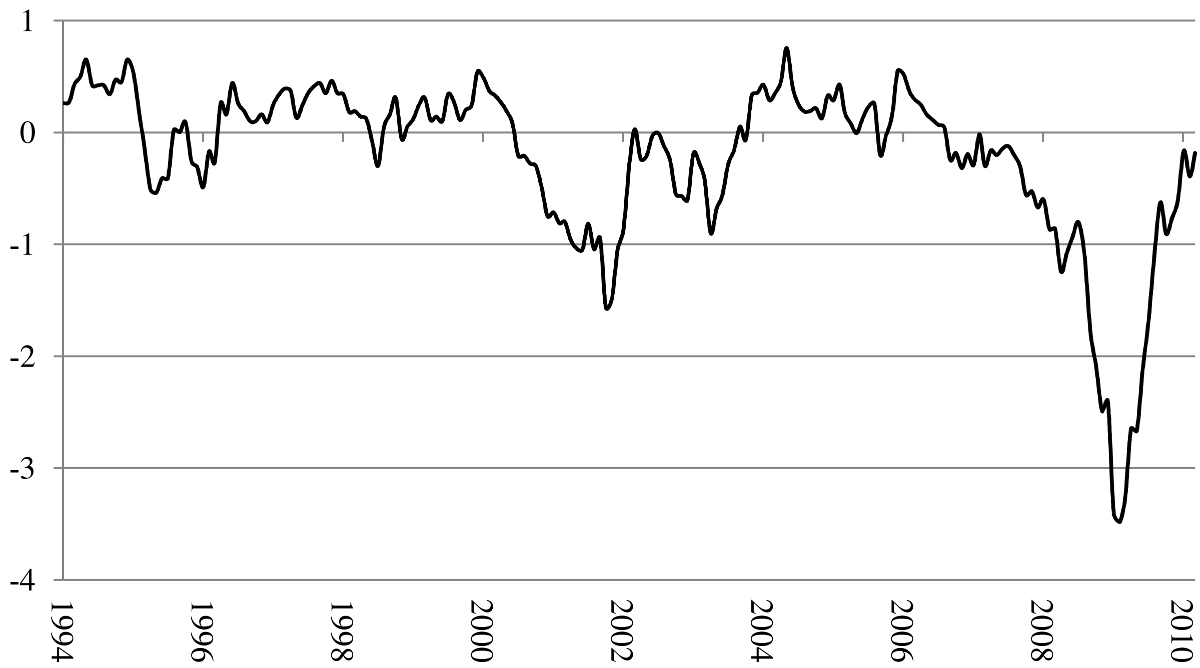
This table reports the abbreviations, units as well as summary statistics, i.e., the mean ($\mu$), standard deviation ($\sigma$), as well as minimum and maximum values of our set of variables. These include the Chicago Fed National Activity Index (CFNAI) and the changes of the respective CFNAI components Production and Income (PI), Employment, Unemployment and Hours (EUH), Personal Consumption and Housing (CH), and Sales, Orders, and Inventories (SOI), the Recommendation Change, the recommendation change if only companies in the S&P 500 are considered (Reco. Change (S&P 500)), the Leading Economic Indicator (LEI), monthly market excess returns calculated as the difference of the monthly market index return and the monthly risk-free rate including all NYSE, AMEX and NASDAQ firms (Market Excess Return), the Dividend Yield, the Default Spread (calculated as the difference between Moody's Baa corporate bond yield and the 10-year Treasury constant maturity rate), the 3-months T-Bill rate, the Term Spread (calculated as the difference between the 10-year Treasury constant maturity rate and the 3-month T-Bill rate), as well as the number of submitted recommendations (monthly Recommendations).

| Variable | Abbr. | Unit | $\mu$ | $\sigma$ | Min | Max |
|---|---|---|---|---|---|---|
| CFNAI | | Change | -0.0023 | 0.2293 | -1.0074 | 0.6226 |
| CFNAI IP | | Change | 0.0003 | 0.1245 | -0.5467 | 0.3333 |
| CFNAI EUH | | Change | -0.0008 | 0.0775 | -0.2934 | 0.2576 |
| CFNAI CH | | Change | -0.0020 | 0.0211 | -0.1058 | 0.0845 |
| CFNAI SOI | | Change | 0.0003 | 0.0323 | -0.1046 | 0.1433 |
| Recommendation Change | $\Delta rec$ | % Change | 0.0002 | 0.0241 | -0.0733 | 0.0629 |
| Reco. Change (S&P 500) | $\Delta rec^{S\&P50}$ | % Change | 0.0000 | 0.1109 | -0.3456 | 0.4194 |
| Leading Economic Indicator | LEI | % Change | 0.0022 | 0.0053 | -0.0147 | 0.0154 |
| Market Excess Return | mex | % Change | 0.41 | 4.67 | -18.54 | 11.04 |
| Dividend Yield | divy | % | 1.87 | 0.53 | 1.11 | 3.60 |
| Default Spread | defs | % | 2.31 | 0.89 | 1.29 | 6.01 |
| 3-months T-Bill rate | 3mtb | % | 3.46 | 1.87 | 0.03 | 6.17 |
| Term Spread | tes | % | 1.64 | 1.18 | -0.53 | 3.70 |
| Monthly Recommendations | | # | 2572 | 644 | 1812 | 7709 |

Our sample ranges from January 1994 to March 2010. Table 4.1 shows means $(\mu)$, standard deviations $(\sigma)$ as well as minimal and maximal values of the variables used in the analysis. We apply the Phillips–Perron test to verify that the time series data do not contain a unit-root and we do not find evidence for non-stationarity.[33]

We obtain vintage data (whenever available) for the CFNAI and its four components[34] from the Federal Reserve Bank of Chicago to adequately describe the available information about the state of the economy at every single point of time.[35] Figure 4.1 shows the CFNAI development from January 1994 to March 2010.

Figure 4.1 CFNAI Development




Analyst recommendations are obtained from the Institutional Brokers' Estimate System (I/B/E/S). We use the following filters for the recommendations: (1) we only keep

[33] Results are reported in Appendix 4 A.

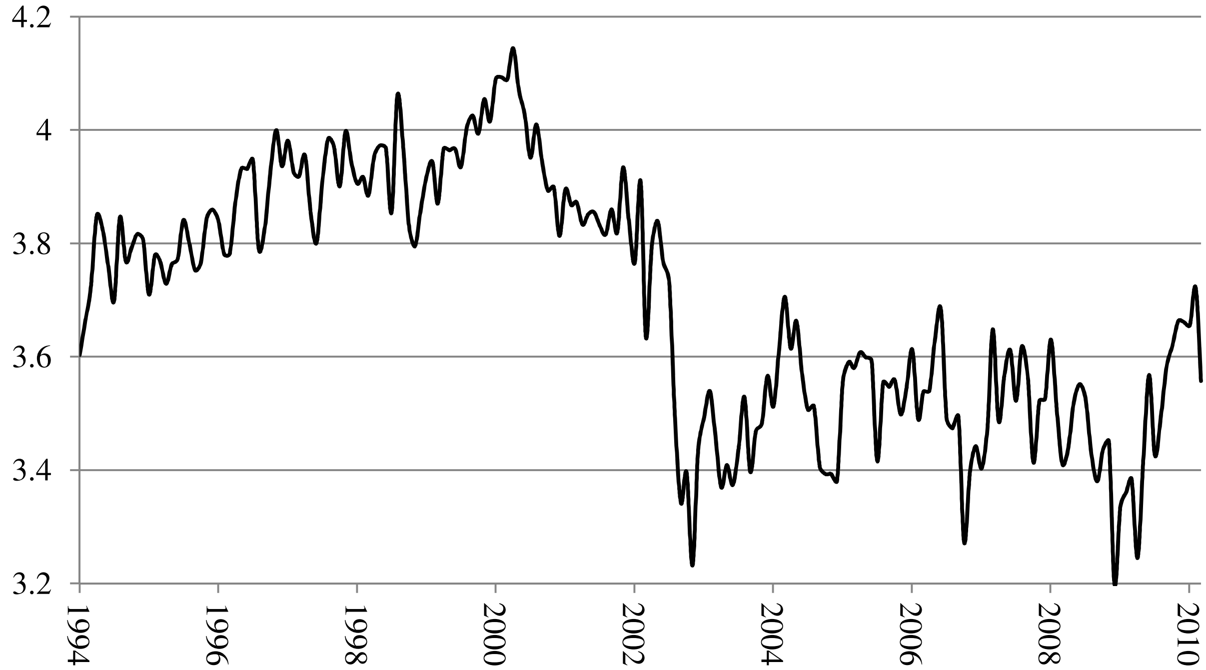
[34] We use the 3-month moving average of the CFNAI and its components.

[35] Federal Reserve Bank of Chicago (2012).

recommendations for US-American firms to be consistent with the CFNAI measure, and (2) we drop 696 recommendations from our sample in order to control for obvious data base errors.[36] Our final sample consists of 499,571 stock recommendations.

We calculate the monthly aggregated mean analyst recommendation with the recommendations issued in the respective month. On average we consider 2,575 recommendations issued per month. Figure 4.2 visualizes the aggregated analyst recommendations from January 1994 to March 2010.

Figure 4.2 Average Consensus Recommendation




In the I/B/E/S database the recommendations are coded as follows: 1= Strong Buy, 2 = Buy, 3 = Hold, 4 = Sell, and 5 = Strong Sell. In line with previous literature (e.g., Jegadeesh and Kim, 2010) we reverse the ratings to facilitate a more intuitive interpretation. A Strong Buy receives the values of 5, a Buy a value of 4, a Hold a value of 3, a Sell a value of 2, and a

[36] Specifically, we drop recommendations that have been issued by the same analyst from the same broker on the same day for the same firm with the same rating.

Strong Sell a value of 1. Therefore, a higher consensus recommendation implies a more optimistic analyst outlook. Figure 4.2 illustrates that the overall recommendation level varies significantly in the analyzed 1994 to 2010 time period. Jagadeesh and Kim (2010) point out that the mean recommendation level dropped significantly in 2002 when analysts were alleged of being overly optimistic by the New York State Attorney General's office.

Data about the Leading Economic Index are obtained from Reuters. Monthly market excess returns are calculated by using the difference of the monthly market index return and the monthly risk-free rate including all NYSE, AMEX and NASDAQ firms. We obtain the data for monthly market excess returns from the website of Kenneth French.[37] We retrieve the monthly dividend yield from the homepage of Robert Shiller.[38] The 3-month T-Bill rate, the default spread (calculated as the difference between Moody's Baa corporate bond yield and the 10-year Treasury constant maturity rate), and the term spread (calculated as the difference between the 10-year Treasury constant maturity rate and the 3-month T-Bill rate) are obtained from the Federal Reserve Bank of St. Louis (2011).

## 4.4 Results

Table 4.2 shows the results of the univariate regression of the recommendation change on the upcoming CFNAI development. The column "Base Model" shows coefficient estimates of the model without any control variables. The results provide strong evidence that aggregate company level information has economic content. The aggregated stock recommendations have predictive power for the upcoming economic activity up to 7 months. Improving company outlooks, i.e., better recommendations ($\Delta rec > 0$) imply positive changes of overall economic activity as measured by CFNAI changes.

[37] French (2011).

[38] Shiller (2011).

Table 4.2 CFNAI Developments and Changing Company Outlooks

This table reports results from the following model:

$$\Delta CFNAI_{t,t+\tau} = \alpha + \beta_1 \cdot \Delta rec_t + \sum_{n=2}^{N} \beta_n \cdot control\, variable_t^n + \varepsilon_t, \quad \forall\, \tau = 1,...,T\,,$$

where the "Base Model" β is estimated without control variables and the "Control Model" β is estimated with control for changes in the leading economic indicator ($\Delta lei$), market excess returns (mex), dividend yield (divy), default spread (defs), term spread (tes), 3-month t-bill rate (3mtb), and three lags of the dependent variable. Inference is based on Newey-West Standard Errors. *, **, *** indicates significance at the 10%, 5% and 1% level, respectively.

| Horizon | Base Model $\beta_1$ | Control Model $\beta_1$ |
|---|---|---|
| 1 | 2.250*** | 2.042** |
| 2 | 2.879*** | 2.635** |
| 3 | 2.981** | 3.294*** |
| 4 | 2.207* | 2.557** |
| 5 | 2.233* | 2.651* |
| 6 | 2.564* | 2.720* |
| 7 | 3.102** | 3.539** |
| 8 | 2.005 | 2.610 |
| 9 | 2.516 | 2.848* |
| 10 | 1.934 | 2.469 |
| 11 | 1.276 | 2.953* |
| 12 | 1.192 | 2.066 |

In a second setting ("Control Model"), we control for well-established macroeconomic variables as discussed in Section 4.2 to ensure that our measure is not just a summary measure of well-known variables which possess macroeconomic prediction power. We use changes in the leading economic indicator ($\Delta lei$),[39] market excess returns (mex), dividend yield (divy), default spread (defs), term spread (tes), 3-month t-bill rate (3mtb), and three lags of the dependent variable as control variables. The correlation matrix between the CFNAI changes and the control variables can be found in Appendix 4 B.

[39] The LEI combines 10 single series with predictive power.

When controlling for macroeconomic variables that are well-known predictors of economic activity the aggregated recommendation change is still significant. Except for the 1- and 2-months horizons the coefficients are even larger. The coefficients of the aggregated recommendation change predict the upcoming 1-month to 11-months changes of economic activity. The results show that our measure contains unique information not captured in the 4 macroeconomic control variables as well as in the LEI, which was specifically designed to predict economic activity. While the control variables contain information about growth (LEI and dividend yield), risk premiums (market excess returns and default spreads) as well as interest rates (3-months t-bill rate and term spread) our results suggest that aggregate company outlooks contain even more information about the development of the economy.

The results are economically significant. For example, the coefficient for the 1-month forward CFNAI change is 2.042, while the standard deviation of the recommendation change is 0.0241. Therefore, a one standard deviation increase in the recommendation change suggests a change in the CFNAI for the upcoming month of 0.05. This is relatively high since the CFNAI has by construction a mean of zero and a standard deviation of one.

Next, we analyze the predictive power of the recommendation change for the four individual CFNAI components. (1) Production and Income (PI) contributes about 33%, (2) Employment, Unemployment, and Hours (EUH) 32%, (3) Personal Consumption and Housing (CH) 14%, and (4) Sales, Orders and Inventories (SOI) 21% to the CFNAI.[40] We use the same control variables as for the analysis of the whole CFNAI. The results are reported in Table 4.3.

[40] The weights are readjusted monthly. The shifts in weights are rather small, see Federal Reserve Bank of Chicago (2012b).

Table 4.3 CFNAI Components Development and Changing Company Outlooks

This table reports results from the following model:

$$\Delta CFNAI\_COMPONENT_{t,t+\tau} = \alpha + \beta_1 \cdot \Delta rec_t + \sum_{n=2}^{N} \beta_n \cdot control\,variable_t^n + \varepsilon_t, \quad \forall \tau = 1,...,T,$$

where the "Control Model" β for the respective CFNAI Components Production and Income (PI), Employment, Unemployment and Hours (EUH), Personal Consumption and Housing (CH), and Sales, Orders, and Inventories (SOI) is estimated with control for changes in the leading economic indicator ($\Delta lei$), market excess returns (mex), dividend yield (divy), default spread (defs), term spread (tes), 3-month t-bill rate (3mtb), and three lags of the dependent variable. Inference is based on Newey-West Standard Errors. $^{*}$, $^{**}$, $^{***}$ indicates significance at the 10%, 5% and 1% level, respectively.

| | Control Model PI | Control Model EUH | Control Model CH | Control Model SOI |
|---|---|---|---|---|
| Horizon | $\beta_1$ | $\beta_1$ | $\beta_1$ | $\beta_1$ |
| 1 | $1.084^{**}$ | 0.121 | $0.118^{*}$ | -0.0118 |
| 2 | $1.814^{***}$ | 0.341 | 0.121 | 0.0841 |
| 3 | $2.063^{***}$ | $0.967^{**}$ | 0.179 | 0.1255 |
| 4 | $1.112^{*}$ | $0.911^{**}$ | 0.102 | $0.228^{**}$ |
| 5 | 1.034 | $1.197^{**}$ | 0.222 | 0.1541 |
| 6 | $1.533^{**}$ | $1.051^{**}$ | 0.202 | $0.237^{**}$ |
| 7 | $2.142^{***}$ | $1.099^{***}$ | 0.239 | $0.202^{*}$ |
| 8 | $1.349^{**}$ | $1.055^{*}$ | 0.153 | $0.218^{*}$ |
| 9 | 1.083 | $0.950^{*}$ | 0.201 | $0.261^{**}$ |
| 10 | 0.864 | $1.209^{**}$ | 0.215 | $0.248^{**}$ |
| 11 | $1.499^{**}$ | $1.220^{**}$ | 0.177 | $0.260^{**}$ |
| 12 | 0.918 | $1.280^{**}$ | 0.114 | 0.157 |

The results show that the predictive power of the recommendation change varies drastically for the four CFNAI components. The size of the coefficients is directly comparable. We find moderate predictive power for PI and EUH, relatively low predictive power for the SOI component, and no predictive power for CH. These results correspond to the weights of the individual series in the CFNAI and therefore to the relative importance for economic activity. Changes in company outlooks have less predictive power for individual measures of economic activity than for the CFNAI as an aggregate measure of economic activity.

To further evaluate the properties of aggregated company outlooks we restrict the sample to firms listed on the S&P 500. We again perform the previous analysis, however, the aggregated recommendation change is calculated only for S&P 500 firms. Using only outlooks of a part of the entire economy the number of observed companies decreases and consequently the number of issued stock recommendations also declines to an average of 552 recommendations per months compared to 2,575 recommendations in the base case where the entire economy is considered. The results are shown in Table 4.4.

Table 4.4 CFNAI Development and Changing Company Outlooks S&P 500 only

This table reports results from the following model:

$$\Delta CFNAI_{t,t+\tau} = \alpha + \beta_1 \cdot \Delta rec_t^{S\&P500} + \sum_{n=2}^{N} \beta_n \cdot control\, variable_t^n + \varepsilon_t, \quad \forall\, \tau = 1,...,T,$$

where the "Base Model" β is estimated without control variables and the "Control Model" β is estimated with control for changes in the leading economic indicator ($\Delta lei$), market excess returns (mex), dividend yield (divy), default spread (defs), term spread (tes), 3-month t-bill rate (3mtb), and three lags of the dependent variable. Only S&P 500 firms are considered in the consensus recommendation calculation. Inference is based on Newey-West Standard Errors. $^{*}$, $^{**}$, $^{***}$ indicates significance at the 10%, 5% and 1% level, respectively.

| Horizon | Base Model $\beta_1$ | Control Model $\beta_1$ |
|---|---|---|
| 1 | 1.655$^{**}$ | 1.755$^{**}$ |
| 2 | 1.709$^{**}$ | 2.177$^{**}$ |
| 3 | 1.4535 | 2.430$^{**}$ |
| 4 | 1.2018 | 2.313$^{**}$ |
| 5 | 1.695$^{*}$ | 2.774$^{**}$ |
| 6 | 1.914$^{*}$ | 2.963$^{**}$ |
| 7 | 1.9161 | 3.151$^{**}$ |
| 8 | 0.8962 | 2.342$^{*}$ |
| 9 | 0.6508 | 1.9422 |
| 10 | 1.4107 | 2.713$^{**}$ |
| 11 | 0.4665 | 2.764$^{**}$ |
| 12 | 0.9074 | 2.727$^{**}$ |

The results point out that our general results are somewhat weaker when using only the firms listed on the S&P 500. The aggregated recommendation change predicts economic activity up to 11 months ("Control Model").[41] The coefficient is smaller for 8 out of 12 months. The results indicate that using outputs from a larger sample leads to more predictive power than using only outlooks of companies listed in the S&P 500. However, since the largest publicly traded companies are included in the S&P 500 the economic difference is rather small.

Overall, our results show that aggregated company outlooks proxied by aggregated stock recommendations are an important macroeconomic predictor. The information set included in the aggregated company outlooks is not reflected in other established macroeconomic variables making it a valuable predictor for macroeconomic developments.

Next, we analyze aggregated company outlooks proxied by aggregated earnings forecasts as an alternative measure for company outlooks. Each month we calculate the median earnings per share forecast based on earnings forecasts issued in the respective month. In order to control for different forecast horizons we calculate the median earnings per share forecast separately for forecasts with a forecast horizon of less or equal to 3 months, less or equal to 6 months, less or equal to 12 months and greater than 12 months. Since earnings per share forecasts show seasonal patterns we calculate the monthly change in the earnings forecasts relatively to the aggregate median earnings forecast twelve months ago. Finally, we apply the Baxter-King filter[42] with a minimum period of oscillation of 18 and a maximum oscillation period of 96 corresponding to monthly date. The calculation of the monthly change in aggregated earnings per share forecasts is shown below:

$$\Delta EPS_t = \frac{Median\_EPS_{t,Horizon} - Median\_EPS_{t-12,Horizon}}{Median\_EPS_{t-12,Horizon}} \tag{4.4}$$

---

[41] Results for the 12-month horizon should be handled with care due to non-stationarity issues according to the Phillips–Perron test.

[42] Baxter and King (1999).

Estimation results are shown in Table 4.5:

Table 4.5 CFNAI Developments and Changing Median Earnings Forecasts

This table reports results from the following model:

$$\Delta CFNAI_{t,t+\tau} = \alpha + \beta_1 \cdot \Delta EPS_t + \sum_{n=2}^{N} \beta_n \cdot control\,variable_t^n + \varepsilon_t, \quad \forall \tau = 1,...,T,$$

where the "Control Model" β is estimated with the respective EPS change (horizon is equal to 3 months, 6 months, 12 months and above 12 months) and controlling for changes in the leading economic indicator ($\Delta lei$), market excess returns (mex), dividend yield (divy), default spread (defs), term spread (tes), 3-month t-bill rate (3mtb), and three lags of the dependent variable. Inference is based on Newey-West Standard Errors. $^{*}$, $^{**}$, $^{***}$ indicates significance at the 10%, 5% and 1% level, respectively.

| | Control Model EPS 3 months | Control Model EPS 6 months | Control Model EPS 12 | Control Model EPS > 12 months |
|---|---|---|---|---|
| Horizon | $\beta_1$ | $\beta_1$ | $\beta_1$ | $\beta_1$ |
| 1 | 0.0004 | -0.0002 | 0.0003 | 0.0007 |
| 2 | 0.0034 | -0.0001 | -0.0002 | 0.0000 |
| 3 | 0.0054 | 0.0003 | -0.0005 | -0.0003 |
| 4 | 0.0040 | 0.0003 | -0.0005 | -0.0002 |
| 5 | -0.0008 | 0.0004 | -0.0007 | -0.0006 |
| 6 | 0.0001 | 0.0002 | -0.0028 | -0.0008 |
| 7 | -0.0044 | 0.0008 | -0.0031 | -0.0015 |
| 8 | 0.0024 | 0.0005 | -0.0028 | -0.0007 |
| 9 | -0.0019 | 0.0006 | -0.0003 | -0.0014 |
| 10 | -0.0006 | -0.0002 | -0.0004 | -0.0012 |
| 11 | -0.0048 | -0.0002 | 0.0001 | -0.0027 |
| 12 | 0.0018 | -0.0010 | 0.0011 | -0.0031 |

We do not find any predictive power for the CFNAI in companies' aggregated earnings per share forecasts. This result is in line with the disadvantages of earnings forecasts discussed in Section 4.2. Most prominently, recommendations contain a larger information set than just an one period earnings forecasts. They basically include earnings forecasts for several years as well as interest rate assumptions and risk premium forecasts. Additionally, earnings forecasts might be based on stale information. Consequently, our results lend support to the notion that more forward-looking information, as included in recommendation changes, is required if one attempts to forecast future macroeconomic developments.

## 4.5 Conclusion

Combining individual company outlooks and macroeconomic developments is a new approach to exploit publicly available, but neglected, information. Our results show that the aggregation of these outlooks has predictive power for future macroeconomic developments of about one year. Due to their strong performance, changing company outlooks might be a promising predictor for different applications, especially in the area of now-casting and mid-term forecasting, e.g., for economic agencies. As shown, our measure contains information content not inherent in well-established macroeconomic predictors. Therefore it seems mandatory to use company outlooks as a control variable in models which could be influenced by macroeconomic developments.

Our results also provide a potential explanation for the old Wall Street saying that the stock market leads the economy. Recommendation changes basically mirror expectation changes of well informed market participants. We document that these changes lead overall macroeconomic conditions. Assuming that the stock market incorporates these new expectations, the overall stock market must lead the real economy. Thus far there is no study showing this direct link.

# References

Agarwal, V., and Hess, D. (2012) Common Factors in Analysts' Earnings Revisions: The Role of Changing Economic Conditions, unpublished working paper, Georgia State University, University of Cologne.

Aggarwal, R., Mohanty, S., and Song, F. (1995) Are Survey Forecasts of Macroeconomic Variables Rational?, *Journal of Business* **68**, 99-119.

Albuquerque, R., and Vega, C. (2009) Economic News and International Stock Market Co-movement, *Review of Finance* **13**, 401-465.

Almeida, A., Goodhart, C., and Payne, R. (1998) The Effects of Macroeconomic News on High Frequency Exchange Rate Behavior, *Journal of Financial and Quantitative Analysis* **33**, 383-408.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Vega, C. (2003) Micro Effects of Macro Announcements: Real-Time Price Discovery in Foreign Exchange, *American Economic Review* **93**, 38-62.

Ariely, D., Loewenstein, G., and Prelec, D. (2003) Coherent Arbitrariness: Stable Demand Curves Without Stable Preferences, *The Quarterly Journal of Economics* **118**, 73-105.

Ashiya, M. (2006) Forecast accuracy and product differentiation of Japanese institutional forecasters, *International Journal of Forecasting* **22**, 395–401.

Balduzzi, P., Elton, E. J., and Green, T. C. (2001) Economic News and Bond Prices: Evidence from the U.S. Treasury Market, *Journal of Financial and Quantitative Analysis* **36**, 523-543.

Barber, B., Lehavy, R., McNichols, M., and Trueman, B. (2001) Can Investors Profit from the Prophets? Security Analyst Recommendations and Stock Returns, Journal of Finance **56**, 531–63.

Baxter, M., and King, R.G. (1999) Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series, The Review of Economics and Statistics **81**, 575–593.

Beber, A., and Brandt, M. W. (2008) Resolving Macroeconomic Uncertainty in Stock and Bond Markets, *Review of Finance* **13**, 1-45.

Beber, A., and Brandt, M. W. (2010) When It Cannot Get Better or Worse: The Asymmetric Impact of Good and Bad News on Bond Returns in Expansions and Recessions, *Review of Finance* **14**, 119-155.

Bernanke, B. S., and Kuttner, K. N. (2005) What Explains the Stock Market's Reaction to Federal Reserve Policy?, *Journal of Finance* **60**, 1221-1257.

Bestelmeyer, G., Dicke, K., and Hess, D. (2012) Implied Risk Premium and the Business Cycle, unpublished working paper, University of Cologne.

Boyd, J. H., Hu, J., and Jagannathan, R. (2005) The stock market's reaction to unemployment news: Why bad news is usually good for stocks, *Journal of Finance* **60 (2)**, 649-672.

Brandt, M. W., and Kavajecz, K. A. (2004) Price Discovery in the U.S. Treasury Market: The Impact of Orderflow and Liquidity on the Yield Curve, *Journal of Finance* **59**, 2623-2654.

Brenner, M., Pasquariello, P., and Subrahmanyam, M. (2009) On the Volatility and Comovement of U.S. Financial Markets around Macroeconomic News Announcements, *Journal of Financial and Quantitative Analysis* **44**, 1265-1289.

Brown, L. D. (1993) Earnings Forecasting Research: Its Implications for Capital Markets Research, *International Journal of Forecasting* **9**, 295-320.

Brown, L. D., and Mohammad, E. (2010) Is Analyst Earnings Forecast Ability Only Firm Specific?, *Contemporary Accounting Research* **27**, 727-750.

Brown, L. D., Gay, G. D., and Turac M. (2008) Creating a "Smart" Conditional Consensus Forecast, *Financial Analysts Journal* **64**, 74-86.

Brown, L. D., Hagerman, R. L., Griffin, P. A., and Zmijewski, M. E. (1987) Security Analysts Superiority Relative to Univariate Time series Models in Forecasting Quarterly Earnings, *Journal of Accounting and Economics* **9**, 61-87.

Campbell, S. D., and Sharpe, S. A. (2009) Anchoring Bias in Consensus Forecasts and Its Effect on Market Prices, *Journal of Financial and Quantitative Analysis* **44**, 369-390.

Cenesizoglu, T. (2011) Size, Book-to-Market Ratio and Macroeconomic News, *Journal of Empirical Finance* **18**, 248-270.

Clement, M. B. (1999) Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter?, *Journal of Accounting and Economics* **27**, 285-303.

Clement, M. B., and Tse, S. Y. (2005) Financial Analyst Characteristics and Herding Behavior in Forecasting, *The Journal of Finance* **60**, 307-341.

Clements, M. P., and Krolzig, H.-M. (1998) A Comparison of the Forecast Performance of Markov-switching and Threshold Autoregressive models of US GNP, *Econometrics Journal* **1**, C47-C75.

Croushore, D., and Stark, T. (2001) A Real-Time Data Set for Macroeconomists, *Journal of Econometrics* **105**, 111-130.

Davis, S. J. (2006) Publicly Traded Versus Privately Held Firms. Capital Ideas, Chicago GSB: http://www.chicagobooth.edu/capideas/sep06/4.aspx

De Goeij, P. C., Hu, J., and Werker, B. J. M. (2009) The Price of Macroeconomic Announcement News, unpublished working paper, Tilburg University.

Dicke, K., and Hess, D. (2012) Information Content Drives Risk Premiums of Macroeconomic News on Bond and Stock Markets, unpublished working paper, University of Cologne.

Diebold, F. X., and Mariano, R. (1995) Comparing Predictive Accuracy, *Journal of Business and Economic Statistics* **13**, 253-265.

Dumont, M., Rayp, G., Thas, O., and Willemé, P. (2005) Correcting Standard Errors in Two-stage Estimation Procedures with Generated Regressands, *Oxford Bulletin of Economics and Statistics* **67**, 421-433.

Elton, E. J. (1999) Expected Return, Realized Return, and Asset Pricing Tests, *Journal of Finance* **54**, 1199-1220.

Evans, C. L., Liu,C. T., and Pham-Kanter, G. (2002) The 2001 recession and the Chicago Fed National Activity Index: Identifying business turning points, Economic Perspectives, Federal Reserve Bank of Chicago **Issue Q III**, 26-43.

Evans, M. D. D., and Lyons, R. K. (2008) How is macro news transmitted to exchange rates?, *Journal of Financial Economics* **88**, 26–50.

Federal Reserve Bank of Chicago (2012): www.chicagofed.org/webpages/research/data/cfnai/current_data.cfm

Federal Reserve Bank of Chicago (2012b) http://www.chicagofed.org/digital_assets/publications/cfnai/background/cfnai_background.pdf

Federal Reserve Bank of St. Louis (2011): http://research.stlouisfed.org/

Flannery, M. J., and Protopapadakis, A. A. (2002) Macroeconomic Factors Do Influence Aggregate Stock Returns, *Review of Financial Studies* **15**, 751-782.

French, Kenneth (2011): http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/

Fried, D., and Givoly, D. (1982) Financial Analysts' Forecasts of Earnings – A Better Surrogate for Market Expectations, *Journal of Accounting and Economics* **4**, 85-107.

Gilbert, T. (2011): Information aggregation around macroeconomic announcements: Revisions matter, *Journal of Financial Economics* **101**, 114-131.

Gilbert, T., Scotti, C., Strasser, G., and Vega, C. (2010) Why Do Certain Macroeconomic News Announcements Have a Big Impact on Asset Prices? Unpublished working paper, University of Washington, Federal Reserve Washington D.C., Boston College, Federal Reserve Washington D.C..

Green, T. C. (2004) Economic News and the Impact of Trading on Bond Prices, *Journal of Finance* **59**, 1201-1233.

Grossman, S. J., and Stiglitz, J. E. (1980) On the Impossibility of Informationally Efficient Markets, *American Economic Review* **70**, 393-408.

Gürkaynak, R. S., Sack, B., and Swanson, E. (2005) The Sensitivity of Long-Term Interest Rates to Economic News: Evidence and Implications for Macroeconomic Models, *American Economic Review* **95**, 425-436.

Harvey, D., Leybourne, S., and Newbold, P. (1997) Testing the equality of prediction mean squared errors, *International Journal of Forecasting* **13**, 281-291.

Hautsch, N., and Hess, D. (2007) Bayesian Learning in Financial Markets – Testing for the Relevance of Information Precision, *Journal of Financial and Quantitative Analysis* **42**, 189-208.

Hautsch, N., Hess, D., and Müller, C. (2012) Price Adjustment to News with Uncertain Precision, *Journal of International Money and Finance* **31**, 337-355.

Hess, D., and Orbe, S. (2013) Irrationality or Efficiency of Macroeconomic Survey Forecasts? Implications from the Anchoring Bias Test, *Review of Finance* Advance Access published January 12, 2013, doi:10.1093/rof/rfs037.

Howe, J. S., Unlu, E., and Yan, X. (2009) The Predictive Content of Aggregate Analyst Recommendations, *Journal of Accounting Research* **47**, 799-821.

Jegadeesh, N., and Kim, W. (2010) Do analysts herd? An analysis of recommendations and market reactions, *Review of Financial Studies* **32**, 901-937.

Jegadeesh, N., Kim, J., Krische, S. D., and Lee, C. M. C. (2004) Analyzing the Analysts: When Do Recommendations Add Value?, *Journal of Finance* **59**, 1083-1124.

Kolb, R.A., and Stekler, H.O. (1996) How well do analysts forecast interest rates?, *Journal of Forecasting* **15**, 385–394.

Marcellino, M. (2008) A Linear Benchmark for Forecasting GDP Growth and Inflation, *Journal of Forecasting* **27**, 305-340.

McNees, S.K. (1987) Consensus forecasts: tyranny of the majority?, *New England Economic Review* **Nov/Dec**, 15–21.

McQueen, G., and Roley, V. V. (1993) Stock Prices, News, and Business Conditions, *Review of Financial Studies* **6**, 683-707.

Mullineaux, D. J. (1978) On Testing for Rationality: Another Look at the Livingston Price Expectations Data, *Journal of Political Economy* **86**, 329-336.

Muth, J. F. (1961) Rational Expectations and the Theory of Price Movements, *Econometrica* **29**, 315-335.

Pasquariello, P., and Vega, C. (2007) Informed and Strategic Order Flow in the Bond Markets, *Review of Financial Studies* **20**, 1975–2019.

Pasquariello, P., and Vega, C. (2009) The On-The-Run Liquidity Phenomenon, *Journal of Financial Economics* **92**, 1-24.

Pearce, D. K., and Roley, V. V. (1985) Stock Prices and Economic News, *Journal of Business* **58**, 49-67.

Pesando, J. E. (1975) A Note on the Rationality of the Livingston Price Expectations, *Journal of Political Economy* **83**, 849-858.

Savor, P., and Wilson, M. (2012) How Much Do Investors Care About Macroeconomic Risk? Evidence From Scheduled Economic Announcements, *Journal of Financial and Quantitative Analysis*, forthcoming.

Schirm, D. C. (2003) A Comparative Analysis of the Rationality of Consensus Forecasts of U.S. Economic Indicators, *Journal of Business* **76**, 547-561.

Shiller, R. (2011): http://www.econ.yale.edu/~shiller/

Sinha, P., Brown, L. D., and Daas, S. (1997) A Re-Examination of Financial Analysts' Differential Earnings Forecast Accuracy, *Contemporary Accounting Research* **14**, 1-42.

Stekler, H.O. (1987) Who forecasts better?, *Journal of Business and Economics Statistics* **5**, 155–158.

Stock, J. H., and Watson, M. W. (1998) Business Cycle Fluctuations in U.S. Macroeconomic Time Series, NBER Working Paper No. 6528.

Stock, J. H., and Watson, M. W. (1999). A Comparison of Linear and Nonlinear Univariate Models for Forecasting Macroeconomic Time Series, in: R. F. Engle and H. White (eds.), *Cointegration, Causality, and Forecasting. A Festschrift in Honour of Clive W.J. Granger*, Oxford: Oxford University Press, 1-44.

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

Swanson, N. R. and White, H. (1997) A Model Selection Approach To Real-Time Macroeconomic Forecasting Using Linear Models And Artificial Neural Networks, *Review of Economics and Statistics* **79**, 540-550.

Tversky, A., and Kahneman, D. (1974) Judgment under Uncertainty: Heuristics and Biases, *Science* **185**, 1124-1131.

Urich, T., and Wachtel, P. (1984) The Effects of Inflation and Money Supply Announcements on Interest Rates, *Journal of Finance* **39**, 1177-1188.

Womack, K. (1996) Do Brokerage Analysts' Recommendations Have Investment Value? *Journal of Finance* **51**, 137–167.

Zarnowitz, V. (1984) The accuracy of individual and group forecasts from business outlook surveys, *Journal of Forecasting* **3**, 11-26.

Zarnowitz, V. and Braun, P. (1993) Twenty-Two Years of the NBER-ASA Quarterly Economic Outlook Surveys: Aspects and Comparisons of Forecasting Performance, In: Stock, J.H., Watson, M.W. (Eds.), Business Cycles, Indicators, and Forecasting. NBER Studies in Business Cycles, 28, University of Chicago, Chicago.

# Appendix to Chapter 2

## Appendix 2 A

The direct estimation of

$$F_{t-\tau}^{t} = \lambda \cdot E\left[A_t \middle| I_{t-\tau}\right] + \left(1-\lambda\right) \cdot \overline{A_h} \tag{A.1}$$

is not possible. However the estimation becomes feasible by means of the well known definition of the unanticipated news component of a macroeconomic release:

$$S_t = A_t - F_t, \tag{A.2}$$

where $S_t$ denotes the unanticipated news component called surprise, $A_t$ the actual announced value of the macroeconomic indicator and $F_t$ the survey based forecast. Taking the conditional expectation of Equation (A.2) and rearranging it leads to:

$$E\left[A_t \middle| I_{t-\tau}\right] = E\left[S_t \middle| I_{t-\tau}\right] + F_{t-\tau}^{t}. \tag{A.3}$$

Substituting $E\left[A_t \middle| I_{t-\tau}\right]$ in (A.1) with (A.3) gives the model for the further investigation:

$$E\left[S_t \middle| I_{t-\tau}\right] = \frac{1-\lambda}{\lambda} \cdot F_{t-\tau}^{t} - \frac{1-\lambda}{\lambda} \cdot \overline{A_h}\,. \tag{A.4}$$

For reasons of clarity we define the slope coefficient in our model as:

$$\gamma \equiv \frac{\left(1-\lambda\right)}{\lambda}\;. \tag{A.5}$$

Therefore the regression model for the test of the anchoring bias is given by[43]:

$$S_t = \gamma \cdot \left(F_t - \bar{A_h}\right) + \eta_t. \tag{A.6}$$

[43] Although Equation (A.6) does not include a constant we always include one in the estimation.

## Appendix 2 B

Assume that $A_t$ follows an ARIMA(p,d,q) process without a constant term, or equivalently, the first difference of $A_t$ follows an ARMA(p,q) process:

$$\begin{aligned}\Delta A_t &= b_1 \cdot \Delta A_{t-1} + b_2 \cdot \Delta A_{t-2} + ... + b_p \cdot \Delta A_{t-p} \\ &\quad + \varepsilon_t + c_1 \cdot \varepsilon_{t-1} + c_2 \cdot \varepsilon_{t-2} + ... + c_q \cdot \varepsilon_{t-q},\end{aligned}$$

with i.i.d. $\varepsilon_t \sim N\left(0,\sigma^2\right)$. At this point we maximally consider the first difference (d=1) because all integrated macroeconomic series are stationary after differencing once. Provided the process is stationary, it can be rewritten as:

$$\Delta A_t = \psi(L)\varepsilon_t \qquad \text{with } \psi(L) = \frac{1 + c_1 \cdot L + c_2 \cdot L^2 + ... + c_q \cdot L^q}{1 - b_1 \cdot L - b_2 \cdot L^2 - ... - b_p \cdot L^p}.$$

For example, for an ARIMA(1,1,1) or equivalently for an ARMA(1,1) of a differenced series we then get:

$$\Delta A_t = \varepsilon_t + \sum_{j=1}^{\infty} \beta_j \cdot \varepsilon_{t-j} \qquad\qquad \text{with } \beta_j = b_1^{\,j-1} \cdot c_1 + b_1^{\,j}.$$

Moreover suppose that analysts use a corresponding ARIMA(p,d,q) model to generate forecasts. However, suppose that analysts can obtain some additional information $Z_t$ useful to predict the innovation $\varepsilon_t$ in $A_t$, e.g., from the inspection of other macroeconomic announcements released earlier. Assume that $corr(\varepsilon_t, Z_t) \neq 0$ and $corr(\varepsilon_{t-j}, Z_t) = 0 \;\forall\; j \geq 1$. Then their forecasts may be written as:

$$F_t = A_{t-1} + \sum_{j=1}^{\infty} \hat{\beta}_j \cdot \varepsilon_{t-j} + Z_t.$$

Substituting the above MA($\infty$) representations of $A_t$ and $F_t$ into the anchoring regression:

$$A_t - F_t = \gamma \cdot \left( F_t - \frac{1}{h} \sum_{i=1}^{h} A_{t-i} \right) + \eta_t$$

yields

$$\begin{aligned}
&\left( A_{t-1} + \varepsilon_t + \sum_{j=1}^{\infty} \beta_j \varepsilon_{t-j} \right) - \left( A_{t-1} + \sum_{j=1}^{\infty} \hat{\beta}_j \cdot \varepsilon_{t-j} + Z_t \right) \\
&= \gamma \cdot \left( \left( A_{t-1} + \sum_{j=1}^{\infty} \hat{\beta}_j \cdot \varepsilon_{t-j} + Z_t \right) - \frac{1}{h} \sum_{i=1}^{h} \left( A_{t-1-i} + \varepsilon_{t-i} + \sum_{j=1}^{\infty} \beta_j \cdot \varepsilon_{t-j-i} \right) \right) + \eta_t
\end{aligned}$$

and after simplifying

$$\begin{aligned}
&\varepsilon_t + \left( \sum_{j=1}^{\infty} \left( \beta_j - \hat{\beta}_j \right) \cdot \varepsilon_{t-j} \right) - Z_t \\
&= \gamma \cdot \left( \left( A_{t-1} + \sum_{j=1}^{\infty} \hat{\beta}_j \cdot \varepsilon_{t-j} + Z_t \right) - \frac{1}{h} \sum_{i=1}^{h} \left( A_{t-1-i} + \varepsilon_{t-i} + \sum_{j=1}^{\infty} \beta_j \cdot \varepsilon_{t-j-i} \right) \right) + \eta_t .
\end{aligned}$$

Now we can rewrite the anchoring bias regression as:

$$\underbrace{\varepsilon_t + y'_{t-1} - Z_t}_{y_t} = \hat{\gamma} \cdot \underbrace{\left( x'_{t-1} + Z_t \right)}_{x_t} + \eta_t$$

$$\text{with } y'_{t-1} \equiv \sum_{j=1}^{\infty} \left( \beta_j - \hat{\beta}_j \right) \cdot \varepsilon_{t-j}$$

and[44]

$$x'_{t-1} \equiv A_{t-1} + \sum_{j=1}^{\infty} \hat{\beta}_j \cdot \varepsilon_{t-j} - \frac{1}{h}\sum_{i=1}^{h}\left( A_{t-1-i} + \varepsilon_{t-i} + \sum_{j=1}^{\infty} \beta_j \cdot \varepsilon_{t-j-i} \right)$$
$$= \sum_{j=1}^{\infty} \hat{\beta}_j \cdot \varepsilon_{t-j} + \frac{1}{h}\sum_{i=1}^{h}\sum_{j=1}^{h-i}\left( \varepsilon_{t-j} + \sum_{k=1}^{\infty} \beta_k \cdot \varepsilon_{t-j-k} \right).$$

Note that $y'_{t-1}$ and $x'_{t-1}$ collect past time series information, or more precisely, terms depending on past innovations $\varepsilon_t$ and (true and estimated) time series parameters ($\hat{\beta}_j$ and $\beta_j$). In contrast, $\varepsilon_t$ captures the innovations of the announcement process, i.e., the component of an announcement which is unpredictable on the basis of past time series information. $Z_t$ is similar to an innovation since it cannot be explained by past announcements. Hence, $Z_t$ reflects deviations of analysts' forecasts from purely time series based forecasts, or the influence of "additional information" (besides past announcements) on analyst' forecasts.

---

[44] Note that $x'_{t-1}$ is stationary. With $\Delta A_{t-i} = \varepsilon_{t-i} + \sum_{j=1}^{\infty} \beta_j \cdot \varepsilon_{t-j-i}$ we can rewrite

$$x'_{t-1} \equiv A_{t-1} + \sum_{j=1}^{\infty} \hat{\beta}_j \cdot \varepsilon_{t-j} - \frac{1}{h}\sum_{i=1}^{h}\left( A_{t-1-i} + \varepsilon_{t-i} + \sum_{j=1}^{\infty} \beta_j \cdot \varepsilon_{t-j-i} \right)$$

to

$$\begin{aligned}
x'_{t-1} &= A_{t-1} + \sum_{j=1}^{\infty} \hat{\beta}_j \cdot \varepsilon_{t-j} - \frac{1}{h}\sum_{i=1}^{h}\left( A_{t-1-i} + \Delta A_{t-i} \right) \\
&= A_{t-1} + \sum_{j=1}^{\infty} \hat{\beta}_j \cdot \varepsilon_{t-j} - \frac{1}{h}\sum_{i=1}^{h} A_{t-i} \\
&= \sum_{j=1}^{\infty} \hat{\beta}_j \cdot \varepsilon_{t-j} + \frac{1}{h}\sum_{i=1}^{h}\left( A_{t-1} - A_{t-i} \right) \\
&= \sum_{j=1}^{\infty} \hat{\beta}_j \cdot \varepsilon_{t-j} + \frac{1}{h}\sum_{i=1}^{h}\sum_{j=1}^{h-i} \Delta A_{t-j} \\
&= \sum_{j=1}^{\infty} \hat{\beta}_j \cdot \varepsilon_{t-j} + \frac{1}{h}\sum_{i=1}^{h}\sum_{j=1}^{h-i}\left( \varepsilon_{t-j} + \sum_{k=1}^{\infty} \beta_k \cdot \varepsilon_{t-j-k} \right)
\end{aligned}$$

The coefficient $\hat{\gamma}$ of the anchoring regression is given by:

$$\hat{\gamma} = \frac{Cov\left(x_t, y_t\right)}{Var\left(x_t\right)}$$

$$= \frac{Cov\left(\varepsilon_t + y'_{t-1} - Z_t, x'_{t-1} + Z_t\right)}{Var\left(x'_{t-1} + Z_t\right)}$$

$$= \frac{Cov\left(y'_{t-1}, x'_{t-1}\right) + Cov\left(\varepsilon_t, Z_t\right) - Var\left(Z_t\right)}{Var\left(x'_{t-1}\right) + Var\left(Z_t\right)},$$

where the last line exploits the fact that $Cov\left(Z_t, y'_{t-1}\right) = 0$, $Cov\left(Z_t, x'_{t-1}\right) = 0$, and $Cov\left(\varepsilon_t, x'_{t-1}\right) = 0$ by construction. We can split this expression into two parts by collecting all terms in the numerator depending on $x'_{t-1}$ and those depending on $Z_t$ :

$$\hat{\gamma}_1 = \frac{Cov\left(y'_{t-1}, x'_{t-1}\right)}{Var\left(x'_{t-1}\right) + Var\left(Z_t\right)}$$

$$\hat{\gamma}_2 = \frac{Cov\left(\varepsilon_t, Z_t\right) - Var\left(Z_t\right)}{Var\left(x'_{t-1}\right) + Var\left(Z_t\right)}.$$

The first component $\hat{\gamma}_1$ captures the influence of (possibly biased) parameters $\hat{\beta}_j$ , while the second component $\hat{\gamma}_2$ captures the influence of $Z_t$ . In the case of non-integrated time series, i.e., if $A_t$ follows an ARMA(p,q) process only the definition of $y'_{t-1}$ and $x'_{t-1}$ changes to:

$$y'_{t-1} \equiv \sum_{j=1}^{\infty}\left(\beta_j - \hat{\beta}_j\right) \cdot \varepsilon_{t-j}$$

and

$$x'_{t-1} \equiv \sum_{j=1}^{\infty} \hat{\beta}_j \cdot \varepsilon_{t-j} - \frac{1}{h} \sum_{i=1}^{h} \left( \varepsilon_{t-i} + \sum_{j=1}^{\infty} \beta_j \cdot \varepsilon_{t-j-i} \right).$$

## Appendix 2 C

Technically, our simulation procedure consists of the following steps:

(1) We start by simulating time series for a macroeconomic release $\tilde{A}_t$ using ARIMA(1,0,0), ARIMA(1,0,1) and ARIMA(1,1,1) models. For brevity, we describe only how we generate ARIMA(1,0,1) processes, i.e., $\tilde{A}_t = \alpha \cdot \tilde{A}_{t-1} + \beta \cdot \tilde{\varepsilon}_{t-1} + \tilde{\varepsilon}_t$. In addition, we also simulate a process for $\tilde{Z}_t$, assuming that $\tilde{Z}_t$ (as well as $\tilde{\varepsilon}_t$) is just a series of normally distributed random variables. While both $\tilde{Z}_t$ and $\tilde{\varepsilon}_t$ are not autocorrelated, we allow for a contemporaneous correlation among them, i.e., for different values of $\rho = \operatorname{corr}\left(\tilde{Z}_t, \tilde{\varepsilon}_t\right)$. Within each simulation run we produce a time series of 2,000 observations for $\tilde{Z}_t$ and $\tilde{A}_t$ (after a spin-in period of 2,000 observations).

(2) Then we perform two rolling estimations (using a rolling estimation window of 1,000 observations) to generate two series (each with $n$ = 1,000) of one-step-ahead out-of-sample predictions: First, we produce a time series of "simple" one-step-ahead forecasts estimating a rolling ARIMA(1,0,1) model,[45] i.e., $\tilde{A}_t = \alpha \cdot \tilde{A}_{t-1} + \beta \cdot \tilde{\varepsilon}_{t-1} + \tilde{\varepsilon}_t$. Second, we generate a series of "sophisticated" one-step-ahead forecasts based on a rolling ARIMAX estimation, i.e., $\tilde{A}_t = \alpha \cdot \tilde{A}_{t-1} + \beta \cdot \tilde{\varepsilon}_{t-1} + \phi \tilde{Z}_t + \tilde{\varepsilon}_t$.

(3) On these two out-of-sample forecast series ($n$ = 1,000) we then perform anchoring tests for $\overline{h}$ = 1, 2 and 3. We retain the test results for the optimal $\overline{h}$, i.e., for which the anchoring regression performed best according to the BIC.

[45] For the other specifications as mentioned in step (1) we estimate the corresponding ARIMA model.

Steps (1) to (3) are repeated 5,000 times for different parameter combinations, i.e., for $\alpha = 0.2$, 0.4, 0.6, and 0.8, $\beta = 0.2$, and $\rho = 0.2$, 0.4, 0.6, and 0.8.

# Appendix to Chapter 3

## Appendix 3 A

Table Appendix 3 A Summary Statistics for Predictability Quartiles

Panel A to D report summary statistics for the set of used variables in descending predictability order. AFErr denotes the scaled absolute forecast error, L.AFErr the absolute forecast error from the previous period, SpExp specific experience, GenExp general experience, GAbil the general ability measure, CovInd the number of covered indicators, and Horizon the forecast horizon.

Panel A: First Predictability Quartile (Highest Predictability)

| | AFErr | L.AFErr | SpExp | GenExp | GAbil | CovInd | Horizon |
|---|---|---|---|---|---|---|---|
| Mean | 1.49 | 1.39 | 43.49 | 55.95 | 0.61 | 18.58 | 8.73 |
| Median | 1.32 | 1.25 | 34.00 | 49.00 | 0.62 | 20.00 | 5.00 |
| 10$^{th}$ Percentile | 0.00 | 0.00 | 5.00 | 8.00 | 0.50 | 12.00 | 1.00 |
| 90$^{th}$ Percentile | 3.10 | 2.93 | 97.00 | 118.00 | 0.72 | 22.00 | 10.00 |

Panel B: Second Predictability Quartile

| | AFErr | L.AFErr | SpExp | GenExp | GAbil | CovInd | Horizon |
|---|---|---|---|---|---|---|---|
| Mean | 1.86 | 1.63 | 42.71 | 57.10 | 0.62 | 18.61 | 7.33 |
| Median | 1.58 | 1.40 | 34.00 | 50.00 | 0.62 | 20.00 | 5.00 |
| 10$^{th}$ Percentile | 0.00 | 0.00 | 5.00 | 9.00 | 0.50 | 13.00 | 1.00 |
| 90$^{th}$ Percentile | 3.77 | 3.33 | 94.00 | 118.00 | 0.72 | 22.00 | 10.00 |

Panel C: Third Predictability Quartile

| | AFErr | L.AFErr | SpExp | GenExp | GAbil | CovInd | Horizon |
|---|---|---|---|---|---|---|---|
| Mean | 2.01 | 1.70 | 44.30 | 55.47 | 0.61 | 18.27 | 8.07 |
| Median | 1.88 | 1.69 | 35.00 | 48.00 | 0.62 | 19.00 | 6.00 |
| 10$^{th}$ Percentile | 0.00 | 0.00 | 5.00 | 8.00 | 0.49 | 12.00 | 2.00 |
| 90$^{th}$ Percentile | 4.12 | 3.46 | 98.00 | 117.00 | 0.72 | 22.00 | 11.00 |

Panel D: Fourth Predictability Quartile (Lowest Predictability)

| | AFErr | L.AFErr | SpExp | GenExp | GAbil | CovInd | Horizon |
|---|---|---|---|---|---|---|---|
| Mean | 2.43 | 2.01 | 43.83 | 55.54 | 0.62 | 18.07 | 5.37 |
| Median | 2.11 | 1.83 | 35.00 | 48.00 | 0.63 | 20.00 | 4.00 |
| 10$^{th}$ Percentile | 0.43 | 0.34 | 5.00 | 9.00 | 0.50 | 12.00 | 1.00 |
| 90$^{th}$ Percentile | 4.80 | 3.88 | 98.00 | 117.00 | 0.73 | 22.00 | 11.00 |

# Appendix to Chapter 4

## Appendix 4 A: Stationarity Tests of the CFNAI Change and its Components

Table Appendix 4 A Stationarity Tests of the CFNAI Change and its Components

This table reports Phillips-Perron $\tau -$ test statistics using the corresponding MacKinnon approximated p-values for the forward looking CFNAI change (CFNAI F), historical CFNAI changes (CFNAI H), and the changes of the respective CFNAI components Production and Income (PI), Employment, Unemployment, and Hours (EUH), Personal Consumption and Housing (CH), and Sales, Orders, and Inventories (SOI). $^{*}$, $^{**}$, $^{***}$ indicates significance at the 10%, 5% and 1% level, respectively.

| Horizon | CFNAI F | CFNAI H | CFNAI PI | CFNAI EUH | CFNAI CH | CFNAI SOI |
|---|---|---|---|---|---|---|
| 1 | -12.601$^{***}$ | -12.601$^{***}$ | -13.499$^{***}$ | -11.267$^{***}$ | -10.983$^{***}$ | -16.450$^{***}$ |
| 2 | -6.178$^{***}$ | -6.178$^{***}$ | -7.146$^{***}$ | -5.325$^{***}$ | -5.618$^{***}$ | -7.614$^{***}$ |
| 3 | -5.922$^{***}$ | -5.922$^{***}$ | -6.921$^{***}$ | -5.586$^{***}$ | -5.535$^{***}$ | -7.785$^{***}$ |
| 4 | -4.786$^{***}$ | -4.786$^{***}$ | -5.282$^{***}$ | -4.736$^{***}$ | -4.684$^{***}$ | -6.454$^{***}$ |
| 5 | -4.597$^{***}$ | -4.597$^{***}$ | -5.227$^{***}$ | -4.620$^{***}$ | -4.450$^{***}$ | -5.760$^{***}$ |
| 6 | -4.266$^{***}$ | -4.266$^{***}$ | -5.047$^{***}$ | -4.195$^{***}$ | -4.084$^{***}$ | -5.953$^{***}$ |
| 7 | -3.948$^{***}$ | -3.948$^{***}$ | -4.611$^{***}$ | -3.952$^{***}$ | -3.711$^{***}$ | -5.167$^{***}$ |
| 8 | -3.447$^{***}$ | -3.447$^{***}$ | -4.247$^{***}$ | -3.662$^{***}$ | -3.411$^{**}$ | -4.934$^{***}$ |
| 9 | -3.044$^{**}$ | -3.044$^{**}$ | -4.022$^{***}$ | -3.512$^{***}$ | -3.134$^{**}$ | -4.677$^{***}$ |
| 10 | -2.663$^{*}$ | -2.663$^{*}$ | -4.044$^{***}$ | -3.379$^{**}$ | -2.904$^{***}$ | -4.320$^{***}$ |
| 11 | -2.633$^{*}$ | -2.633$^{*}$ | -3.882$^{***}$ | -3.240$^{**}$ | -2.648$^{*}$ | -4.488$^{***}$ |
| 12 | -2.205 | -2.205 | -3.828$^{***}$ | -3.146$^{**}$ | -2.477 | -4.074$^{***}$ |

## Appendix 4 B: Correlation Matrix

Table Appendix 4 B Correlation Matrix

This table reports the correlations coefficients between the 12 forward looking CFNAI changes (CFNAI1 to CFNAI12), the recommendation change (Δrec), the change in the leading economic indicator (Δlei), market excess returns (mex), dividend yield (divy), default spread (defs), term spread (tes), and the 3-month T-Bill rate (3mtb).

| | CFNAI | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | Δrec | Δlei | mex | divy | defs | 3mtb | tes |
| CFNAI1 | 1.00 | | | | | | | | | | | | | | | | | | |
| CFNAI2 | 0.72 | 1.00 | | | | | | | | | | | | | | | | | |
| CFNAI3 | 0.65 | 0.85 | 1.00 | | | | | | | | | | | | | | | | |
| CFNAI4 | 0.45 | 0.72 | 0.88 | 1.00 | | | | | | | | | | | | | | | |
| CFNAI5 | 0.44 | 0.59 | 0.78 | 0.91 | 1.00 | | | | | | | | | | | | | | |
| CFNAI6 | 0.39 | 0.55 | 0.67 | 0.83 | 0.93 | 1.00 | | | | | | | | | | | | | |
| CFNAI7 | 0.35 | 0.50 | 0.63 | 0.73 | 0.86 | 0.94 | 1.00 | | | | | | | | | | | | |
| CFNAI8 | 0.27 | 0.43 | 0.56 | 0.68 | 0.77 | 0.87 | 0.94 | 1.00 | | | | | | | | | | | |
| CFNAI9 | 0.23 | 0.34 | 0.49 | 0.60 | 0.71 | 0.79 | 0.89 | 0.95 | 1.00 | | | | | | | | | | |
| CFNAI10 | 0.22 | 0.31 | 0.42 | 0.54 | 0.65 | 0.74 | 0.81 | 0.90 | 0.95 | 1.00 | | | | | | | | | |
| CFNAI11 | 0.20 | 0.29 | 0.39 | 0.48 | 0.59 | 0.69 | 0.77 | 0.83 | 0.91 | 0.96 | 1.00 | | | | | | | | |
| CFNAI12 | 0.19 | 0.28 | 0.37 | 0.45 | 0.54 | 0.64 | 0.72 | 0.79 | 0.85 | 0.92 | 0.96 | 1.00 | | | | | | | |
| Δrec | 0.22 | 0.20 | 0.13 | 0.07 | 0.08 | 0.08 | 0.09 | 0.05 | 0.05 | 0.05 | 0.07 | 0.04 | 1.00 | | | | | | |
| Δlei | 0.32 | 0.48 | 0.34 | 0.31 | 0.22 | 0.23 | 0.21 | 0.19 | 0.14 | 0.12 | 0.08 | 0.08 | 0.08 | 1.00 | | | | | |
| mex | 0.16 | 0.23 | 0.34 | 0.30 | 0.21 | 0.12 | 0.10 | 0.10 | 0.05 | 0.02 | -0.02 | -0.03 | -0.12 | 0.28 | 1.00 | | | | |
| divy | -0.08 | -0.10 | -0.07 | -0.01 | 0.06 | 0.11 | 0.14 | 0.17 | 0.20 | 0.22 | 0.24 | 0.26 | 0.01 | -0.15 | -0.08 | 1.00 | | | |
| defs | -0.08 | -0.07 | -0.01 | 0.07 | 0.15 | 0.22 | 0.26 | 0.31 | 0.36 | 0.41 | 0.44 | 0.47 | -0.08 | -0.20 | -0.27 | 0.17 | 1.00 | | |
| 3mtb | -0.04 | -0.07 | -0.10 | -0.14 | -0.17 | -0.20 | -0.22 | -0.25 | -0.27 | -0.29 | -0.31 | -0.33 | 0.03 | -0.09 | 0.13 | -0.15 | -0.68 | 1.00 | |
| tes | 0.08 | 0.09 | 0.09 | 0.10 | 0.12 | 0.13 | 0.15 | 0.18 | 0.20 | 0.21 | 0.22 | 0.25 | -0.01 | 0.25 | -0.08 | 0.31 | 0.39 | -0.76 | 1.00 |